

**CARLOS EDUARDO VIEIRA RAMOS**

**O DIREITO DAS PLATAFORMAS**

Procedimento, legitimidade e constitucionalização na regulação  
privada da liberdade de expressão na internet

Dissertação de Mestrado

Orientador: Professor Dr. Orlando Villas Bôas Filho

**UNIVERSIDADE DE SÃO PAULO**  
**FACULDADE DE DIREITO**

SÃO PAULO – SP

2020



**CARLOS EDUARDO VIEIRA RAMOS**

**O DIREITO DAS PLATAFORMAS**

Procedimento, legitimidade e constitucionalização na regulação  
privada da liberdade de expressão na internet

Dissertação apresentada à Banca Examinadora do Programa de Pós-Graduação da Faculdade de Direito da Universidade de São Paulo como exigência parcial para obtenção do título de Mestre em Direito, na área de concentração Filosofia e Teoria Geral de Direito, sob orientação do Professor Dr. Orlando Villas Bôas Filho.

**UNIVERSIDADE DE SÃO PAULO**  
**FACULDADE DE DIREITO**

SÃO PAULO – SP

2020

Catálogo da Publicação  
Serviço de Biblioteca e Documentação  
Faculdade de Direito da Universidade de São Paulo

---

Vieira Ramos, Carlos Eduardo

O Direito das Plataformas: Procedimento, legitimidade e constitucionalização na regulação privada da liberdade de expressão na internet ; Carlos Eduardo Vieira Ramos ; orientador Orlando Villas Bôas Filho -- São Paulo, 2020.

465 pp.

Dissertação (Mestrado - Programa de Pós-Graduação em Filosofia do Direito e Teoria Geral do Direito) - Faculdade de Direito, Universidade de São Paulo, 2020.

1. Moderação de Conteúdo. 2. Plataformas da Internet. 3. Liberdade de Expressão. 4. Sociologia Jurídica. 5. Niklas Luhmann. I. Villas Bôas Filho, Orlando , orient. II. Título.

---

**CARLOS EDUARDO VIEIRA RAMOS**

**O Direito das Plataformas: Procedimento, legitimidade e constitucionalização na  
regulação privada da liberdade de expressão na internet**

Dissertação apresentada à Banca Examinadora do Programa de Pós-Graduação da Faculdade de Direito da Universidade de São Paulo como exigência parcial para obtenção do título de Mestre em Direito, na área de concentração Filosofia e Teoria Geral de Direito, sob orientação do Professor Dr. Orlando Villas Bôas Filho.

Data de aprovação: \_\_\_\_\_

**Banca Examinadora**

Prof. (a) Dr. (a): \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. (a) Dr. (a): \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. (a) Dr. (a): \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. (a) Dr. (a): \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura: \_\_\_\_\_

**São Paulo – SP**

## AGRADECIMENTOS

Esta dissertação e toda a minha formação universitária não seriam possíveis sem a universidade pública de qualidade, da qual tive a oportunidade de fazer parte tanto nos meus estudos de graduação, quanto nos de pós-graduação. A ela, devo minha formação – e, por isso, uma espécie de “*agradecimento zero*” por este trabalho.

Agradeço ao Professor Orlando Villas Bôas Filho, orientador deste trabalho – e de boa parte do meu percurso, desde a graduação, em que me deu a primeira oportunidade para ser monitor em uma disciplina de sociologia jurídica. Por meio dela – e de outras tantas experiências que partiram daí – encontrei a verdadeira paixão de ensinar algo a alguém. Fica, por isso, o meu muito obrigado.

Aos Professores Ronaldo Porto Macedo Jr. e Celso Fernandes Campilongo, pelas observações na qualificação deste mestrado, que permitiram uma mudança de rumo na pesquisa, sem a qual o trabalho não seria possível. Ao Professor Ronaldo agradeço, em particular, pelas oportunidades e pela generosidade, tanto na graduação, quanto na pós-graduação, de participar dos grupos de estudo e das monitorias, que muito contribuíram para minha formação.

Agradeço, também, aos colegas de trabalho, do Tribunal Regional Federal da 3.<sup>a</sup> Região, em particular à sua Presidente, Desembargadora Federal Therezinha Cazerta, que compreendeu como ninguém mais teria compreendido a necessidade de ajustar os horários do seu assessor às demandas de um mestrado.

A todos os colegas de faculdade, de graduação e de pós-graduação, que contribuíram ao longo deste trabalho, lendo, opinando e fornecendo críticas àquilo que escrevi. Em particular, agradeço a Arthur Prado, Caio Ferrari, Luiz Fernando e Rodrigo Nitrini.

Por fim, o mais importante: agradeço, profundamente, a toda a minha família pelo apoio e paciência, em particular nos momentos mais difíceis. Não consigo pensar em como isso seria possível sem vocês. À minha mãe, Marilúcia; ao meu pai, José Maria. E a Maísa Verdugo, dedico, adaptadas, as palavras de Javier Torres Nafarrate em uma coletânea de aulas de Luhmann – de quem você tanto cansou de me ouvir falar: em nenhuma linha desta dissertação aparece seu nome, mas, sem você, não teria sido possível uma única destas palavras.

*“Technologists need to realize that when they are building a platform, they are also building a world. And policy-makers need to realize that technologists are capable of building a world”.*

**Bruce Schneier**

## RESUMO

Vieira Ramos, Carlos Eduardo. O Direito das Plataformas: Procedimento, legitimidade e constitucionalização na regulação privada da liberdade de expressão na internet. 465 pp. Mestrado – Faculdade de Direito, Universidade de São Paulo, São Paulo, 2020.

Esta é uma dissertação sobre o papel que os atores privados exercem na regulação da liberdade de expressão. Em específico, ela investiga o funcionamento da moderação de conteúdo das plataformas da internet, o mecanismo decisório estruturado pelas empresas que estão por trás de espaços como o Facebook, o YouTube e o Twitter para controlar aquilo que as pessoas podem dizer. Para fazer isso, o trabalho parte, no Capítulo 1, de um problema, baseado em duas constatações. Primeiro, a de que as transformações tecnológicas que fizeram com que as pessoas utilizassem redes sociais como o Facebook para se expressar não criaram apenas novos *espaços de manifestação* – as plataformas da internet –, mas também se traduziram em novos *espaços de decisão* – estruturas que, integradas às plataformas, são mantidas pelas empresas que as criaram para exercer poder sobre aquilo que as pessoas dizem ali. Segundo, a de que esse mecanismo privado de regulação da liberdade de expressão tem se transformado de um modo que desafia, com sucesso, a forma como ele é compreendido. Em particular, as plataformas da internet têm transitado de um modelo decisório antes caracterizado pela opacidade, pela ausência de participação dos seus usuários e pela irrecorribilidade de suas decisões, para outro, marcado por procedimentos transparentes, que viabilizam a participação dos usuários e que tornam possível questionar – até mesmo para instâncias externas e independentes – suas decisões. Que razões existiriam para que as plataformas, depois de desenvolverem um sistema próprio de moderação de conteúdo – construído para que tivessem o total controle sobre os espaços virtuais que administram – agora se empenhassem em transformá-lo, perdendo, com isso o domínio que têm sobre o que é dito na internet? Por que essas empresas estão abrindo mão do seu poder? A dissertação se propõe a responder a essas perguntas em dois passos: primeiro, no Capítulo 2, empreendendo uma investigação a respeito de como os mecanismos de moderação de conteúdo funcionam, sistematizando-os para lançar luz à forma como essas decisões são tomadas. Segundo, utilizando a sociologia jurídica – e, em particular, a Teoria dos Sistemas, justificada e reconstruída, respectivamente, nos Capítulos 3 e 4 – para entender as razões por trás desse movimento de transformação. Para isso, o trabalho acopla o funcionalismo estrutural de Niklas



Luhmann – capaz de identificar o papel que essa forma de tomar decisões tem em seu lugar de origem, o Estado – com o aporte feito à teoria por Gunther Teubner – que, reinterpretando o pluralismo jurídico sob uma perspectiva sistêmica, tornou a Teoria dos Sistemas capaz de analisar o direito também como um fenômeno não-estatal. Construída essa estrutura analítica, o Capítulo 5 da dissertação sustenta que a utilização de técnicas procedimentais estatais para tomar decisões nas plataformas tem dois significados: primeiro, essa é uma solução para um problema de falta de legitimidade decisória, porque procedimentos transparentes e participativos levam as pessoas – tanto nos Estados, quanto nas plataformas – a aceitarem as decisões, mesmo que lhes sejam desfavoráveis. Segundo, ela é parte de um movimento mais amplo, uma manifestação da constitucionalização do mundo digital, pela qual as plataformas da internet se mobilizam para adotar formas procedimentais capazes de imunizar as decisões da moderação de conteúdo da intervenção de outros atores institucionais que atuam regulando a liberdade de expressão, como o Estado. O trabalho conclui, portanto, que por trás dessa transformação inexplicada há um *tradeoff* oculto que a justifica: quanto maior a autonomia que as plataformas conferem à moderação de conteúdo, construindo procedimentos que a aproximam do modelo decisório do Estado, mais ela se constitui como um direito propriamente dito – o Direito das Plataformas – vocacionando-se a ser legítimo e independente em relação às influências tanto internas quanto externas à estrutura empresarial que o criou.

**Palavras-chave:** Moderação de Conteúdo. Plataformas da Internet. Liberdade de Expressão. Sociologia Jurídica. Niklas Luhmann.

## ABSTRACT

Vieira Ramos, Carlos Eduardo. The Law of the Platforms: Procedure, legitimacy and constitutionalization in the private regulation of the freedom of expression on the internet. 465 pp. Master's Degree – Faculty of Law, University of São Paulo, São Paulo, 2020.

This is a dissertation about the role that private actors play in regulating freedom of expression. Specifically, it investigates how content moderation – the decision-making mechanism structured by the companies behind spaces like Facebook, YouTube and Twitter to control what people can say – works on Internet platforms. In order to do this, it starts, in Chapter 1, with a problem based on two findings. First, it states that the technological transformations that made people use social networks like Facebook to express themselves not only created new *spaces for manifestation* – the internet platforms – but also new *spaces for decision-making* - structures that, integrated to the platforms, were created by the companies to exercise power over what people say. Second, that this private mechanism that regulates freedom of expression has been transformed in a way that successfully challenges our thoughts about it. In particular, Internet platforms have shifted from a decision-making model previously characterized by opacity, lack of user participation, and impossibility of appeal, to one marked by transparent procedures that enable users to participate and even to question these decisions. What reasons would exist for platforms, after developing their own content moderation system – designed so they would have full control over the spaces managed by them – now strive to transform it, thereby losing their dominance over them? Why are these companies giving up their power? This dissertation intends to answer these questions in two steps: first, in Chapter 2, by undertaking an investigation into how content moderation mechanisms work, systematizing them and shedding light on how these decisions are made. Second, using sociology of Law – and, in particular, Systems Theory, justified and reconstructed respectively in Chapters 3 and 4 – to understand the reasons behind this platform-driven movement of transformation. In order to do this, this work couples the structural functionalism of Niklas Luhmann – which is able to identify the role that this form of decision-making has in its place of origin, the State – with the contributions made to the theory by Gunther Teubner – who, reinterpreting the legal pluralism from a systemic perspective, made Systems Theory capable of analyzing law as a non-state phenomenon. Having built this analytical framework, Chapter 5 of the dissertation holds that the use of

state procedural techniques in platform decisions has two meanings. First, it is a solution for a problem of lack of decision-making legitimacy, as transparent and participatory procedures lead people – in States, as well as in platforms – to accept decisions, even if they are unfavorable to them. Second, this transition is part of a larger movement of constitutionalization of the digital world, by which internet platforms adopt procedural forms capable of immunizing their content moderation decisions from the intervention of other institutional actors that regulate freedom of expression. This work concludes that, behind this unexplained transformation, there is a tradeoff that justifies it: the greater the autonomy that platforms give to content moderation, the more it constitutes itself as a kind of Law – the Law of Platforms – becoming both legitimate and autonomous regarding the influences that are internal and external to the corporate structure that created it.

**Keywords:** Content Moderation. Internet platforms. Freedom of expression. Sociology of Law. Niklas Luhmann.

## LISTA DE FIGURAS

Figura 1 – Imagens publicadas por Lipman_____	16
Figura 2 – O Horror da Guerra_____	70
Figura 3 – Capa do jornal Aftenposten _____	72
Figura 4 – Publicações da Primeira-Ministra norueguesa no Facebook, antes e depois da ação da moderação de conteúdo da plataforma _____	74
Figura 5 – Mensagem no YouTube informando a remoção de um vídeo do Infowars ____	78
Figura 6 – Mensagens de bloqueio no Twitter das contas de Alex Jones e do Infowars__	79
Figura 7 – Exemplos de resultados do hashtag “holohoax”_____	87
Figura 8 – Páginas de PowerPoint do material utilizado para treinar os moderadores ____	90
Figura 9 – Páginas de PowerPoint do material utilizado para treinar os moderadores ____	91
Figura 10 – Modelo pluralista de regulação da liberdade de expressão _____	114
Figura 11 – “A lista do ódio” _____	149
Figura 12 – Publicação no Facebook _____	158
Figura 13 – Publicação no Twitter_____	159
Figura 14 – Publicação no YouTube _____	160
Figura 15 – Caixa de seleção e formulário para se reportar uma publicação no Facebook _____	173
Figura 16 – Caixa de seleção e formulário para se reportar uma publicação no Twitter_	174
Figura 17 – Caixa de seleção e formulário para se reportar uma publicação no YouTube _____	174
Figura 18 – Restrição de acesso por idade no YouTube _____	197
Figura 19 – Transparência e recorribilidade no YouTube _____	239
Figura 20 – Informações sobre o histórico de denúncias_____	240
Figura 21 – Procedimento de Apelação no Twitter _____	242
Figura 22 – Comitê de Supervisão para Decisões de Conteúdo _____	244

## LISTA DE QUADROS

Quadro 1 – Mudança social: infraestruturas de liberdade de expressão _____	51
Quadro 2 – Mudança regulatória: as regulações da liberdade de expressão _____	61
Quadro 3 – Mudança social e mudança regulatória _____	63
Quadro 4 – Narrativa comum _____	98
Quadro 5 – “ <i>Algumas diferenças esquemáticas entre a Web 1.0 e a Web 2.0</i> ” _____	129
Quadro 6 – A variabilidade normativa nas plataformas da internet _____	152
Quadro 7 – Fase n.º 1: A detecção de conteúdo no procedimento adjudicatório _____	179
Quadro 8 – Descentralização de poderes no Facebook _____	182
Quadro 9 – Fase n.º 2: A decisão quanto ao conteúdo no procedimento adjudicatório _	193
Quadro 10 – Fase n.º 3: A aplicação de sanções no procedimento adjudicatório _____	199
Quadro 11 – A moderação de conteúdo nas plataformas da internet _____	219
Quadro 12 – A moderação de conteúdo em transformação a partir dos Santa Clara	
Principles _____	248

## SUMÁRIO

<b>INTRODUÇÃO</b>	<b>15</b>
<b>PARTE I   POR UMA DESCRIÇÃO</b>	<b>40</b>
<b>CAPÍTULO 1 – A MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET COMO UMA FORMA DE REGULAÇÃO DA LIBERDADE DE EXPRESSÃO</b>	<b>40</b>
1. INTRODUÇÃO	40
2. DA MUDANÇA SOCIAL À MUDANÇA REGULATÓRIA: BALKIN E A REGULAÇÃO ESTATAL DA LIBERDADE DE EXPRESSÃO	44
3. A MODERAÇÃO DE CONTEÚDO NAS PLATAFORMAS DA INTERNET: UMA JUSTIFICATIVA	64
3.1. <i>Introdução</i>	64
3.2 <i>Três casos concretos</i>	68
3.2.1 <i>Introdução</i>	68
3.2.2 <i>Três Casos</i>	69
3.2.2.1 <i>Caso n.º 1: O Horror da Guerra</i>	69
3.2.2.2 <i>Caso n.º 2: Alex Jones e o Infowars</i>	76
3.2.2.3 <i>Caso n.º 3: Holocausto negado</i>	83
3.3 <i>A Moderação de conteúdo das plataformas da internet</i>	91
3.3.1 <i>Introdução</i>	91
3.3.2 <i>Por uma narrativa comum</i>	93
3.3.3 <i>O problema de Balkin</i>	99
3.3.4 <i>A defesa de Balkin</i>	105
3.3.5 <i>A solução de Balkin</i>	112
4. CONCLUSÃO	118
<b>CAPÍTULO 2 – DA OPACIDADE À CLAREZA: COMO FUNCIONA A MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET</b>	<b>121</b>
1. INTRODUÇÃO	121
2. A MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET: UMA ABORDAGEM SISTEMÁTICA	124
2.1 <i>Introdução</i>	124

2.2 <i>Da promessa de democratização à moderação de conteúdo: dois conceitos</i>	124
2.3 <i>A moderação de conteúdo como normatização</i>	145
2.3.1 <i>As regras</i>	145
2.3.2 <i>Os exemplos</i>	152
2.4 <i>A moderação de conteúdo como adjudicação</i>	157
2.4.1 <i>Introdução</i>	157
2.4.2 <i>As Fases da Moderação de Conteúdo</i>	161
2.4.2.1 <i>Fase n.º 1: Detectar</i>	161
2.4.2.2 <i>Fase n.º 2: Decidir</i>	180
2.4.2.3 <i>Fase n.º 3: Sancionar</i>	194
3. <b>DA PRIMEIRA EMENDA À GOVERNANÇA PRIVADA: DUAS LEITURAS SOBRE A MODERAÇÃO DE CONTEÚDO</b>	200
3.1 <i>Introdução</i>	200
3.2 <i>Duas correntes</i>	202
3.2.1 <i>A moderação de conteúdo como uma categoria jurisprudencial</i>	202
3.2.2 <i>A moderação de conteúdo como uma forma de governança privada</i>	211
4. <b>CONCLUSÃO</b>	216
<b>PARTE II   POR UMA REINTERPRETAÇÃO</b>	<b>224</b>
<b>CAPÍTULO 3 – PROBLEMATIZANDO A MODERAÇÃO DE CONTEÚDO</b>	<b>224</b>
1. <b>INTRODUÇÃO</b>	224
2. <b>PROBLEMA E HIPÓTESES</b>	226
3. <b>MÉTODO</b>	258
3.1 <i>Por que sociologia?</i>	258
3.2 <i>Por que Teoria dos Sistemas?</i>	263
4. <b>CONCLUSÃO</b>	267
<b>CAPÍTULO 4 – DO PROCEDIMENTO À LEGITIMIDADE</b>	<b>269</b>
1. <b>INTRODUÇÃO</b>	269
2. <b>LEGITIMAÇÃO PELO PROCEDIMENTO</b>	270
2.1 <i>Introdução</i>	270
2.2 <i>Problema e hipótese: a legitimação pelo procedimento na teoria</i>	271
2.2.1 <i>O problema</i>	271

2.2.2 A hipótese	279
2.3 <i>Três casos: a legitimação pelo procedimento na prática</i>	294
2.3.1 Os processos judiciais	294
2.3.2 Os processos eleitorais e legislativos	305
3. CONCLUSÃO	313
<b>CAPÍTULO 5 – DA LEGITIMIDADE ÀS PLATAFORMAS</b>	<b>319</b>
1. INTRODUÇÃO	319
2. POR UMA TEORIA DOS SISTEMAS DIFERENTE	322
2.1 <i>Introdução</i>	323
2.2 <i>Por um direito fora do Estado</i>	324
2.2.1 O problema de Teubner	324
2.2.2 A solução de Teubner	341
2.2.3 Os benefícios da solução de Teubner	349
3. POR UMA NOVA PERSPECTIVA QUANTO À MODERAÇÃO DE CONTEÚDO	368
3.1 <i>Introdução</i>	368
3.2 <i>O procedimento como legitimação</i>	371
3.3 <i>O procedimento como constitucionalização</i>	397
3.4 <i>Os próximos passos</i>	409
4. CONCLUSÃO	417
<b>CONCLUSÃO</b>	<b>421</b>
<b>REFERÊNCIAS</b>	<b>431</b>



## INTRODUÇÃO

Em maio de 2016, um grupo nacionalista israelense publicou no Facebook uma imagem em que fazia referência ao território em disputa na região, com um texto afirmando que “[o local] *é chamado de Israel, e não ‘Palestina’! Compartilhe para concordar*”<sup>1</sup>. A mensagem foi submetida à moderação de conteúdo da plataforma e removida porque ela não “*seguiria os Padrões de Comunidade do Facebook*”. Além disso, os administradores da página foram proibidos de fazer postagens por três dias e ordenados a retirar outras publicações consideradas “*problemáticas*”. Caso isso não acontecesse – e os moderadores encontrassem mais alguma mensagem que infringisse as regras – a página seria excluída.

Insatisfeitos com a decisão – mas sem ter como revertê-la –, os administradores do grupo pediram ajuda a Dov Lipman, um membro do parlamento israelense, que decidiu testar se as restrições aplicadas pelo Facebook àquilo que era dito na plataforma eram construídas de forma consistente. Ele criou uma página na rede social e postou uma imagem quase igual à anterior, trocando apenas os lugares das palavras “*Israel*” e “*Palestina*”. O resultado foi um texto que afirmava que “[o local] *é chamado de Palestina, e não ‘Israel’*. *Compartilhe para concordar*”. Em seguida, Lipman reportou a imagem à mesma moderação de conteúdo do Facebook – que decidiu que a postagem “*não violava os nossos Padrões de Comunidade*”.

---

<sup>1</sup> A retirada do conteúdo e seus desdobramentos foram amplamente noticiados à época; *cf.*, por exemplo, BEARAK, M. Facebook apologizes for taking down a pro-Israel post. **The Washington Post**, 15.6.2016, disponível em <[https://www.washingtonpost.com/news/worldviews/wp/2016/06/15/facebook-apologizes-for-taking-down-a-pro-israel-post/?noredirect=on&utm\\_term=.867baac48fec](https://www.washingtonpost.com/news/worldviews/wp/2016/06/15/facebook-apologizes-for-taking-down-a-pro-israel-post/?noredirect=on&utm_term=.867baac48fec)>; NEWMAN, L. Facebook apologizes for removing pro-Israel post. **The Jerusalem Post**, 15.6.2016, disponível em <<https://www.jpost.com/Diaspora/Facebook-apologizes-for-removing-pro-Israel-post-456774>>; a questão teve repercussão também em sites de inclinação nacionalista israelense; *cf.*, por exemplo, SPEYER, L. Sting Operation Reveals Facebook’s Double ‘Community Standards’ in Relation to Pro-Israel Posts. **The Algemeiner**, 15.6.2016, disponível em <<https://www.algemeiner.com/2016/06/15/sting-operation-reveals-facebooks-double-community-standards-in-relation-to-pro-israel-posts/>>, acessos em 30.7.2019.

**Figura 1 – Imagens publicadas por Lipman**



Fonte: The Jerusalem Post<sup>2</sup>

Após tentar, sem sucesso, entrar em contato com a plataforma, Lipman decidiu publicar uma carta aberta em um jornal local. No texto, ele reconhecia que “o Facebook está, sem dúvidas, em uma posição difícil, tentando balancear a liberdade de expressão enquanto também tenta prevenir a incitação à violência e outros resultados negativos das postagens das pessoas”, mas pedia explicações quanto às razões que fundamentaram a decisão de manter uma postagem e excluir a outra, porque “não é possível que se tenha um padrão para aqueles que apoiam o território que está sendo chamada de Israel, e um padrão diferente para aqueles que apoiam o território que está sendo chamado de Palestina”<sup>3</sup>. Depois da carta, o Facebook decidiu reestabelecer o conteúdo retirado, encaminhando à página uma mensagem padronizada de desculpas<sup>4</sup> – sem explicar, entretanto, o que exatamente tinha acontecido.

---

<sup>2</sup> LIPMAN, D. To whom it may concern – an open letter to Facebook. **The Jerusalem Post**, 9.6.2016, disponível em <<https://www.jpost.com/Opinion/To-whom-it-may-concern-an-open-letter-to-Facebook-456388>>, acesso em 30.7.2019.

<sup>3</sup> LIPMAN, D. To whom it may concern – an open letter to Facebook. **The Jerusalem Post**, 9.6.2016, disponível em <<https://www.jpost.com/Opinion/To-whom-it-may-concern-an-open-letter-to-Facebook-456388>>, acesso em 30.7.2019.

<sup>4</sup> Trata-se da mensagem genérica utilizada pela plataforma para essas situações (“A member of our team accidentally removed something you posted on Facebook. This was a mistake, and we sincerely apologize for this error. We’ve since restored the content, and you should now be able to see it.”). Cf., por exemplo, exatamente o mesmo conteúdo em um caso de 2013, descrito em WEMPLE, E. Facebook admits ‘mistake’ in blocking Fox News Radio commentator. **The Washington Post**, 29.6.2013, disponível em

O caso é importante por três razões. Primeiro, ele demonstra o poder que o Facebook e outras plataformas, como o Twitter e o YouTube – que enfrentaram situações semelhantes recentemente<sup>5</sup> –, têm sobre aquilo que é dito na internet. Elas não são folhas em branco, em que as pessoas expressam livremente aquilo que pensam, mas a parte visível de companhias capazes de exercer poder sobre os espaços em que as pessoas passaram a se manifestar<sup>6</sup>. Como o aviso feito pelos moderadores expôs, essas empresas não decidem apenas *o que* fica ou não em suas plataformas, mas também determinam *como* as pessoas podem se manifestar nesses ambientes e *quem* pode tomar parte deles, respondendo a participantes que descumprem suas regras com sanções que vão da retirada do conteúdo ao banimento do usuário que o disponibilizou.

Segundo, ao exercer esse poder, essas empresas fazem juízos complexos quanto ao direito à liberdade de expressão, que repercutem em espaços que vão muito além das comunidades virtuais que criaram. Como a reclamação feita por Lipman expõe, tomar decisões a esse respeito gera controvérsias não apenas quanto a se uma publicação deve ou não permanecer no Facebook, mas, principalmente, a se pessoas foram tolhidas pela empresa na liberdade que têm de se expressar, fazendo-as mobilizar uma gramática específica para defendê-la: segundo Lipman, “*a postagem certamente faz uma afirmação*

---

<[https://www.washingtonpost.com/blogs/erik-wemple/wp/2013/06/29/facebook-admits-mistake-in-blocking-fox-news-radio-commentator/?noredirect=on&utm\\_term=.61936934d549](https://www.washingtonpost.com/blogs/erik-wemple/wp/2013/06/29/facebook-admits-mistake-in-blocking-fox-news-radio-commentator/?noredirect=on&utm_term=.61936934d549)>, acesso em 30.7.2019.

<sup>5</sup> Cf., a esse respeito, respectivamente, quanto ao Twitter e o YouTube, ROGERS, K. Twitter ‘sorry’ for suspending Guy Adams as NBC withdraws complaint. **The Guardian**, 31.7.2012, disponível em <<https://www.theguardian.com/technology/2012/jul/31/guy-adams-twitter-growing-pains>>; e BERGEN, M. YouTube's New Moderators Mistakenly Pull Right-Wing Channels. **Bloomberg**, 28.2.2018, disponível em <<https://www.bloomberg.com/news/articles/2018-02-28/youtube-s-new-moderators-mistakenly-pull-right-wing-channels>>. A constância em que o Facebook enfrenta problemas semelhantes fica evidente pela existência de uma situação muito parecida poucos meses antes do incidente, também envolvendo a diferença de tratamento entre páginas anti-israelenses e anti-palestinos, cf. BLUM, R. Israeli ‘Lawfare’ Group Says Experiment Reveals Anti-Jewish Bias on Facebook. **The Algemeiner**, 05.01.2016, disponível em <<https://www.algemeiner.com/2016/01/05/israeli-lawfare-group-experiment-reveals-anti-jewish-bias-on-facebook/>>, acessos em 30.7.2019.

<sup>6</sup> “*By understanding moderation not just as an occasional act platforms must engage in but as a fundamental aspect of their service and a fundamental part of their place in public discourse, we can reconsider what platforms are, and ask new questions about their power in society*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 14 (Kindle, n. p.); no mesmo sentido, cf. AGRE, P. *The Internet and Public Discourse*. **First Monday**, v. 3, n. 3, n. p., 1998.

*política, mas esse é o ponto central da liberdade de expressão – a habilidade de expressar a visão de alguém desde que ela não incita alguém a agir com violência*”. Se a popularização da internet desencadeou a criação de novos espaços de manifestação para as pessoas, levando-as das ruas às redes sociais – as “*praças públicas da modernidade*”<sup>7</sup> –, ela também fez das decisões tomadas a respeito do conteúdo que é permitido nessas plataformas algo muito maior do que a forma interna de se organizar um negócio. Quando uma publicação é retirada do Facebook, do Twitter ou do YouTube, ela não é excluída apenas de um banco de dados na Califórnia: ela deixa de existir no mundo enquanto uma forma de expressão<sup>8</sup>.

Terceiro, o caso expõe que o aumento da importância das decisões tomadas por essas empresas é inversamente proporcional àquilo que se sabe sobre como esses julgamentos são feitos. Mesmo após uma carta aberta, o reestabelecimento da postagem e o pedido de desculpas feito pelo Facebook, ficamos sem entender o que realmente aconteceu. Por que a publicação foi inicialmente excluída? Que tipo de regra ela violou? Quem decidiu removê-la? Por que isso foi considerado um erro após a demonstração, feita por Lipman, de que havia uma inconsistência com o juízo feito em uma situação semelhante? Por que, ao invés de reestabelecer a primeira publicação, o Facebook não decidiu banir as duas? Em síntese: como, afinal de contas, decisões quanto àquilo que as pessoas podem dizer são tomadas pelas plataformas da internet?<sup>9</sup>

---

<sup>7</sup> Cf. a referência feita pela Suprema Corte no caso *Packingham v. North Carolina*, 582 U. S. \_\_\_\_ (2017), em que uma lei da Carolina do Norte que impedia, genericamente, o acesso de pessoas condenadas por crimes sexuais a redes sociais foi declarada inconstitucional por violar a Primeira Emenda; no mesmo sentido, KLONICK, K. *The New Governors: The People, Rules, And Processes Governing Online Speech*. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.611.

<sup>8</sup> Para uma análise detalhada desses impactos, cf., por exemplo, BALKIN, J. *Old-School / New-School Speech Regulation*. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.306 e ss.

<sup>9</sup> Para uma análise geral quanto a esse ponto, cf. o Projeto “*Tossed Out*”, iniciado pela Electronic Frontier Foundation com o objetivo de “*highlights the various ways in which Terms of Service (TOS) and other speech moderation rules are unevenly enforced, with little to no transparency, against a wide spectrum of people*” (“*destacar as várias formas pelas quais os Termos de Serviço (TS) e outras regras de moderação da expressão são aplicadas, com nenhuma ou pouquíssima transparência, contra um grande espectro de pessoas*”). ELECTRONIC FRONTIER FOUNDATION. **Tossed Out**, disponível em <<https://www.eff.org/tossedout>>, acesso em 8.1.2020.

Durante um simpósio dedicado à liberdade de imprensa e organizado pela *Harvard Law Review* em 2014<sup>10</sup>, um dos convidados, Marvin Ammori – um acadêmico e advogado de atuação vinculada à proteção dos direitos civis – propôs uma investigação para responder a esse tipo de pergunta<sup>11</sup>. Na apresentação, Ammori sugeriu uma via não tradicional para compreender a relação entre internet e liberdade de expressão: ele se dispôs a investigar como a popularização da internet transformou a advocacia especializada na liberdade de expressão nos Estados Unidos. Partindo de uma reconstrução biográfica de alguns casos, Ammori argumentou que os estudantes de direito que sonhavam em escrever os próximos capítulos da defesa da liberdade de expressão na década de 90 – e, por isso, iniciaram suas carreiras em escritórios que representavam judicialmente conglomerados como o *The New York Times*, litigantes nos principais precedentes da época<sup>12</sup> –, deixaram paulatinamente para trás esse tipo de atuação, construindo carreiras de sucesso não como representantes da mídia tradicional – jornais, revistas e cadeias televisivas –, mas se tornando advogados em empresas de tecnologia, como o Facebook, o Twitter e o Google.

Ammori descobriu, além disso, que essa mudança foi mais radical do que parecia à primeira vista. Esses advogados não se tornaram gestores de litígios judiciais, que trocaram as ações indenizatórias contra o *The New York Times* pelos processos em que o Facebook é parte. Eles não foram chamados para *representar* essas novas empresas em juízo, mas para *construir*, dentro delas, soluções independentes que permitissem resolver o

---

<sup>10</sup> Trata-se do simpósio “*Freedom of Press*”, organizado em fevereiro de 2014 pela *Harvard Law Review*. Os textos estão disponíveis em <<https://harvardlawreview.org/roundtable/symposium-freedom-of-the-press/>>, acesso em 30.7.2019.

<sup>11</sup> As referências foram retiradas da versão revisada da apresentação, publicada no artigo AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, pp. 2.259/2.295, 2014.

<sup>12</sup> *Cf.*, por exemplo, o precedente citado por Ammori no texto, *New York Times Co. v. Sullivan*, 376 US 254 (1964), em que uma controvérsia quanto a um anúncio de um grupo de direitos civis no jornal levou à fixação, pela Suprema Corte, de um critério estreito de responsabilização da imprensa por aquilo que publica. Há, entretanto, inúmeros outros julgados da época em que questões quanto à liberdade de expressão foram fixados em casos nos quais a imprensa era parte, como *New York Times Co. v. United States*, 403 U.S. 713 (1971), *Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241 (1974), *Time, Inc. v. Firestone*, 424 U.S. 448 (1976), *Nebraska Press Association v. Stuart*, 427 U.S. 539 (1976), *Richmond Newspapers Inc. v. Virginia*, 448 U.S. 555 (1980) e *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46 (1988).

problema que Lipman levantou: definir, consistentemente, aquilo que os usuários dessas plataformas podem ou não dizer na internet.

Os advogados de Ammori, por isso, não se especializaram em *defender* a liberdade de expressão em juízo, mas em elaborar estruturas que permitissem *decidir* se, por exemplo, uma postagem feita no Facebook por um site neonazista, que ofendia uma mulher morta em protestos conservadores nos Estados Unidos, deveria ou não permanecer ativa<sup>13</sup>; ou, então, o que fazer com publicações que incitavam, por meio de um sistemático discurso de ódio, a violência étnica contra minorias islâmicas em Mianmar – algo que, mais tarde, revelou-se parte de uma estratégia militar que levou ao êxodo de mais de 700 mil pessoas<sup>14</sup>; ou, por fim, se imagens de mulheres amamentando – integrantes de grupos de apoio mútuo feminino – ou de uma criança sem roupas – fugindo de um ataque de napalm no Vietnã – representavam uma violação à política de vedação à nudez do Facebook<sup>15</sup>.

Decidir o que fazer nesses casos é uma tarefa, em si, difícil, que teve que ser conciliada com uma característica típica da internet: a escala<sup>16</sup>. Justamente o que faz das plataformas espaços de expressão relevantes e lucrativos – o número de usuários,

---

<sup>13</sup> NEWTON, C. Facebook is deleting links to a viral attack on a Charlottesville victim. **The Verge**, 14.8.2017, disponível em <<https://www.theverge.com/2017/8/14/16147126/facebook-delete-viral-post-charlottesville-daily-stormer>>, acesso em 30.7.2019. Para uma análise quanto ao papel desempenhado pela moderação de conteúdo no caso, cf. KLONICK, K. Facebook v. Sullivan. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, p. 2 e ss e nota n. 2, disponível em <<https://knightcolumbia.org/content/facebook-v-sullivan>>, acesso em 30.7.2019.

<sup>14</sup> “Facebook was used to incite violence and hatred against the Muslim minority group [de Mianmar]. The platform, she said, had ‘turned into a beast.’”. STECKLOW, S. Why Facebook is losing the war on hate speech in Myanmar. **Reuters**, 15.8.2018, disponível em <<https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>>, acesso em 30.7.2019.

<sup>15</sup> Cf., respectivamente, os relatos que Gillespie fornece quanto ao “*Caso Breastfeeding*” e o “*Caso Napalm*” – em que a foto “*O Terror da Guerra*” foi reiteradamente excluída por moderadores do Facebook – em GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 1 e ss e 141 e ss (Kindle, n. p.); cf., também, *infra*, Cap. I.1, 3.2.2.1.

<sup>16</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 74 e ss (Kindle, n. p.).

atualmente estimados em 2,27 bilhões no Facebook<sup>17</sup>, e a constância da interação entre eles – tornou o problema de moderação de conteúdo particularmente desafiador: entre julho e setembro de 2018, por exemplo, o Facebook agiu para aplicar sua política de proibição ao discurso de ódio em mais de 3 milhões de publicações<sup>18</sup>; o Twitter, por sua vez, interviu em mais de 600 mil contas entre janeiro e junho de 2018<sup>19</sup>; e o YouTube desativou mais de 1,6 milhão de canais entre julho e setembro de 2018<sup>20</sup>. A construção de uma solução profissionalizada de tomada de decisões sobre o conteúdo disponibilizado nessas plataformas – isto é, de uma estrutura capaz de analisar, situação a situação, aquilo que era veiculado por seus usuários na internet – teve que levar em conta tanto a complexidade dos casos, quanto o volume de material sujeito a esse tipo de juízo, algo inerente a um modelo de negócios que se baseia no estímulo à participação e à interação nessas comunidades.

A partir desse quadro – e do que se sabe com dados obtidos essencialmente por vazamentos internos e divulgações tão recentes quanto limitadas<sup>21</sup> – é possível esboçar um início de resposta à pergunta do caso de Lipman. Plataformas como o Facebook, o Twitter e o YouTube definem, abstratamente, aquilo que pode ou não ser veiculado nas suas páginas, estabelecendo regras gerais – os Padrões de Comunidade do Facebook, por exemplo – como “*Não publique [...] Qualquer discurso violento ou apoio de forma escrita*

---

<sup>17</sup> FACEBOOK. Stats. **Facebook Newsroom**, 2018, disponível em <<https://newsroom.fb.com/company-info/#statistics>>, acesso em 23.7.2019; um usuário mensalmente ativo é aquele que entrou pelo menos uma vez na plataforma no último mês.

<sup>18</sup> FACEBOOK. Community Standards Enforcement Report. **Facebook Transparency**, 2018, disponível em <<https://transparency.facebook.com/community-standards-enforcement#hate-speech>>, acesso em 30.7.2019.

<sup>19</sup> TWITTER. Twitter Rules enforcement. **Twitter Transparency Report**, 2018, disponível em <<https://transparency.twitter.com/en/twitter-rules-enforcement.html>>, acesso em 30.7.2019.

<sup>20</sup> GOOGLE. Aplicação das diretrizes da comunidade do YouTube. **Google Transparency Report**, 2018, disponível em <<https://transparencyreport.google.com/youtube-policy/removals>>, acesso em 30.7.2019.

<sup>21</sup> Cf., por exemplo, a controvérsia quanto à divulgação das regras utilizadas para julgar o conteúdo disponibilizado pelo Facebook, que foram vazadas inicialmente por uma reportagem do *The Guardian* em maio de 2017 (os “*Facebook Files*”), e acabaram parcialmente disponibilizadas pela plataforma quase um ano depois, em abril de 2018. Cf. HOPKINS, N. Revealed: Facebook's internal rulebook on sex, terrorism and violence. **The Guardian**, 21.5.2017, disponível em <<https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence>>; ZUCKERBERG, M. (Mark Zuckerberg). Sem título. Sem local, 24 de abril de 2018. **Facebook**: zuck. Disponível em <<https://www.facebook.com/zuck/posts/10104874769784071>> e FACEBOOK. **Padrões de Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acessos em 30.7.2019.

ou *visual*”<sup>22</sup>. Grupos internos interpretam esses Padrões, formulando esquematizações dessas regras – que não são divulgados – traduzindo-as em um formato mais diretamente aplicável: no caso do discurso de ódio, há tabelas indicando símbolos que, se publicados, devem ser removidos da plataforma<sup>23</sup>. Esses padrões são repassados a equipes que recebem e analisam conteúdos identificados por usuários – ou pela própria plataforma – como violadores de suas regras. Se um membro dessa equipe – ou um algoritmo parametrizado para isso – detecta a violação das regras da plataforma, ele aplica um conjunto de sanções – da retirada da publicação à exclusão da comunidade – e avisa o usuário com uma mensagem genérica. Não havia, quando Ammori escreveu o texto, transparência quanto às justificativas e ao procedimento, mecanismos de contraditório nem a possibilidade de se recorrer desse tipo de decisão<sup>24</sup>.

A pergunta de Lipman e a resposta de Ammori chamam a atenção para algo que vai além de uma divergência política na internet ou de uma análise sociológica quanto à evolução de uma profissão. Por trás de ambas está uma constatação comum: a de que internet não criou apenas um novo *espaço de manifestação*, mas também gerou um novo *espaço de decisão* quanto àquilo que as pessoas podem dizer. Isso significa que, ao estruturar negócios baseados em plataformas nas quais as pessoas pudessem disponibilizar aquilo que pensam, companhias como o Facebook, o Twitter e o YouTube viram-se diante do problema de ter que decidir aquilo que seria ou não aceitável que seus usuários

---

<sup>22</sup> FACEBOOK. **Padrões de Comunidade: discurso de ódio**. Disponível em <[https://pt-br.facebook.com/communitystandards/hate\\_speech](https://pt-br.facebook.com/communitystandards/hate_speech)>, acesso em 30.7.2019.

<sup>23</sup> Existe uma diferença entre os Padrões de Comunidade dessas plataformas – atualmente divulgados – e as orientações que os moderadores efetivamente recebem para tomar as decisões, que permanecem ocultas – o acesso que se tem a respeito é feito via, em regra, vazamentos internos, como os documentos pelo Motherboard e pelo The New York Times. Cf., respectivamente, KOEBLER, J. e COX, J. The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People. **Motherboard**, 23.8.2018, disponível em <[https://motherboard.vice.com/en\\_us/article/xwk9zd/how-facebook-content-moderation-works](https://motherboard.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works)>; e, mais recentemente, FISHER, M. Inside Facebook’s Secret Rulebook for Global Political Speech. **The New York Times**, 27.12.2018, disponível em <<https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>>, acessos em 22.7.2019; cf. *infra*, Cap. I.2, 2.3.1.

<sup>24</sup> Como exponho adiante, esse é um procedimento em transformação; o Facebook, por exemplo, implantou no início de 2018 uma sistemática de apelação de suas decisões. Cf., para o detalhamento dessa análise, BICKERT, M. Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process. **Facebook Newsroom**, 24.4.2018, disponível em <<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>>, acesso em 30.7.2019; e *infra*, Cap. II.3, 2.



veiculassem na rede, compatibilizando a dificuldade inerente a esse tipo de questão com a escala que é típica da internet.

Quando a solução ficou pronta, ela tomou a forma da moderação de conteúdo das plataformas da internet: um complexo e profissionalizado mecanismo decisório que produz milhares de juízos diariamente quanto aos limites daquilo que as pessoas podem dizer em um dos principais espaços de expressão atual – a própria internet. Por ele, regras são criadas, violações são apuradas e julgamentos são proferidos a partir de uma estrutura particular e, por isso, essencialmente transnacional, privada e opaca – muito diferente, por isso, do juiz local, investido pelo Estado e que decide segundo o devido processo legal. Se a internet deslocou o espaço de manifestação das pessoas da praça ao Facebook, ela também mudou a forma como parte importante das decisões a esse respeito são tomadas – levando-as do juízo à plataforma e constituindo, com isso, uma nova instância decisória que exerce poder sobre aquilo que as pessoas podem ou não dizer:

*“[...] Os Estados Unidos agora têm dois sistemas que adjudicam disputas decorrentes de linguagem nociva quanto às outras pessoas. A primeira é mais antiga e familiar: o sistema de indenizações em que juízes resolvem demandas apresentadas sob leis de difamação e privacidade. O segundo é mais novo e menos entendido: as políticas de moderação de conteúdo e as práticas de plataformas privadas como o Facebook.”<sup>25</sup>*

O problema é que esse novo espaço de decisão passa por transformações que têm desafiado, com sucesso, a forma como nós o compreendemos. Ele é, em síntese, tão importante quanto desconhecido – seja porque opaco, seja porque se transforma a partir de vetores que as teorias que temos disponíveis para analisá-lo não conseguem identificar.

Para entender melhor o que isso significa, é importante compreender que a moderação de conteúdo não é uma criação original, mas uma reunião de escolhas e

---

<sup>25</sup> “[...] *the United States now has two systems to adjudicate disputes arising from harmful speech about other people. The first is older and more familiar: the tort system in which judges resolve claims brought under state defamation and privacy law. The second is newer and less well understood: the content moderation policies and practices of private platforms such as Facebook*”. KLONICK, K. Facebook v. Sullivan. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, p. 2, disponível em <<https://knightcolumbia.org/content/facebook-v-sullivan>>, acesso em 30.7.2019. A percepção quanto a esse papel regulatório das plataformas não é recente; *cf.*, por exemplo, BELLI, L., FRANCISCO, P. A., ZINGALES, N. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police in BELLI, L., ZINGALES, N. (ed.) **Platform Regulators: How Platforms are Regulated and How They Regulate Us**. Rio de Janeiro: FGV, 2017, pp. 41/64.

adaptações de modelos anteriores, que, de acordo com o contexto, forneceram respostas contingentes à pergunta de como decidir se alguém pode ou não dizer algo. Essa seletividade fica clara quando se percebe que, assim como nas decisões tomadas em processos judiciais, há, na moderação de conteúdo, questões *materiais* e *processuais* relativamente independentes: determinar se um tipo de discurso de ódio será ou não permitido no Twitter é uma tarefa diferente de se saber se o usuário poderá se manifestar quanto à decisão, da companhia, de excluir algo que ele compartilhou.

Essa distinção é importante porque expõe a seletividade com que essa forma de decidir foi construída: se no *campo material* esse modelo toma como referência o quadro geral da forma como o Estado – adotando, por exemplo, a “*doutrina da Primeira Emenda*”<sup>26</sup> – lida com controvérsias quanto à liberdade de expressão, o mesmo não acontece quanto ao *campo procedimental*. Como expus acima, esse é um mecanismo construído em um contexto privado, o que viabilizou que seus idealizadores estruturassem um procedimento muito mais próximo daquele que produz uma diretriz corporativa do que uma decisão judicial: modera-se conteúdo a partir de um processo em que não há transparência, contraditório, nem recorribilidade, uma forma de decidir essencialmente diferente da que produz decisões vinculantes em contextos regulatórios estatais e dos quais as plataformas extraem a base material do que fazem. É como se as plataformas tivessem implantado um sistema híbrido, que diz garantir direitos como os da Constituição dos Estados Unidos tomando decisões pelo procedimento utilizado por um Tribunal da Arábia Saudita.

O problema é que essa distinção de influências entre os campos material e procedimental tem sido cada vez mais atenuada por uma série de transformações exigidas pelos usuários e operadas pelas companhias que organizam essas plataformas.

---

<sup>26</sup> Por referência não quero dizer identidade: mesmo que em menor grau, há divergências quanto às políticas de conteúdo das plataformas e a “*doutrina da Primeira Emenda*”, notadamente quanto ao discurso de ódio; como Ammori aponta, trata-se mais de uma influência quanto ao “*modo de pensar*” do que como uma submissão jurídica entre a política de conteúdo dessas plataformas às normas de um país – no caso, a jurisprudência da Suprema Corte. Cf., p. ex., AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.274/2.275 e 2.283; ARMIJO, E. Meet the New Governors, Same as the Old Governors. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, disponível em <<https://knightcolumbia.org/content/meet-new-governors-same-old-governors>>, acesso em 30.7.2019; e *infra*, Cap. I.2, 3.2.1.

Por um lado, há a pressão exercida por setores sociais quanto ao modo como essas decisões são tomadas, exigindo delas cada vez menos uma decisão corporativa e cada vez mais uma decisão “judicial”: no início de 2018, por exemplo, um grupo de organizações, acadêmicos e advogados, dentre os quais a União Americana pelas Liberdades Cívicas, aprovou um conjunto de princípios quanto à transparência e a responsabilidade na moderação de conteúdo, dentre os quais a necessidade de “*informar a cada usuário cujo conteúdo é retirado ou a conta suspensa a razão pela retirada ou suspensão*”, e fornecer a “*real oportunidade de se apelar de qualquer determinação de remoção de conteúdo ou suspensão de conta*”<sup>27</sup>. Alguns meses depois, 88 organizações de defesa dos direitos civis enviaram uma carta aberta ao Facebook, sustentando que a plataforma “*está muito atrás de seus competidores em fornecer aos seus usuários o devido processo legal [quanto à moderação de conteúdo]*”, instando-a a garantir aos participantes de sua comunidade a possibilidade de recorrer de suas decisões<sup>28</sup>.

Por outro, é incontroverso que essas companhias têm realizado modificações estruturais alinhadas com essas exigências, apreendendo elementos do modo como controvérsias quanto à liberdade de expressão são tratadas nos tribunais e aplicando-os no procedimento decisório da moderação de conteúdo.

O Facebook, por exemplo – em uma mudança semelhante às constatadas no Twitter<sup>29</sup> e no YouTube<sup>30</sup> – iniciou, em 2018, um movimento para transformar a

---

<sup>27</sup> “*Companies should provide notice to each user whose content is taken down or account is suspended about the reason for the removal or suspension*”; “*Companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 30.7.2019.

<sup>28</sup> ELECTRONIC FRONTIER FOUNDATION. 88 rights groups call for Facebook to implement appeals process for removed content. **IFEX**, 13.11.2018, disponível em <[https://ifex.org/digital\\_rights/2018/11/13/facebook\\_eff\\_letter/](https://ifex.org/digital_rights/2018/11/13/facebook_eff_letter/)>, acesso em 30.7.2019.

<sup>29</sup> O Twitter publica desde 2012 relatórios de transparência, mas 2018 foi o primeiro ano em que dados quanto à execução dos Termos de Serviço da plataforma – isto é, quanto à moderação de conteúdo – foram disponibilizados. TWITTER. Twitter Rules enforcement. **Twitter Transparency Report**, 2018, disponível em <<https://transparency.twitter.com/en/twitter-rules-enforcement.html>>, acesso em 30.7.2019; cf. *infra*, Cap. II.3, 2.

<sup>30</sup> Cf., por exemplo, a publicação, pelo Google, do primeiro relatório de transparência quanto à remoção do conteúdo no Youtube, em abril de 2018. GOOGLE. Aplicação das diretrizes da comunidade do YouTube. **Google Transparency Report**, 2018, disponível em <<https://transparencyreport.google.com/youtube-policy/removals>>. A despeito da escassez de trabalhos propriamente acadêmicos quanto a esse movimento,

moderação de conteúdo em uma forma de julgar que garanta algum nível de transparência, contraditório, recorribilidade às suas decisões. O ano começou com o anúncio de que a plataforma publicaria aos seus usuários os seus Padrões de Comunidade – que orientam as decisões tomadas pelos moderadores – e instituiria um procedimento de apelação quanto ao conteúdo que eles excluam, direcionando-o para a própria plataforma<sup>31</sup>. Cumpridas as promessas, a plataforma avançou e estabeleceu uma sistemática de relatórios periódicos quanto à execução dos Padrões de Comunidade pela moderação de conteúdo, produzindo um primeiro documento em maio de 2018<sup>32</sup>, seguido do anúncio, em novembro, de um novo patamar de transparência quanto a esses dados, viabilizando pesquisas quantitativas a respeito; da ampliação do procedimento de apelação, para abarcar praticamente todas as decisões da moderação de conteúdo; e, por fim, da formação de um corpo decisório independente de sua estrutura empresarial – a “*Suprema Corte do Facebook*”<sup>33</sup> – que, com poder sobre as decisões tomadas na moderação de conteúdo da plataforma, teria o objetivo “*defender o princípio de conferir voz às pessoas, ao mesmo tempo em que reconhece a necessidade real de manter as pessoas seguras*”<sup>34</sup>. O projeto foi à frente e, em 28 de

---

ele é largamente documentado na mídia especializada; *cf.*, por exemplo, BANKSTON, K. SINGH, S. Facebook And Google Finally Take First Steps On Road To Transparency About Content Moderation. **Techdirt**, 26.4.2018, disponível em <<https://www.techdirt.com/articles/20180426/10164939724/facebook-google-finally-take-first-steps-road-to-transparency-about-content-moderation.shtml>>, acessos em 30.7.2019.

<sup>31</sup> ZUCKERBERG, M. (Mark Zuckerberg). Sem título. Sem local, 24 de abril de 2018. **Facebook**: zuck. Disponível em <<https://www.facebook.com/zuck/posts/10104874769784071>>; e BICKERT, M. Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process. **Facebook Newsroom**, 24.4.2018, disponível em <<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>>; acessos em 30.7.2019.

<sup>32</sup> FACEBOOK. Community Standards Enforcement Report. **Facebook Transparency**, 2018, disponível em <<https://transparency.facebook.com/community-standards-enforcement#hate-speech>>, acesso em 30.7.2019.

<sup>33</sup> A ideia – e o termo – já haviam sido utilizados por Zuckerberg antes (*cf.*, por exemplo, uma entrevista a respeito em abril do mesmo ano em KLEIN, E. Mark Zuckerberg on Facebook’s hardest year, and what comes next. **Vox**, 2.4.2018, disponível em <<https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge>>), mas apenas após o anúncio oficial em novembro de 2018 que ganhou repercussão, inclusive com discussões a respeito de como desenhar institucionalmente essa solução. *Cf.* KLONICK, K., KADRI, T. How to Make Facebook’s ‘Supreme Court’ Work. **The New York Times**, 17.11.2018, disponível em <<https://www.nytimes.com/2018/11/17/opinion/facebook-supreme-court-speech.html>>, acessos em 30.7.2019; e *infra*, Cap. II.3, 2.

<sup>34</sup> ZUCKERBERG, M. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

janeiro de 2019, a plataforma anunciou a formação, ao longo do ano, do “*Comitê de Supervisão para Decisões de Conteúdo*”, um órgão independente para “*analisar as decisões de conteúdo do Facebook*”, “*Reverter decisões do Facebook quando necessário*” e “*ser uma autoridade independente externa ao Facebook*”<sup>35</sup> – uma espécie de corte arbitral internacional da liberdade de expressão. Pouco tempo depois, a plataforma abriu uma consulta pública sobre o modelo, na qual representados 88 países, via 28 workshops e mesas-redondas, colhendo informações com mais de 2.000 pessoas e que resultou em um documento<sup>36</sup> no qual consolidadas as propostas a serem colocadas em prática nas próximas etapas<sup>37</sup>. Por fim, em dezembro de 2019, o Facebook anunciou que constituiu um fundo independente (*trust*) de 130 milhões de dólares que financiará as atividades dos seis anos iniciais do painel – que devem se iniciar ainda em 2020 –, e atuará sem vinculação à empresa, atendendo diretamente aos custos do Conselho de Governança<sup>38</sup>.

Essa é uma mudança usualmente interpretada a partir de uma narrativa otimista, em parte induzida pelas próprias plataformas: atentas à sua responsabilidade social e movidas por pressões de grupos organizados, empresas como o Facebook teriam decidido modificar a moderação de conteúdo, adequando-a a um modelo que melhor atendesse àquilo que esperam seus usuários, em uma perspectiva amplamente explorada, por exemplo, no relatório que mencionei acima<sup>39</sup>. Assim, esse movimento seria,

---

<sup>35</sup> CLEGG, N. Charting a Course for an Oversight Board for Content Decisions. **Facebook Newsroom**, 28.1.2019, disponível em <<https://newsroom.fb.com/news/2019/01/oversight-board/>>, acesso em 30.7.2019.

<sup>36</sup> FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 29.7.2019.

<sup>37</sup> FACEBOOK. Global Feedback and Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, disponível em <<https://newsroom.fb.com/news/2019/06/global-feedback-on-oversight-board/>>, acesso em 29.7.2019.

<sup>38</sup> HARRIS, B. An Update on Building a Global Oversight Board. **Facebook Newsroom**, 12.12.2019, disponível em <<https://about.fb.com/news/2019/12/oversight-board-update/>>, acesso em 8.1.2020.

<sup>39</sup> “*I’ve always cared about this and that’s why the first words of our mission have always been ‘give people the power’. I believe the world is better when more people have a voice to share their experiences, and when traditional gatekeepers like governments and media companies don’t control what ideas can be expressed*” (“*Eu sempre me importei sobre isso e é por isso que as primeiras palavras de nossa missão sempre foi ‘dê às pessoas o poder’. Eu acredito que o mundo é melhor quando mais pessoas têm voz para compartilhar suas experiências e quando os gatekeepers tradicionais como governos e companhias de mídia não controlam quais ideias podem ser expressadas*”). ZUCKERBERG, M. A **Blueprint for Content Governance and**

simultaneamente, um testemunho do sucesso desses movimentos sociais e um esteio para um argumento comum a essas companhias, de que não são meras estruturas empresariais, mas comunidades moldadas por seus participantes e que, como indica o slogan do Facebook, servem de “*utilidade social que conecta você às pessoas ao seu redor*”. Da mudança, portanto, ficaria a evidência do quanto essas companhias dependem dos usuários e permanecem atentas àquilo que lhes fornece uma experiência melhor, mesmo que isso signifique abrir mão de uma parte importante do poder que têm sobre aquilo que circula na internet por suas plataformas.

O problema é que esse paralelismo intuitivo entre, de um lado, demanda dos usuários e, de outro, mudança nas plataformas – que domina o modo como essa transformação é interpretada<sup>40</sup> –, é enganoso, porque sugere uma causalidade única onde pode haver só coexistência, correlação ou, no mínimo, uma causalidade mais complexa entre esses fatores, que permanece oculta se, aceitando aquilo que dizem essas companhias, entendemos essas modificações como um simples atendimento àquilo que os membros dessas comunidades exigiram. Mais do que isso: ela toma a perspectiva dos usuários – ou seja, daqueles que pressionam as plataformas por direitos e que, por isso, teriam sido atendidos – como privilegiada para explicar essa transformação, uma postura que gera dúvidas porque, no fim, são as próprias plataformas que definem os rumos da moderação de conteúdo. Assim, e por mais atrativa que seja essa narrativa, há indicativos de que, por trás dela, há algo mais complexo que deixamos de compreender quando a aceitamos sem questionar.

Isso porque, por um lado, a estruturação das políticas de moderação de conteúdo não foi um processo inocente, mas um modelo decisório cuidadosamente criado por decisões estratégicas da cúpula dessas companhias com o objetivo de reter o controle do que permanece ou não nas plataformas. A centralização era tão marcante que a executiva responsável pela política de conteúdo do Google, Nicole Wong, ficou conhecida

---

**Enforcement.** 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

<sup>40</sup> Cf., por exemplo, as dúvidas levantadas após a publicação dos Santa Clara Principles, quanto a em que medida eles seriam ou não atendidos pelas plataformas, como se essa mudança fosse uma espécie de carta de exigências apresentada pelos usuários a essas companhias. BRANDOM, R. New rules challenge Google and Facebook to change the way they moderate users. **The Verge**, 7.5.2018, disponível em <<https://www.theverge.com/2018/5/7/17328764/santa-clara-principles-platform-moderation-ban-google-facebook-twitter>>, acesso em 30.7.2019.

como “*aquela que decide*”<sup>41</sup> (“*The Decider*”), por ter a palavra final sobre tudo que passava pelas plataformas da empresa – algo compreensível, inclusive sob a perspectiva da mitigação de riscos decorrentes da responsabilidade civil, porque nem todos os países têm, como os Estados Unidos, uma política de imunidade em relação àquilo que é veiculado por terceiros nos sites dessas companhias<sup>42</sup>. Faz sentido, assim, que essas empresas queiram ter o controle sobre aquilo que é veiculado na internet por suas plataformas: o que surpreende, nessa transformação, não é a dimensão do controle exercido sobre o conteúdo, mas sim a disposição para abrir mão dele. É exatamente por isso que é preciso ter cautela com uma leitura que afirme que essas companhias, depois de mobilizarem um corpo especializado de advogados, estruturar um formato próprio de análise de conteúdo e operá-lo por anos segundo uma – justificável – lógica de opacidade, tenham agora percebido que a forma mais adequada de se chegar a uma decisão é a que já existia muito antes de conceberem seus modelos de moderação de conteúdo.

Se o descompasso entre uma base material associada à Primeira Emenda e uma base processual fundamentalmente distinta do devido processo legal não foi uma obra do acaso – mas o produto de uma forma estratégica de pensar essas plataformas – é bem possível que a recente mudança nesse paradigma também tenha sido muito mais que um gesto de simpatia a usuários que exigiam isso, sobretudo pela forma como o Facebook lida seletivamente com as pressões que constantemente sofre de grupos sociais organizados<sup>43</sup>.

---

<sup>41</sup> “*During her tenure at Google, Wong (who later moved from Google to Twitter, and is currently working for the White House) was vested with the authority to decide ‘what goes up or comes down’ on Google’s various sites – including the popular video sharing site, YouTube*”. ZARSKY, T. Social Justice, Social Norms and the Governance of Social Media. **Pace Law Review**, v. 35, n. 1, 2014, p. 162.

<sup>42</sup> *Cf.*, a esse respeito, o papel central que a Seção 230 da Lei de Telecomunicações dos Estados Unidos têm na moderação de conteúdo ao inviabilizar a responsabilização de plataformas como o Facebook por aquilo que é disponibilizado nela por terceiros – uma regra, entretanto, que não está necessariamente presente em todas as jurisdições que a plataforma atua, algo que pode gerar estímulos no sentido de adotar políticas mais restritivas de moderação de conteúdo. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1604 e ss; GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 29 e ss (Kindle, n. p.).

<sup>43</sup> *Cf.*, por exemplo, a longa controvérsia – que remonta, pelo menos, às eleições dos Estados Unidos em 2016 – a respeito da transparência com que o Facebook lida com os anúncios em sua plataforma, em específico os associados a campanhas políticas. Nesse sentido, mesmo após alterar suas políticas a respeito, a plataforma ainda resiste em liberar dados estatísticos que seriam fundamentais para pesquisas de ciência política a respeito. INGRAM, D. Facebook to keep wraps on political ads data despite researchers' demands.

Tanto essa dualidade entre exigência dos usuários e atendimento das plataformas, quanto a narrativa mutuamente benéfica que decorre dela merecem, no mínimo, desconfiança: parece haver, por trás desse movimento todo, alguma espécie de cálculo estratégico vendido como atenção ao usuário que passa despercebido pelo modo como essa mudança tem sido interpretada.

Por outro lado, mesmo que a interpretação vigente – de que essa transformação é o resultado de uma dinâmica de exigência do usuário e atendimento da plataforma – esgote o modo como os participantes desse processo o interpretam, isso não significa que essa seja a visão mais completa do que ocorre, atualmente, com o modelo regulatório privado da liberdade de expressão na internet. Como os inúmeros problemas que têm surgido do modelo atual – dentre os quais o exemplificado pelo caso Lipman – demonstram, o domínio que as plataformas exercem sobre a moderação de conteúdo e a participação de seus usuários nesse mecanismo é limitado, e não significa que ambos sejam oniscientes quanto a esse modelo regulatório e suas transformações. Assim como entrevistar um parlamentar quanto às suas intenções sobre uma legislação que auxiliou a aprovar não é um dado suficiente para entender os movimentos que a gestaram e as suas consequências, colher aquilo que os executivos do Facebook esperam atingir com as mudanças que implantaram na moderação de conteúdo é algo insuficiente para compreendê-la totalmente. É preciso, portanto, tomar distância dessa mudança, analisando-a não como um conjunto de intencionalidades provenientes de plataformas e usuários, mas como um movimento contingente de um modelo decisório.

Seja, portanto, porque a narrativa sobre aquilo que está por trás desse movimento das plataformas é incompleta, seja porque se tem dado muita relevância à interpretação que os atores desse processo de mudança têm fornecido a seu respeito, o fato é que a compreensão que temos disponível quanto a esse modelo regulatório e suas transformações é deficitária. E isso é um problema, porque, à medida que esses modelos não conseguem apreender as transformações pelas quais a moderação de conteúdo tem passado, deixa-se de entender, paulatinamente, como funciona o mecanismo que, pelo menos na última década, tem decidido – em uma extensão territorial e humana muito maior

---

**Reuters**, 22.6.2017, disponível em <<https://www.reuters.com/article/us-usa-politics-facebook/facebook-to-keep-wraps-on-political-ads-data-despite-researchers-demands-idUSKBN19D1CN>>; **FACEBOOK**. Q & A on Transparency for Electoral and Issue Ads. **Facebook Newsroom**, 24.5.2018, disponível em <<https://newsroom.fb.com/news/2018/05/q-and-a-on-ads-transparency/>>, acessos em 30.7.2019.



que o modelo jurisdicional tradicional – aquilo que as pessoas podem dizer no principal âmbito de manifestação atual: a internet.

Faz sentido, por isso, investigar melhor essa forma de regular o que as pessoas podem dizer, empreendendo uma tentativa de compreender as razões por trás de suas transformações para esclarecer o quanto estabelecer, na moderação de conteúdo, uma forma de decidir mais próxima à do Estado significa algo além de um encontro de vontades entre usuários e plataformas. Por que as plataformas decidiram mudar o seu modelo decisório, guinando-o a um formato baseado naquele que obriga os Estados? Que razões existem para que elas, depois de desenvolverem um sistema próprio de moderação de conteúdo – construído para que tivessem o total controle sobre os seus resultados – agora se empenhem em transformá-lo em algo mais próximo do modelo decisório estatal, perdendo, com isso o domínio que têm sobre o ambiente que administram? Por que – e, em que medida – essas empresas estão abrindo mão do seu poder?<sup>44</sup>

Neste trabalho, meu objetivo é abordar a moderação de conteúdo como uma nova forma de criar regras, adjudicar condutas e aplicar punições quanto àquilo que as pessoas podem dizer, concebendo-a como um direito das plataformas. Minha proposta é fazer isso a partir da transformação pela qual a moderação de conteúdo atualmente passa, de um procedimento opaco para uma forma de decidir que se aproxima, cada vez mais, do paradigma pelo qual os Estados tomam decisões vinculantes quanto à liberdade de expressão em juízo, oferecendo explicações que, concorrendo com as que temos disponíveis, auxiliem a compreender essa mudança, esclarecendo o papel que essa forma de decidir tem nessa nova estrutura regulatória.

Para isso, adoto como perspectiva principal de minha análise não a dos usuários, que exigem os seus direitos – o que é a regra em abordagens sobre a moderação de conteúdo –, mas a das plataformas, como as responsáveis por tomarem essas decisões. Faço isso porque essa é uma transformação que, até o momento, é o resultado de uma ação

---

<sup>44</sup> Essa perda de poder – decorrente de um modelo decisório que retira, ao menos parcialmente, o controle da plataforma sobre o que permanece no espaço que administra – tem traços ainda mais surpreendentes porque vai de encontro com prognósticos da literatura especializada, que previam um movimento contrário, isto é, de aumento no poder dessas empresas. Nesse sentido, *cf.*, por exemplo, a afirmação de David Kaye – em favor de uma postura regulatória mais restritiva –, segundo o qual “*If left alone, these companies will gain even greater power over expression in public sphere*” (“*Se deixadas sozinhas, essas empresas vão ganhar ainda mais poder sobre a expressão na esfera pública*”). KAYE, D. **Speech Police: The Global Struggle to Govern the Internet**. New York: Columbia Global Reports, 2019, p. 147 (Kindle).

das empresas que estão por trás das plataformas da internet – como é o caso do Facebook e o seu Comitê de Supervisão para Decisões de Conteúdo – que, portanto, devem ser tomadas como o principal material de análise para a explicar. Dentro desses contornos, minha proposta é estudar as transformações pelas quais passa a moderação de conteúdo das plataformas da internet para descobrir quais são as razões por trás do movimento liderado por essas companhias. Espero, dessa forma, conseguir explicar melhor como funciona e como se transforma, enquanto um espaço de decisão, o mecanismo que regula a liberdade de expressão nas plataformas da internet.

Minha hipótese, a esse respeito, é que a transformação pela qual passa a moderação de conteúdo tem dois significados. Primeiro, argumento que o movimento das plataformas de adotar procedimentos decisórios mais próximos ao paradigma do devido processo legal não é uma forma de atender às exigências de seus usuários, mas a solução para um problema que enfrentam ao tomar esse tipo de decisão: a falta de legitimidade<sup>45</sup>. Sustento, nesse sentido, a partir de uma perspectiva funcionalista estrutural, que as plataformas se dispõem, nesse novo modelo, a abrir mão do poder de que dispunham sobre as manifestações de seus usuários em nome de um procedimento que os leva a aceitar as suas decisões, tomando-as como legítimas, mesmo que lhes sejam desfavoráveis. Como ocorre nos processos judiciais, estabelecer um procedimento decisório baseado em ideias como transparência, recorribilidade e contraditório nas plataformas da internet tem uma função-chave: conferir, à decisão que resulta desse processo, um corpo de legitimidade, isto é, uma maior possibilidade de que ela seja aceita, neutralizando o potencial destrutivo que tem sido ocasionado – como no caso Lipman – por divergências quanto às políticas de conteúdo na internet.

Segundo, argumento que essa transformação significa algo a mais: ela é parte de um processo de constitucionalização do mundo digital. Sustento que se a moderação de conteúdo for analisada sob uma perspectiva do pluralismo jurídico que confere um significado não-estatal tanto à ideia de direito, quanto à de constituição, é possível ler essa mudança como parte de uma transformação mais ampla que, viabilizada pela internet, permite perceber não só a formação de novos polos decisórios não-estatais – como a moderação de conteúdo –, mas também a constitucionalização dessas formas de decidir. Afirmo, nesse sentido, que essa constitucionalização digital é perceptível nas

---

<sup>45</sup> AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. *Harvard Law Review*, v. 127, n. 8, 2014, pp. 2.276/2.277.

plataformas pela adoção de formatos procedimentais capazes de imunizar decisões de moderação de conteúdo das estruturas corporativas em que estão inseridos, autonomizando-as tanto em relação às plataformas da internet, quanto de outros atores institucionais, como os Estados – algo que auxilia a compreender, com mais precisão, os passos seguintes desse modelo regulatório.

Para demonstrar isso, proponho uma abordagem diferente da que tem sido utilizada para entender a moderação de conteúdo. Como uma resposta à opacidade que acompanha esse modelo decisório desde o início, as pesquisas que o analisam adotam uma perspectiva centrada na obtenção de informações diretamente com os principais atores dessa atividade, via, por exemplo, entrevistas com os executivos dessas empresas<sup>46</sup>. Neste trabalho, minha proposta é diferente: não quero simplesmente expor aquilo que estava na mente dos executivos do Facebook quando decidiram adotar mudanças compatíveis com o que pediam seus usuários – algo que, no limite, forneceria narrativas que já temos, limitadas à suposta sensibilidade das plataformas às exigências de seus participantes.

Proponho, pelo contrário, que ferramentas da sociologia jurídica funcionalista sejam utilizadas para abordar tanto o processo decisório estatal quanto a moderação de conteúdo. Minha estratégia, nesse sentido, é suprir as deficiências das análises que temos disponíveis a partir de uma perspectiva funcionalista teórica que permita entender, em um primeiro momento – e independentemente da leitura que os próprios atores têm a esse respeito<sup>47</sup> –, qual é o papel que esse formato específico de decidir – resumidamente identificado sob a ideia de um devido processo legal – tem em sua origem: as decisões jurisdicionais tomadas no âmbito do Estado. Com isso, espero isolar a questão de particularidades locais – a regulamentação, caso a caso, desse tipo de procedimento –, extraíndo, a partir de um modelo geral, a função que esse formato decisório tem no seu âmbito de origem, algo precioso para a análise de contextos

---

<sup>46</sup> *Cf.*, a esse respeito, os dois artigos mais abrangentes quanto à questão, citados acima (AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.259/2.295; e KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1598/1670) e o recente livro dedicado à moderação de conteúdo (GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, 288 pp. (Kindle, n. p.)), todos estruturados a partir de entrevistas com pessoas que ocuparam postos-chave na moderação de conteúdo de plataformas da internet; e *infra*, Cap. I.2, 2 e 3.2.2.

<sup>47</sup> LUHMANN, N. **Ausdifferenzierung des Rechts: Beiträge zur Rechtssoziologie und Rechtstheorie**. Frankfurt: Suhrkamp, 1999, p. 281.

transnacionais, como o das plataformas da internet, em que as disposições normativas locais têm importância limitada. Essa é precisamente a vantagem da análise sociológica funcionalista, que a justifica como método: a capacidade de, sob uma perspectiva externa, responder a perguntas quanto ao papel que institutos jurídicos – como, por exemplo, o devido processo legal – têm em fenômenos sociais – a exemplo da tomada de decisões vinculantes – relacionando-os sem o compromisso, inerente às perspectivas internas, de se atingir uma decisão quanto ao direito aplicável ao caso concreto.

Em um segundo momento, minha expectativa é conseguir levar esse diagnóstico do Estado às plataformas, identificando como essa função inicialmente desempenhada no âmbito decisório estatal esclarece as mudanças pelas quais passa a moderação de conteúdo, obtendo conclusões que dependam menos de percepções pessoais de atores envolvidos nesse processo e dos direitos que eles exigem e sejam mais ligadas à percepção sociológica de como essas empresas e seus modelos jurídicos se transformam. Como exponho adiante, essa transição do Estado à plataforma é um movimento a ser feito com cautela – e que exigirá um passo teórico adicional –, porque as elaborações funcionalistas que temos disponíveis quanto ao devido processo legal foram produzidas a partir de contextos estatais, anteriores mesmo às versões iniciais de movimentos como o do pluralismo jurídico, em que questões quanto à existência de uma produção jurídica em contextos diferentes do estatal foram enfrentadas.

Dentre as análises sociológicas funcionalistas, adoto, neste trabalho, a Teoria dos Sistemas de Niklas Luhmann. Faço isso porque, primeiro, a Teoria dos Sistemas é uma perspectiva sociológica baseada no funcionalismo estrutural<sup>48</sup>, e, por isso, é uma ferramenta que viabiliza a estratégia que propus acima, ao permitir tomar as características dos procedimentos decisórios jurisdicionais – como a recorribilidade, o contraditório e a independência na tomada de decisões – como contingentes, analisando-as

---

<sup>48</sup> Cf. a declaração inicial de comprometimento da Teoria dos Sistemas de Luhmann com o funcionalismo estrutural em LUHMANN, N. **Soziologische Aufklärung 1**, Opladen: Westdeutscher Verlag, pp. 31/53, 1970, cuja versão aperfeiçoada consta em LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 36/44.; para uma análise contextualizada desse posicionamento com outras perspectivas sociológicas cf. NASSEHI, A. *Funktionale Analyse in* JAHRAUS, O, NASSEHI, A., GRIZELK, M., SAAKE, I., KIRCHMEIER, C., MÜLLER, J. **Luhmann-Handbuch: Leben, Werk, Wirkung**. Stuttgart: Springer, 2012, pp. 83/84; a introdução de Stephen Fuchs a LUHMANN, N. *Tautology and Paradox in the Self-Descriptions of Modern Society*. **Sociological Theory**, v. 6, n. 1, pp. 21/37, 1988; e ALCOVER, P. G. **El Derecho en La Teoría de la Sociedad de Niklas Luhmann**. Barcelona: J. M. Bosch, 1993, pp. 36 e ss.

não como consequências de um ordenamento jurídico específico – afirmando, por exemplo, que decisões são tomadas com contraditório porque esse é um direito constitucionalmente reconhecido no Brasil –, mas como engrenagens que têm funções externamente identificáveis e que, uma vez isoladas, podem esclarecer as razões pelas quais essas mesmas estruturas são mobilizadas em contextos diferentes do estatal.

Segundo, porque uma preocupação central da Teoria dos Sistemas é aplicar essa perspectiva funcionalista ao direito, explicando, em específico, o papel desempenhado por elementos característicos de sua estrutura, como o positivismo<sup>49</sup>, a justiça<sup>50</sup> e, para os fins desse trabalho, o procedimento decisório judicial<sup>51</sup>. É, portanto, dessa possibilidade de se extrair da Teoria conclusões já existentes e que resultaram da aplicação de suas ferramentas no fenômeno decisório jurisdicional – explicando funcionalmente o papel que o devido processo legal tem no direito moderno – que espero conseguir estabelecer hipóteses quanto ao papel que essas mesmas estruturas têm em contextos decisórios privados, como é o caso da moderação de conteúdo.

Terceiro, adotar a Teoria dos Sistemas é uma estratégia justificável porque enfrentar o problema que proponho neste trabalho exige uma análise de transformações que ocorrem em um espaço de manifestação naturalmente transnacional: as comunidades criadas por plataformas como o Facebook, o Twitter e o YouTube na internet. Isso se traduz não apenas em uma mudança do espaço de decisão – que, como argumentei acima, foi do juízo às plataformas, via moderação de conteúdo –, mas na exigência de ferramentas teóricas capazes de lidar com um formato de criar regras, apurar condutas e fazer julgamentos que é essencialmente desvinculado da ação estatal.

Por isso, se, por um lado, a análise sustentada por este trabalho precisa uma sociologia jurídica que tenha uma robusta estrutura funcionalista – que, com isso, ofereça uma alternativa às abordagens que mencionei antes; por outro, ela demanda ferramentas capazes de lidar com esse novo fenômeno decisório não como algo à margem do direito tomado em sua perspectiva tradicional – isto é, estatal –, mas levando a sério essa nova forma de julgar aquilo que as pessoas podem ou não dizer nesse espaço de manifestação. Para tanto, não basta o funcionalismo – e as conclusões que ele já fornece quanto ao papel desempenhado pelas características do procedimento decisório jurisdicional: é preciso

---

<sup>49</sup> LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, pp. 207/216; LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, pp. 38 e ss.

<sup>50</sup> LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, pp. 214/238.

<sup>51</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 55/129.

acoplá-lo com alguma modalidade de entendimento plural do fenômeno jurídico, que, ao dissociar direito e Estado, permita utilizar essas ferramentas sociológicas em espaços transnacionais – como o que produz as decisões da moderação de conteúdo.

Essa transição entre uma sociologia jurídica funcionalista, mas que vê no direito um fenômeno estatal, e uma sociologia jurídica que permanece funcionalista, mas que absorve elementos do pluralismo jurídico, é justamente o que marca a Teoria dos Sistemas nos últimos anos de seu desenvolvimento. Isso porque, ainda que sob pressupostos inovadores, essa é uma teoria que nasce, com Luhmann, ainda na década de 70, em um momento anterior ao florescimento do pluralismo jurídico. Por isso, ela foi construída sobre uma compreensão do direito essencialmente vinculada ao Estado, a partir de uma associação, mesmo que implícita, entre os fenômenos jurídico e estatal. Isso desencadeou na Teoria dos Sistemas exatamente o mesmo movimento teórico que é necessário fazer, neste trabalho, para enfrentar o problema que descrevi acima: com a formação de novos espaços decisórios desvinculados do Estado, um dos principais intérpretes de Luhmann, Gunther Teubner, viu-se na necessidade de operar a transição entre um direito essencialmente vinculado ao fenômeno estatal para a possibilidade de uma criação jurídica<sup>52</sup> – e mesmo constitucional<sup>53</sup> – não-estatal.

Assim, a Teoria dos Sistemas detém, enquanto sociologia jurídica, uma estrutura marcada por uma trajetória que tanto reteve os seus aspectos funcionalistas – capazes, portanto, de fornecer interpretações quanto ao papel de institutos jurídicos como o devido processo legal – quanto inovou pela apreensão do pluralismo jurídico, nos termos em que reconstruída por Teubner – permitindo que novos fenômenos decisórios como o que me proponho a estudar não sejam vistos como periféricos, mas também constitutivos do direito. Ela ficou, em síntese, com o melhor dos dois mundos – e, por isso, junge as ferramentas adequadas para um trabalho que propõe fazer uma análise funcionalista de um fenômeno decisório transnacional.

---

<sup>52</sup> Cf., para uma abordagem inicial da questão, TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, pp. 255/290, 1996; e TEUBNER, G. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne, in SCHIMDT, E., WEYERS, H. (ed.) **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, pp. 191/214, 1995.

<sup>53</sup> Cf., também introdutoriamente, TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, pp. 1/28, 2003.

Estruturalmente, dividi o trabalho em duas partes. Na primeira, concentrei um esforço descritivo, cujo objetivo é detalhar o tema da dissertação. Diferente do que ocorre nos Estados, no caso da moderação de conteúdo a regra é a ausência de protocolos de transparência que permitam compreender como essas decisões são tomadas. Não há um Código de Processo Civil da moderação de conteúdo – e, por isso, a opacidade é, em si, um problema: é difícil mesmo se estabelecer um ponto de partida seguro com o qual seja possível elaborar problemas e hipóteses de pesquisa.

Por isso, o primeiro momento deste trabalho será integralmente dedicado a superar a característica – e problemática – opacidade da moderação de conteúdo, construindo uma base conceitual sobre ela. Ele está dividido em dois capítulos: no primeiro, justifico e posiciono uma pesquisa enfocada na moderação de conteúdo. Faço isso a partir da constatação de que análises exclusivamente voltadas às ferramentas regulatórias do Estado têm um problema: elas não conseguem explicar adequadamente casos em que as empresas de tecnologia por trás de plataformas na internet – como o Facebook, o Twitter e o YouTube – decidem, de forma independente – e paralela ao Estado –, controvérsias quanto àquilo que as pessoas podem ou não dizer nesses ambientes. Sustento, nesse sentido, tanto a dificuldade dessas perspectivas tradicionais de compreenderem esses mecanismos decisórios – isto é, a moderação de conteúdo –, quanto a necessidade de abordá-los, fazendo-o para apreender adequadamente como decisões que impactam a liberdade que as pessoas têm de se expressar são tomadas, nessas plataformas, por agentes privados.

No segundo capítulo, abordo e sistematizo a moderação de conteúdo como uma nova estrutura regulatória da liberdade de expressão. Proponho tanto conceitos capazes de delimitar as suas diferentes vertentes, quanto uma taxonomia das formas pelas quais essa atividade decisória ocorre, destrinchando-a em uma classificação que lança luz, por exemplo, sobre aquilo que aconteceu em casos como o de Lipman. Em seguida, apresento com mais detalhes as formas pelas quais a moderação de conteúdo tem sido compreendida pela literatura especializada, fazendo-o a partir das duas correntes que predominam nesse debate.

Reunidos, portanto, o primeiro e o segundo capítulo explicam tanto *o que é* a moderação de conteúdo – resolvendo, assim, um problema de opacidade –, quanto *como* a moderação de conteúdo é atualmente percebida – fornecendo, por essa via, um retrato dos aportes e das limitações das vertentes disponíveis para compreendê-la.

Na segunda parte do trabalho, concentrei um esforço teórico para construir e aplicar à moderação de conteúdo uma estrutura baseada na sociologia jurídica que permita superar as dificuldades que as formas disponíveis têm de interpretá-la. No terceiro capítulo do trabalho, apresentei essas dificuldades: inicialmente, exponho a transformação pela qual passa a moderação de conteúdo, transitando de um procedimento arbitrariamente definido pelas plataformas para um modelo decisório que, cada vez mais, apreende elementos da forma pela qual decisões judiciais vinculantes são produzidas nos Estados. Em seguida, demonstro a insuficiência das abordagens que temos para entender esse movimento e proponho, por isso, uma perspectiva diferente desse problema, sustentada na Teoria dos Sistemas.

Para isso, faço, no quarto capítulo, uma reconstrução em que analiso tanto uma versão original da análise funcionalista relativamente ao papel desse procedimento decisório, quanto a necessidade de, operando uma transição a partir de Teubner, atualizá-la para que faça sentido abordar com ela questões decisórias vinculadas a estruturas jurídicas não-estatais, como é o caso da moderação de conteúdo.

No quinto capítulo, complemento essa versão original da Teoria dos Sistemas com a abordagem de Teubner, aplicando-a à moderação de conteúdo e ao seu movimento de transformação para testar as hipóteses que descrevi acima. Em seguida, ofereço quatro perspectivas de continuidade do trabalho, associadas ao movimento de constitucionalização material das plataformas; à possibilidade de que novos instrumentos do direito sejam por elas adaptados – a exemplo dos mecanismos de legitimidade democrática; à intensificação da colisão entre regimes transnacionais e estatais; e à reintrodução de divisões políticas e econômicas locais no ambiente digital. O trabalho se encerra com uma conclusão, em que retomo os pontos trazidos nesta introdução a partir dos resultados produzidos ao longo de seu desenvolvimento.

Quanto à forma de apresentação da dissertação, para tornar a sua exposição a mais clara possível, optei por escrevê-la em primeira pessoa, privilegiando a forma direta do discurso e evitando, na medida do possível, interrupções que prejudiquem a fluidez do texto. Como, por um lado, as referências estão frequentemente acompanhadas de notas explicativas a respeito de seu conteúdo – que, por vezes, são extensas; e, por outro, traduzi – tanto de forma livre, se não disponível versão em português, quanto me reportando às versões traduzidas, se disponíveis – todos os textos em língua estrangeira, apresentando o original nas notas de rodapé para cotejo, optei por adotar o estilo “*notas de referência*” e não o modelo “*autor-data*”. Por isso, as citações foram referenciadas de forma completa



em notas de rodapé, recurso que também utilizei para as notas explicativas, conforme a NBR n.º 10.520. No final do trabalho, há uma lista consolidando todas as referências consultadas.

## PARTE I | POR UMA DESCRIÇÃO

### CAPÍTULO 1 – A MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET COMO UMA FORMA DE REGULAÇÃO DA LIBERDADE DE EXPRESSÃO

#### 1. INTRODUÇÃO

A primeira parte deste trabalho tem dois objetivos. O primeiro, ao qual dedico este capítulo, é o de justificar o tema da dissertação: a moderação de conteúdo das plataformas da internet. O segundo, tratado no capítulo seguinte, é o de destrincha-lo, lançando luz à opacidade que é típica desse mecanismo decisório. Como justifiquei na Introdução, construir esse acesso inicial é indispensável porque as características da moderação de conteúdo – em particular a ausência de protocolos de transparência nas plataformas da internet – fazem da opacidade um problema em si, na medida em que ela inviabiliza a formação de uma base comum sobre a qual problemas e hipóteses de pesquisa mais complexos possam ser construídos. É preciso, portanto, antes de problematizar a moderação de conteúdo, compreender o que ela é e por que faz sentido escrever uma dissertação sobre isso.

Não há uma forma única de responder a essas perguntas. Em particular, existem diversas maneiras de explicar o que é a moderação de conteúdo<sup>54</sup>. Na introdução<sup>55</sup>,

---

<sup>54</sup> Cf., a esse respeito, diferentes abordagens que introduzem a moderação de conteúdo como uma nova forma de governança da internet (KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1616 e ss e DENARDIS, L., HACKL, A. M. Internet governance by social media platforms. **Telecommunications Policy**, v. 39, n. 9, pp. 761/770, 2015); um instrumento de censura privada (HEINS, M. The Brave New World of Social Media Censorship. **Harvard Law Review Forum**, v. 127, n. 8, pp. 325/330, 2014) e, em contextos não acadêmicos, mesmo como um exercício de poder dito como quase monárquico (FARREL, H., LEVI, M., O'REILLY, T. Mark Zuckerberg runs a nation-state and he's the king. **Vox**, 10.4.2018, disponível em <<https://www.vox.com/the-big-idea/2018/4/9/17214752/zuckerberg-facebook-power-regulation-data-privacy-control-political-theory-data-breach-king>>, acesso em 13.6.2019).

<sup>55</sup> Cf. *supra*, Introdução, “Durante um simpósio dedicado à liberdade de imprensa e organizado pela Harvard Law Review em 2014, um dos convidados, Marvin Ammori [...]”, uma referência ao texto AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, pp. 2.259/2.295, 2014.

por exemplo, optei por uma abordagem resumida, expondo-a como uma espécie de necessidade decisória, que emergiu do modelo de negócio praticado por empresas como o Facebook, o Twitter e o Google. Neste capítulo, minha estratégia será diferente: minha intenção não é abordar a moderação de conteúdo como a solução de um problema empresarial dessas companhias, capaz de garantir que continuem lucrativas<sup>56</sup>. Antes, quero contextualizá-la em um quadro maior: o de como mudanças tecnológicas impactaram a regulação da liberdade de expressão de uma maneira geral.

Isso me permitirá, por um lado, abordar as mudanças sociais que estão por trás do surgimento da moderação de conteúdo, explicando-a como consequência de uma transformação muito mais ampla, que diz respeito fundamentalmente às novas dinâmicas de poder atuantes nos espaços em que as pessoas se manifestam. Por outro, espero expor, com essa forma de apresentar a moderação de conteúdo, as razões que justificam fazer uma dissertação com esse tema, isto é, enfocada na regulação da liberdade de expressão em um contexto essencialmente não estatal. Faço isso demonstrando que abordagens que tratam a ação do Estado como a única fonte a partir da qual os impactos da era digital para a regulação da liberdade de expressão podem ser diagnosticados fornecem resultados incompletos, na medida em que deixam de lado a verdadeira transformação que decorreu dessa evolução tecnológica. Isso porque, como exponho a seguir, a “*era digital*” que vivemos não se traduziu apenas na criação de novos *espaços de manifestação* – plataformas na internet em que as pessoas se manifestam tal como faziam nas praças públicas –, mas também significou a emergência de novos *espaços de decisão* – estruturas empresariais de moderação de conteúdo que exercem poder sobre esses ambientes, decidindo aquilo que as pessoas podem ou não dizer de forma relativamente independente – e que, por isso, demandam uma análise capaz de apreender as suas especificidades em relação à forma estatal de se regular a liberdade de expressão.

---

<sup>56</sup> Cf., por exemplo, a leitura que vê as regras que orientam a exclusão de conteúdos nessas plataformas também como o resultado das pressões econômicas que essas empresas sofrem, como em GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 19/20 (Kindle, n. p.). (“*A platform is a product of the company that runs it, so there is a certain logic that it should be the company’s values and interests that determine what is acceptable and what should be removed. [...] For social media platforms, what ends up standing as “our values” is not some moral core that exists beneath these many competing pressures. It is whatever solution can resolve those pressures—perhaps presented in a language of ‘the right thing to do’ but already accounting for the competing economic and institutional demands these platforms face*”).

Minha intenção, com isso, é mostrar como dessa mudança social tecnológica fez-se a mudança regulatória – estatal ou não –, em uma abordagem que, insisto, não é necessária, mas se mostrará conveniente, porque mais do que *apresentar* o tema deste trabalho, *posiciona-o e justifica-o* em uma esquematização mais ampla, que ajuda esclarecer como a moderação de conteúdo funciona. Deixarei, nesse sentido, para o próximo capítulo – momento em que o recorte temático da dissertação já estiver justificado – uma abordagem conceitual e taxonômica da moderação de conteúdo, concentrando-me, primeiro, em contextualizá-la no conjunto de mudanças sociais que a ocasionaram.

Para fazer isso, trabalharei, ao longo deste capítulo, a transição feita por um dos autores de referência na área: Jack Balkin. Balkin escreveu, em 2014, o texto *Old-School/New-School Speech Regulation*<sup>57</sup>, um artigo paradigmático<sup>58</sup> em que ofereceu uma contribuição para a compreensão dos impactos da mudança tecnológica na regulação da liberdade de expressão.

Na época, ele se dedicou a construir uma esquematização da atuação regulatória nesse novo cenário tecnológico intencionalmente restrita ao Estado, porque, segundo a posição que adota no texto, a regulação da liberdade de expressão seria uma

---

<sup>57</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.296/2.342.

<sup>58</sup> O texto de Balkin é uma constante em artigos que tratam da regulação da liberdade de expressão na internet. Exemplificadamente, ele é utilizado ao menos como um ponto de partida em CITRON, D. Extremist Speech, Compelled Conformity and Censorship Creep. **Notre Dame Law Review**, v. 93, n. 3, pp. 1.035/1.072, 2018, nota n.º 23, p. ex.; DENARDIS, L., HACKL, A. M. Internet governance by social media platforms. **Telecommunications Policy**, v. 39, n. 9, pp. 761/770, 2015 (p. 762, p. ex.); GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, 289 pp. (Cap. 7, nota 18, p ex.; Kindle, n. p.); e WU, T. Is The First Amendment Obsolete? **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2017, notas n.º 2 e 28, p. ex., disponível em <<https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>>, acesso em 13.6.2019. Em específico, Kate Klonick, tensiona a abordagem de Balkin no exato mesmo ponto que faço neste capítulo, isto é, a extensão da ideia de regulação da liberdade de expressão para além dos atores estatais; diferente dela, entretanto, argumento a seguir que a questão não é de coloquialismo, mas sim de precisão conceitual em uma nova realidade tecnológica (“*Where I do use ‘regulate,’ I do so in a more colloquial sense and not the way in which Professor Jack Balkin uses the term ‘speech regulation,’ which concerns government regulation of speech or government cooperation, coercion, or partnership with private entities to reflect government ends*”) KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.601, nota n.º 11.

atividade exclusivamente estatal. Quatro anos depois, entretanto, Balkin escreve outro artigo – *Free Speech is a Triangle*<sup>59</sup> –, em que, deixando de lado a restrição do texto anterior, sustenta a existência de um “*modelo pluralista de regulação da liberdade de expressão*”<sup>60</sup>, pelo qual essa atividade regulatória tanto acontece entre Estados e particulares, quanto entre empresas – como as que estão por trás de plataformas da internet – e particulares, as quais teriam constituído uma espécie de “*governança privada*” da liberdade de expressão, pela qual fazem valer “*os valores e normas que representam as suas comunidades*”<sup>61</sup>.

Neste capítulo, minha proposta é utilizar essa lacuna entre os textos – que, de início, parece arbitrária – para explicar e posicionar a moderação de conteúdo das plataformas da internet. Quero expô-la, nesse sentido, como uma consequência das mudanças tecnológicas inicialmente detectadas pelo Balkin de 2014 e que justificam o abandono de sua posição inicial, fazendo-o em nome da abordagem pluralista da regulação da liberdade de expressão do Balkin de 2018. Sustento, assim, que a realidade da moderação de conteúdo tanto expõe as deficiências de uma abordagem como a do Balkin de 2014, quanto justifica um trabalho – como o que proponho aqui – que trilha o caminho indicado pelo Balkin de 2018.

Minha intenção, portanto, é fazer uma reconstrução que conecte essas duas versões do mesmo autor, utilizando-a para justificar este trabalho. Farei isso, de início, acessando o texto *Old-School/New-School Speech Regulation* e o utilizando como um exemplo da abordagem que usualmente é feita quanto à regulação da liberdade de

---

<sup>59</sup> BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, 2018, pp. 2.011/2.056. As mesmas ideias estão presentes no texto *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, publicado por Balkin também em 2018. Ao longo deste capítulo, farei, por uma questão de brevidade, referência apenas a *Free Speech is a Triangle*, mas as afirmações são extensíveis também a esse segundo trabalho. Cf., a esse respeito, BALKIN, J. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **UC Davis Law Review**, v. 51, 2018, pp. 1.182/1.199, em específico o seguinte trecho: “*This brings us to the third feature of new school speech regulation: private governance. Private governance means that the infrastructure provider governs the flow of information through the infrastructure that it owns, and it governs the behavior of the end-users and customers who employ the digital infrastructure*” (p. 1.182).

<sup>60</sup> BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, 2018, p. 2.014.

<sup>61</sup> “[...] *these companies learned that they had to govern—that is, promulgate and enforce the values and norms that their communities stood for*”. BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, 2018, p. 2.018.

expressão, que se restringe, via de regra, a uma análise de como a ação do Estado impacta o direito de particulares (2).

Em seguida (3), apresentarei a moderação de conteúdo (3.1), introduzindo-a a partir de três casos concretos (3.2), que servem de contraexemplos a abordagens como a do Balkin de *Old-School*, porque expõem que, na era digital, também atores privados – no caso, as empresas que administram as plataformas da internet – exercem poder sobre os espaços em que as pessoas se expressam, tomando decisões quanto àquilo que podem dizer e enfrentando, por isso, controvérsias quanto ao direito à liberdade de expressão. O próximo passo (3.3) será identificar na moderação de conteúdo o mecanismo decisório que operacionaliza essa tomada de decisão, problematizando a abordagem inicial do Balkin de 2014 (3.3.1 a 3.3.4) a partir da necessidade de se analisar também esses meios regulatórios privados para que seja possível ter um cenário suficientemente completo da dinâmica regulatória implantada pela era digital. Por fim (3.3.5), apresento o Balkin de *Free Speech is a Triangle* como uma solução a esse problema, finalizando a reconstrução dessa transição para justificar o que espero fazer neste trabalho: percorrer o caminho tomado como inexistente pelo Balkin de 2014, mas sinalizado pelo de 2018, no sentido de se analisar a estrutura constituída por agentes privados – notadamente as plataformas da internet – para regular a liberdade de expressão. Finalizo o capítulo (4), sintetizando as suas conclusões e abrindo caminho para o seguinte, em que sistematizada a moderação de conteúdo e as perspectivas atuais que temos para compreendê-la.

## 2. DA MUDANÇA SOCIAL À MUDANÇA REGULATÓRIA: BALKIN E A REGULAÇÃO ESTATAL DA LIBERDADE DE EXPRESSÃO

Como mudanças tecnológicas impactam a regulação da liberdade de expressão? Nesta seção, minha intenção é reconstruir a resposta fornecida por Jack Balkin a essa pergunta no texto *Old-School / New-School Speech Regulation*, um trabalho de referência no tema<sup>62</sup>. Fiz essa escolha, por um lado, porque Balkin apresenta e sintetiza, sob o conceito de infraestrutura da liberdade de expressão, a *mudança social* que está por trás da moderação de conteúdo – em síntese, a emergência de novos espaços de manifestação associados às tecnologias que viabilizaram a internet, consolidadas sob a

---

<sup>62</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.296/2.342; cf. *supra*, nota n.º 52.

ideia de uma “*era digital*”. Por outro, porque o texto deixa claro o argumento de Balkin de que dessa *mudança social* fez-se uma *mudança regulatória*, restrita, segundo ele, às transformações pelas quais passaram os instrumentos que o Estado utiliza para regular a liberdade de expressão. Balkin abandona essa posição quatro anos mais tarde, em nome de uma abordagem que, pelas razões que exponho a seguir, adoto neste trabalho. No que segue deste capítulo, utilizo esse percurso como uma forma de justificar a perspectiva que proponho na dissertação.

Balkin inicia seu artigo com uma constatação<sup>63</sup>: é um erro compreender um contexto em que há liberdade de expressão como o resultado apenas da ausência de proibições quanto àquilo que se pode dizer. Expressar-se sobre algo – seja, por exemplo, por meio de um jornal que circula localmente, um filme que está em cartaz pelo país ou uma pintura exposta em uma galeria – depende também de um contexto institucional que viabiliza que aquilo que as pessoas querem dizer seja veiculado socialmente – isto é, que uma notícia seja publicada em um jornal, que um filme seja gravado ou que um artista confeccione e exponha um quadro. Assim, “*o The New York Times da metade do século 20 que aparece nos casos Sullivan e Pentagon Papers não era um conjunto de páginas com tinta*”<sup>64</sup>, mas “*o produto cumulativo de editores, repórteres, salas de notícia, escritórios, agências de notícias, impressoras, sindicatos, caminhões de entregas e serviços de assinatura; e também dependia de um conjunto maior de negócios, arranjos contratuais, costumes e convenções para produzir ‘as notícias prontas para serem impressas’*”<sup>65</sup>.

É essa estrutura institucional – que perpassa qualquer tipo de manifestação – que Balkin denomina infraestrutura da liberdade de expressão. Ele a define, em um texto anterior, como “*os tipos de meios de comunicação e instituições de conhecimento, criação*

---

<sup>63</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.301.

<sup>64</sup> “[...] *the New York Times of the mid-twentieth century featured in Sullivan and Pentagon Papers was not simply a set of pages with ink*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.301.

<sup>65</sup> “[...] *the cumulative product of editors, reporters, newsrooms, bureaus, wire services, printing machines, labor unions, delivery trucks, and subscription services; and it too depended on a larger set of businesses, contractual arrangements, customs, and conventions to produce ‘[a]ll the news that’s fit to print’*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.301.

*e disseminação que estão disponíveis em um dado momento do tempo*”<sup>66</sup>, isto é, uma espécie de estrutura socialmente construída que, previamente a um determinado “*resultado final*” comunicativo – uma notícia, um filme, um post no Facebook –, viabiliza-o ao oferecer as condições materiais para a sua existência.

Constatar isso é algo útil por duas razões. Primeiro, porque identificar a infraestrutura por trás das formas de expressão esclarece como regulações que afetam esse contexto são capazes de impactar a liberdade de expressão. Como aponta Balkin, frequentemente as investidas contra a liberdade de expressão – e, por consequência, os casos judiciais que decorrem dessas situações – não dizem respeito a proibições que impactam manifestações específicas já produzidas. É claro que isso também ocorre: assim como no caso *New York Times Co. v. United States (“Pentagon Papers”)*<sup>67</sup> discutia-se a proibição governamental de distribuir de documentos tidos como secretos pelo Estado, no caso *Ellwanger*<sup>68</sup> se decidiu sobre possibilidade de se publicar um material com viés antissemita. Existem casos, entretanto, que dizem respeito à proteção da liberdade de expressão, mas em que a controvérsia não é se uma determinada manifestação deve ser proibida ou autorizada, mas sim em que medida condutas governamentais que afetam o contexto institucional que viabiliza às pessoas se manifestarem – isto é, a infraestrutura da liberdade de expressão – devem ser permitidas ou não:

*“Em Hague v. CIO, por exemplo, o governo tentou evitar a reunião das pessoas negando aos manifestantes sindicais o acesso às ruas públicas e aos parques; em Schneider v. State,*

---

<sup>66</sup> “[...] *the kinds of media and institutions for knowledge, creation and dissemination that are available at any point in time.*” BALKIN, J. The Future of Free Expression in the Digital Age. **Pepperdine Law Review**, v. 36, 2009, p. 432.

<sup>67</sup> *New York Times Co. v. United States*, 403 U.S. 713 (1971). No caso, um funcionário do governo dos Estados Unidos que trabalhou na confecção de um relatório quanto à atuação militar do país na Indochina no pós-Segunda Guerra Mundial – em específico, na Guerra do Vietnã – entregou parte desses documentos a um jornalista do *The New York Times*, que começou a publicá-los em 13 de junho de 1971. Três dias depois, o Departamento de Justiça dos Estados Unidos obteve uma ordem restritiva prévia em uma Corte Local, que proibiu a publicação dos documentos sob o fundamento de que seriam secretos e poderiam gerar danos imediatos e irreparáveis ao país. A ordem veio a ser revertida no precedente da Suprema Corte, que fixou que a existência de limitações prévias à publicação de um material “*carrega o pesado fardo de justificar a imposição desse tipo de restrição*”.

<sup>68</sup> BRASIL. Supremo Tribunal Federal. Acórdão no Habeas Corpus n.º 82.424/RS. Relator: ALVES, M. Publicado no DJ de 19.3.2004, p. 24. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=2052452>>, acesso em 13.6.2019.



*o Estado de New Jersey tentou banir a distribuição de folhetos; em Grosjean v. American Press Co., o Estado da Louisiana tentou tributar jornais*<sup>69</sup>.

Isso também é comum no Brasil: no julgado que estendeu a imunidade tributária aos livros digitais, por exemplo, o Supremo Tribunal Federal baseou sua decisão em um argumento a favor da liberdade de expressão. O precedente fixou que se a proteção desse direito depende da não incidência de impostos sobre livros físicos, então cobrá-los de livros digitais – onerando, de igual modo, quem os adquire – significa afetar a liberdade de expressão da mesma maneira. Se a “*teleologia da imunidade contida no art. 150, VI, d, da Constituição, aponta para a proteção de valores, princípios e ideias de elevada importância, tais como a liberdade de expressão, voltada à democratização e à difusão da cultura*”, então “*O corpo mecânico não é o essencial ou o condicionante para o gozo da imunidade, pois a variedade de tipos de suporte (tangível ou intangível) que um livro pode ter aponta para a direção de que ele só pode ser considerado como elemento acidental no conceito de livro*”<sup>70</sup>. Da proteção a esses elementos prévios, mas indispensáveis à manifestação propriamente dita, fez-se a tutela da própria liberdade de expressão.

Portanto, proibir diretamente alguém de dizer algo não é única forma de se regular a liberdade de expressão. Como esse é um direito que pressupõe uma estrutura material – uma infraestrutura da liberdade de expressão – para ser exercido, tomar decisões quanto a esse pano de fundo institucional também é uma forma, mesmo que indireta, de regulá-lo.

A segunda razão que torna o conceito de infraestrutura da liberdade de expressão útil é a sua capacidade de viabilizar uma análise que relacione, de um lado, *mudança social* e, de outro, *mudança regulatória*. Isso porque quando controvérsias

---

<sup>69</sup> “*In Hague v. CIO, for example, the government sought to prevent assembly by denying labor protesters access to public streets and parks; in Schneider v. State, New Jersey attempted to ban the distribution of handbills; in Grosjean v. American Press Co., Louisiana sought to tax newspapers*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.303. Exemplificadamente, no caso Grosjean v. American Press Co., 297 U.S. 233 (1936), a Suprema Corte entendeu ser inconstitucional estabelecer um imposto sobre jornais que tinham uma tiragem superior a 20 mil exemplares. O tributo havia sido criado pelo Governador da Louisiana, com a intenção de restringir a circulação de notícias críticas a ele – veiculadas, à época, em jornais de grande circulação no meio urbano.

<sup>70</sup> BRASIL. Supremo Tribunal Federal. Acórdão no Recurso Extraordinário n.º 330.817/RJ. Relator: TOFFOLI, D. Publicado no DJe de 27.11.2017. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=1984213>>, acesso em 16.6.2019.

quanto à liberdade de expressão são tomadas apenas sob a perspectiva da existência ou não de uma proibição governamental à divulgação de algo, corre-se o risco de descontextualizá-las do conjunto de mudanças tecnológicas que, ao longo do tempo, impactaram a forma como as pessoas veiculam suas manifestações e, conseqüentemente, no modo como Estados tentam impedir que isso aconteça.

Com isso, não quero dizer que os argumentos jurídicos utilizados para tomar uma decisão sobre a liberdade de expressão em um caso sobre manifestações publicadas em um jornal – como em *New York Times v. Sullivan*<sup>71</sup> – não possam ser corretamente utilizados quando se discute, por exemplo, o acesso das pessoas a material pornográfico na internet – *Reno v. ACLU*<sup>72</sup> – ou o direito de um condenado participar de uma rede social na internet – *Packingham v. North Carolina*<sup>73</sup>. Essa é uma questão à parte, que não é, ao menos diretamente, o tema deste trabalho<sup>74</sup>. Se interessa saber, entretanto, em que medida

---

<sup>71</sup> *New York Times Co. v. Sullivan*, 376 US 254 (1964). Cf. *supra*, nota n.º 12. Uma descrição de um caso contextualizado com a era digital pode ser encontrada em KLONICK, K. *Facebook v. Sullivan*. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, 19 pp., disponível em <<https://knightcolumbia.org/content/facebook-v-sullivan>>.

<sup>72</sup> *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997). No caso, a Suprema Corte declarou inconstitucional dispositivos do *Communications Decency Act*, a primeira lei federal destinada à regulamentação da pornografia na internet. Em específico, a Corte entendeu que a proibição ao envio de material “*obsceno ou indecente*” a menores de idade, como posta na legislação “*lacks the precision that the First Amendment requires when a statute regulates the content of speech*” (“*não tem a precisão que a Primeira Emenda requer quando uma lei regula o conteúdo da expressão*”). Para uma análise a esse respeito, cf. LESSIG, L. What Things Regulate Speech: CDA 2.0 v.s. Filtering. **Jurimetrics**, v. 38, n. 4, pp. 629/670, 1998; também, LESSIG, L. The Law of the Horse: What Cyberlaw Might Teach. **Harvard Law Review**, v. 113, n. 2, 1999, p. 517; e KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.660/1.661.

<sup>73</sup> *Packingham v. North Carolina*, 582 U. S. \_\_\_\_ (2017). Cf. *supra*, nota n.º 7 e WHITNEY, H. Search Engines, Social Media, and the Editorial Analogy. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, p. 23, disponível em <<https://knightcolumbia.org/content/search-engines-social-media-and-editorial-analogy>>.

<sup>74</sup> Como esclareço no Capítulo 3, este é um trabalho que propõe uma abordagem sociológica a um formato regulatório da liberdade de expressão que surge com a era digital. Ele toma, portanto, essa mudança social não sob uma perspectiva interna – que é, via de regra, a uma demanda por uma decisão e por isso tenta encontrar os argumentos jurídicos que melhor se aplicam a essa nova hipótese –; mas sim sob uma perspectiva externa – em que o que se questiona é como essa nova dinâmica de se decidir funciona, independentemente de sua adequação a um código normativo específico. Sobre a diferença entre perspectiva externa e interna que adoto neste trabalho, cf. a distinção que Luhmann estabelece entre teoria do direito e

determinadas mudanças tecnológicas impactaram a forma como a liberdade de expressão é regulada – pergunta com a qual comecei essa seção – então é fundamental que exista um conceito capaz de conectar a primeira – a *mudança social* – à segunda – a *mudança regulatória*.

É exatamente isso que Balkin faz com o conceito de infraestrutura da liberdade de expressão, em um argumento que se desdobra em dois passos. No primeiro, ele identifica quais foram as mudanças recentes na infraestrutura da liberdade de expressão capazes de gerar impacto na forma como esse tipo de questão é regulada pelos governos – expondo, assim, a *mudança social*. No segundo, ele detecta, a partir dessas modificações, como a regulação da liberdade de expressão se transformou nesse mesmo período – depreendendo a *mudança regulatória*.

O primeiro passo de Balkin parte de um recorte específico: as transformações tecnológicas operadas pelo que denomina ser a “*era digital*”. Balkin utiliza esse termo como uma referência a um conjunto de inovações tecnológicas que impactaram, em específico, as telecomunicações, transformando o modo como as pessoas se manifestam. Para ele, “*a marca registrada da era digital é uma revolução na infraestrutura da liberdade de expressão*”<sup>75</sup> e, por isso, para compreender as alterações desse modelo regulatório é preciso analisar como a era digital transformou a infraestrutura da liberdade de expressão, segundo argumenta, ao viabilizar plataformas na internet pelas quais a produção e a transmissão da informação é democratizada<sup>76</sup>.

Balkin sustenta, nesse sentido, que a era digital significou uma radical transformação no contexto mobilizado pelas pessoas para se expressarem. Se antes essa infraestrutura era identificável a partir do que viabiliza ao *The New York Times* produzir e

---

sociologia jurídica (“*Dies liegt besonders im Verhältnis von Rechtswissenschaft und Soziologie nahe. Für die Rechtswissenschaft geht es um eine normative Ordnung. Für die Soziologie je nach Theorierichtung um soziales Verhalten, Institutionen, soziale Systeme - also um etwas, was so ist, wie es ist, und allenfalls zur Prognose und Erklärung herausfordert*” [“*Essa [questão] aparece de forma especialmente clara na relação entre Teoria do Direito e Sociologia. Para a Teoria do Direito, há uma ordem normativa. Para a sociologia a preocupação, a depender da perspectiva teórica, é com as relações sociais, as instituições, os sistemas sociais – coisas que são o que são e, no máximo, exigem [herausfordert] prognósticos ou explicações*”] LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, p. 14).

<sup>75</sup> “[...] *the hallmark of the digital age is a revolution in the infrastructure of free expression*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.303.

<sup>76</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.301.

distribuir jornais, agora ela está associada a um conjunto de tecnologias fundamentalmente ligadas à internet, tanto sob a perspectiva de uma base tecnológica que viabiliza o seu funcionamento – “o sistema de nomes de domínio (DNS), os protocolos da Internet, os padrões tecnológicos de armazenamento e transmissão de informação, o backbone da Internet, as redes e companhias de banda larga, as companhias de hospedagem na internet e os serviços de nuvem para armazenar, criar, expor e transmitir documentos”<sup>77</sup> e “as plataformas de hardware: computadores, tablets e, especialmente, smartphones, que se tornaram dispositivos de informação e comunicação voltados a todos os propósitos”<sup>78</sup> – quanto, de um modo mais direto, sob plataformas que tornam possível a manifestação das pessoas, como é o caso de redes sociais como o Facebook, o Twitter e o YouTube. Assim como a notícia publicada no jornal não era um misto de tinta e papel, mas o resultado de um tecido institucional anterior, as manifestações que surgiram no início do século XXI – a publicação no Facebook, por exemplo – não são um conjunto de informações armazenadas em um banco de dados, mas o produto de uma estrutura prévia que permitiu que esse tipo de expressão viesse à tona.

Essa nova infraestrutura, por sua vez, não está mais nas mãos de conglomerados que produzem material para jornais, revistas ou programas de televisão, mas sim na de empresas de tecnologia como o Facebook e o Google<sup>79</sup>, responsáveis por disponibilizar o necessário para que a operação dessas plataformas na internet ocorra. Além disso, e, ao contrário do que se poderia supor, essa transformação não se resume a um passo adiante quanto àquilo que se tinha disponível antes, fazendo do Google um novo *The New York Times*, distinguível apenas pelo nível de sofisticação tecnológica. Trata-se, pelo contrário, da constituição de uma infraestrutura da liberdade de expressão essencialmente diferente da que se tinha, porque voltada à disponibilização de espaços para a manifestação não daqueles que detêm o controle desses ambientes, mas de seus usuários:

---

<sup>77</sup> “[...] *the domain name system (DNS), Internet protocols, technological standards for storage and transmission of information, the Internet backbone, broadband networks and broadband companies, web-hosting companies, and cloud services for storing, creating, displaying, and transmitting documents*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.303.

<sup>78</sup> “[...] *hardware platforms: computers, tablets, and especially smartphones, which have become all-purpose information and communication devices*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.301.

<sup>79</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.302.

em um movimento captado pela ideia da Web 2.0<sup>80</sup> – e denominado, por Balkin, como “*dos editores às plataformas*”<sup>81</sup> –, companhias como o Google, o Facebook e o Twitter confeccionaram plataformas em que o conteúdo oferecido não é o produzido por elas – como acontece em um jornal –, mas sim o elaborado por qualquer pessoa que, com acesso à internet, disponha-se a compartilhar as suas próprias informações. A infraestrutura da liberdade de expressão mudou não apenas porque ela avançou à internet, mas porque ela se presta a algo diferente – disponibilizar, a praticamente qualquer pessoa que queira, um espaço de fala.

O quadro abaixo sintetiza essas conclusões, resumindo a transição de infraestruturas de liberdade da expressão em três níveis:

**Quadro 1 – Mudança social: infraestruturas de liberdade de expressão**

	<b>Infraestrutura Pré-Digital</b>	<b>Infraestrutura Digital</b>
<b>Quanto à tecnologia:</b>	Associada às mídias impressas (jornais e revistas, p. ex.) e às cadeias televisivas.	Associada à internet, em específico ao que viabiliza plataformas como Facebook, Twitter e YouTube.

<sup>80</sup> Em síntese, “*Web 2.0*” é um termo popularizado em 2005 por Tim O’Reilly e que é utilizado, desde então, para contrastar fases distintas no desenvolvimento da internet, em específico quanto à forma como o conteúdo que é disponibilizado na rede é produzido. Em um momento inicial – a Web 1.0 – a internet era constituída, de um lado, por desenvolvedores, responsáveis por produzir e disponibilizar o conteúdo e, de outro, pelos usuários, que se limitavam a acessar as informações. A Web 2.0 marca a transição desse modelo para a ideia da internet como uma plataforma em que os seus usuários atuam como produtores de conteúdo em uma comunidade virtual. Em um exemplo claro disso, a enciclopédia mais acessada da internet – a Wikipedia – não é o resultado de informações inseridas pelos desenvolvedores que a administram. Pelo contrário: eles se restringem a viabilizar uma plataforma colaborativa, configurada para que os seus próprios usuários sejam os responsáveis pela inserção e atualização de cada verbete. *Cf.*, a esse respeito, e inicialmente, a apresentação inicial da distinção em O’REILLY, T. *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. **International Journal of Digital Economics**, n. 65, pp. 17/37, 2007; uma análise sociológica quanto a ela BEER, D.; BURROWS, R. *Sociology and, of and in Web 2.0: Some Initial Considerations*. **Sociological Research Online**, v. 12, n. 5, 2007; e uma contextualização da Web 2.0 com a moderação de conteúdo, tema deste trabalho, em GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 213 (Kindle, n. p.); *cf.*, também, *infra*, Capítulo I.2, 2.

<sup>81</sup> “[...] *from publishers to platforms*.” BALKIN, J. *Old-School/New-School Speech Regulation*. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.304.

<b>Quanto àquele que se manifesta:</b>	Disponibilização de espaços de manifestação para aqueles que detêm a infraestrutura (“ <i>publishers</i> ”).	Disponibilização de espaços de manifestação para os usuários de serviços daqueles que detêm a infraestrutura (“ <i>platforms</i> ”).
<b>Quanto aos seus principais detentores:</b>	Conglomerados de mídias impressas e televisivas ( <i>The New York Times</i> e <i>CNN</i> , p. ex.)	Conglomerados de tecnologia na internet (Google, Facebook e Twitter, p. ex.).

Encapsulada a mudança social como uma transição entre infraestruturas de liberdade de expressão, a etapa seguinte – o segundo passo do argumento – é responder à pergunta: o que isso significou para a regulação da liberdade de expressão? Em que, exatamente, essa *mudança social* se traduziu em uma *mudança regulatória*? Uma aproximação inicial da resposta é construída, por Balkin, a partir da percepção de que conferir voz diretamente aos usuários dessas plataformas – suprimindo, assim, a figura do “*publisher*” – dificulta significativamente a atividade de regular aquilo que as pessoas dizem. Isso porque, no modelo de infraestrutura anterior, a existência centralizada de fontes das quais as informações se originavam – jornais, revistas, cadeias de televisão – facilitava a ação do Estado: regular a liberdade de expressão se traduzia em regular uma mídia identificável e incorporada em conglomerados enraizados em estruturas locais acessíveis aos meios coercitivos que governos usualmente têm à mão. Era suficiente, portanto, estruturar um mecanismo regulatório capaz de focar esses núcleos produtores de informação, aproveitando-se da centralização da mídia para também centralizar os mecanismos que a controlavam. A estratégia por trás dos casos *Pentagon Papers* – que se desenvolveram no início da década de 70 – é um bom exemplo disso: para conter a divulgação de documentos secretos sobre a atuação dos Estados Unidos na Guerra do Vietnã, bastava ao governo conseguir, em cortes locais, ordens restritivas contra os específicos veículos de mídia que detinham o material – no caso, o *The New York Times* e o *Washington Post*<sup>82</sup>.

Esse cenário muda com era digital e sua correspondente infraestrutura da liberdade de expressão, que transferem aos usuários das plataformas o papel de produtores da informação. Isso porque a supressão da figura do “*publisher*” retira a eficácia de uma

<sup>82</sup> O caso que descrevi *supra*, na nota n.º 67 (*New York Times Co. v. United States*, 403 U.S. 713 (1971)) é a consolidação de duas decisões divergentes, uma que proibiu a divulgação das informações, e direcionada ao *The New York Times*; e outra que a permitiu e da qual era parte o *Washington Post*.

atuação direta contra essas empresas de mídia: como afirma Balkin, “*indivíduos que disseminam o conteúdo que o Estado quer controlar podem ser anônimos ou pseudoanônimos, ou localizados além do alcance territorial dos governos*”<sup>83</sup>. Cumprir uma ordem restritiva contra o *The New York Times* – uma companhia com localização física, contas bancárias rastreáveis e responsáveis editoriais identificados – é muito mais simples do que fazê-lo para conter um conjunto não identificado de pessoas que, distribuídas pelo mundo, compartilham material utilizando estruturas disseminadas em inúmeras jurisdições.

Isso não significa, entretanto, o fim dos meios regulatórios da liberdade de expressão, mas a sua transição para um modelo diferente. Aos casos *Pentagon Papers*, nesse sentido, opõe-se a situação do *Wikileaks*<sup>84</sup> – um paralelo útil, porque, em ambos, a questão de fundo refere-se à divulgação de documentos governamentais de guerra tidos como sensíveis. No *Wikileaks*, a dificuldade de restringir diretamente o compartilhamento das informações porque elas eram transmitidas por pessoas e equipamentos localizados além do território dos Estados Unidos não levou ao fim do controle exercido sobre a liberdade de expressão, mas à construção de estratégias capazes de atingir essa nova infraestrutura. Assim, a ineficácia dos mecanismos anteriores de controle foi contrabalançada com a pressão governamental exercida sobre plataformas como o PayPal – que operacionaliza a transferência de valores pela internet –, inviabilizando que doações chegassem aos mantenedores da página<sup>85</sup>.

---

<sup>83</sup> “*Individuals who disseminate content that the state wants to control may be anonymous or pseudonymous, or located beyond the reach of territorial governments*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.304.

<sup>84</sup> O Wikileaks é uma organização internacional sediada na Suécia e que se especializou em publicar informações confidenciais de Estados e empresas, que são repassadas e veiculadas pela internet, utilizando-se tecnologias que asseguram tanto o sigilo das fontes quanto a acessibilidade do conteúdo de qualquer lugar. O caso mencionado por Balkin é o do vazamento de telegramas diplomáticos dos Estados Unidos pelo militar Chelsea Manning, conhecido como “*Cablegate*”. Cf. WIKILEAKS. **What is Wikileaks**. Disponível em <<https://wikileaks.org/What-is-WikiLeaks.html>>, acesso em 16.6.2019; BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.306 e ss. Para uma abordagem mais extensa a respeito do Wikileaks, cf. LEIGH, D., HARDING, L. **WikiLeaks: Inside Julian Assange's War on Secrecy**. London: Guardian Books, 2011, 352 pp.

<sup>85</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.328; para uma descrição mais extensa – citada por Balkin – a respeito da atuação do PayPal, cf. BENKLER, Y. *A Free Irresponsible Press: Wikileaks and the Battle over the Soul of the Networked Fourth Estate*. **Harvard Civil Rights-Civil Liberties Law Review**, v. 46, n. 2, 2011, pp. 341 e ss.

Casos como esse demonstram como mudanças na infraestrutura da liberdade de expressão impactaram os mecanismos que a regulam: elas tornaram parte das ferramentas utilizadas pelo Estado obsoletas, exigindo que novos instrumentos de controle fossem construídos. Isso se tornou possível a partir da percepção de que essa nova infraestrutura não significou a *eliminação dos intermediários* – antes identificados com as empresas de mídia tradicional –, mas sim a *criação de novos intermediários* – empresas como o Google, o Facebook e o Twitter –, que, justamente por terem à mão aquilo que viabiliza o exercício da liberdade de expressão por meio da internet, tornaram-se os alvos diretos da regulação. Assim, “*os Estados cada vez mais procuram atingir a infraestrutura digital, não apenas porque a maioria das pessoas fala por meio dela, mas também porque atingir a infraestrutura é o modo mais fácil de controle*”<sup>86</sup>. Eles fazem dos mecanismos que foram inicialmente postos como instrumentos de democratização do lugar de fala ferramentas de controle da liberdade de expressão, colocados à sua disposição por um complexo mecanismo de cooptação das empresas privadas responsáveis por manter essa infraestrutura<sup>87</sup>.

Essa mudança do foco, entretanto, exigiu modificações na própria forma pela qual se regula a liberdade de expressão. Balkin esquematiza essas mudanças por meio de dois rótulos: o primeiro, que denomina “*old-school speech regulation*”, identifica as técnicas regulatórias utilizadas na era pré-digital; o segundo, denominado de “*new-school speech regulation*”, refere às inovações adotadas por governos para controlar aquilo que as pessoas veiculam por meios tipicamente digitais, como as plataformas da internet. Ambos podem ser esquematizados em quatro eixos, que estruturam esse antes e depois regulatório da liberdade de expressão, sempre sob o pressuposto de que os mecanismos “*old-school*” e “*new-school*” são complementares entre si, coexistindo em uma realidade na qual o advento da era digital não significou o fim nem da mídia tradicional nem dos meios utilizados para regulá-la.

---

<sup>86</sup> “[...] *states increasingly target digital infrastructure, not only because most people are speaking through it, but also because targeting infrastructure is the easiest method of control*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.304.

<sup>87</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.327.



O primeiro eixo<sup>88</sup> diz respeito ao âmbito espacial da ação regulatória. Nos modelos anteriores – “*old-school*” –, os alvos da ação estatal consistiam, basicamente, (i) na base material utilizada pela mídia tradicional para viabilizar a disseminação da comunicação, isto é, “*espaços públicos, a imprensa escrita, filmes, telegrafia, telefones e as transmissões de rádio e televisão*”<sup>89</sup>; e (ii) na base humana desse tipo de atividade, notadamente jornalistas cooptados em troca de acesso irrestrito a determinadas fontes. Exemplos dessa prática podem ser encontrados no Brasil, seja no vigente esquema de concessões públicas do sistema televisivo<sup>90</sup> – em que o acesso a essa estrutura de transmissão depende de uma chancela do Estado – ou na exigência, superada pelo Supremo Tribunal Federal<sup>91</sup>, de que o exercício da profissão de jornalista fosse condicionado a um diploma na área. Em ambos os casos, o método de regulação praticado se resume ao poder que detém o Estado de restringir o acesso à infraestrutura que viabiliza a liberdade de expressão, condicionando-o ao atendimento de requisitos que podem ser utilizados para evitar posições contrárias às oficiais.

No modelo posterior – “*new-school*” –, a multiplicidade de pessoas que têm voz em um mundo digital – em que plataformas como o Facebook disponibilizam um espaço de fala para virtualmente qualquer indivíduo – torna difícil controlá-las particularmente. Se o Estado recusa a alguém a qualidade de jornalista e isso inviabiliza que essa pessoa assine uma reportagem em um jornal, basta abrir uma conta no Facebook e começar a publicar. Por isso, “*no lugar de atingir aqueles que se manifestam, os Estados*

---

<sup>88</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.308/2.314.

<sup>89</sup> “[...] *public spaces, post offices, printing press, movies, telegraphy, telephony, and radio and television broadcasting*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.308.

<sup>90</sup> Na sistemática vigente no Brasil, o serviço de radiodifusão televisivo é sujeito à autorização do Poder Executivo, com a aprovação do Senado, que outorgam concessões em número limitado e por prazo determinado aos interessados nesse tipo de serviço, que devem, periodicamente, renová-las. *Cf.*, a esse respeito, o art. 223, da Constituição Federal (“*Compete ao Poder Executivo outorgar e renovar concessão, permissão e autorização para o serviço de radiodifusão sonora e de sons e imagens, observado o princípio da complementaridade dos sistemas privado, público e estatal*”); e o art. 33, § 3.º, da Lei n.º 4.117/1962 (“*Os prazos de concessão, permissão e autorização serão de dez anos para o serviço de radiodifusão sonora e de quinze anos para o de televisão, podendo ser renovados por períodos sucessivos e iguais*”).

<sup>91</sup> BRASIL. Supremo Tribunal Federal. Acórdão no Recurso Extraordinário n.º 511.961/SP. Relator: MENDES, G. Publicado no DJ de 13.11.2009, p. 692. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=2469175>>, acesso em 17.6.2019.

*podem mirar os intermediários e os proprietários de serviços auxiliares*”<sup>92</sup>, operando desde mecanismos explícitos – como o bloqueio de sinais de internet, de celulares, ou mesmo inviabilizando o acesso a aplicativos específicos<sup>93</sup> – quanto implícitos – a censura colateral. Nesse último caso, os Estados utilizam mecanismos regulatórios que induzem os intermediários responsáveis pela infraestrutura da liberdade de expressão a adotarem políticas restritivas quanto àquilo que os usuários disponibilizam na internet, levando-os a atuar como censores por delegação. A lógica é relativamente simples: como o número de usuários que disponibilizam conteúdo é praticamente ilimitado, mas o de empresas que viabilizam isso – provedores de internet, serviços de hospedagens de sites, e plataformas com o Facebook, o Twitter e o YouTube – não é, é mais fácil exercer controle sobre os últimos, induzindo-os, a partir de estímulos regulatórios específicos, a controlar aqueles que realmente produzem o conteúdo.

Um exemplo disso é a responsabilização civil. Dependendo da legislação do país, é possível tornar esses intermediários digitais responsáveis, sob a perspectiva indenizatória, pelo que é veiculado na internet por meio deles<sup>94</sup>. Assim, de acordo com o modelo regulatório, uma vez disponibilizado um material contrário à lei, tanto usuário quanto plataforma podem ser solidariamente responsáveis pelo fato – sendo, por isso, igualmente acionáveis para indenizar o dano. Cria-se, assim, um tipo de cooperação forçada entre o Estado e empresas, na medida em que se fabricam estímulos para que as corporações adotem políticas restritivas quanto àquilo que é veiculado por meio delas, justamente para evitar um nível de responsabilização cível que inviabilize economicamente a sua operação. É exatamente em razão do poder que a censura colateral tem, como um instrumento dessa nova escola de regulação da liberdade de expressão, que companhias como o Google confessadamente<sup>95</sup> expõem, no caso dos Estados Unidos, a importância da

---

<sup>92</sup> “[...] *instead of targeting speakers, states can aim at intermediaries and owners of auxiliary services*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.308.

<sup>93</sup> Cf., por exemplo, quanto à atuação do Estado brasileiro, ABREU, J. S. Disrupting the disruptive: making sense of app blocking in Brazil. **Internet Policy Review**, v. 7, n. 3, pp. 1/16, 2018, disponível em <<https://policyreview.info/articles/analysis/disrupting-disruptive-making-sense-app-blocking-brazil>>.

<sup>94</sup> Citando exemplos da Itália e da União Europeia, cf. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.313, nota n.º 67.

<sup>95</sup> “*Ari Shahdadi, the general counsel of Tumblr, claims that ‘Tumblr would not exist, at least not as it does, without CDA 230.’ Ben Lee of Twitter told me: ‘My gosh, every possible erosion of CDA 230 that I’ve seen misses the point of 230 as the cornerstone for a functioning Internet.’*”. AMMORI, M. The ‘New’ New York

Seção 230, do *Communications Decency Act*<sup>96</sup>. A regulamentação, imuniza, em certa medida<sup>97</sup>, provedores e plataformas de ações de responsabilização civil por aquilo que um terceiro veicula por meio de suas estruturas, criando uma distinção entre o autor do conteúdo – um usuário, por exemplo, que faz uma publicação difamatória – e o veículo da internet utilizado para tanto – a empresa Facebook, que apenas o hospeda.

O segundo eixo<sup>98</sup> concerne ao âmbito temporal da ação regulatória. No modelo anterior, o principal momento restritivo quanto à liberdade de expressão era o anterior à publicação, aplicando-se restrições prévias àquilo que vinha ao público. Balkin distingue, nesse sentido, seis razões pelas quais esse é um método eficiente, mesmo atualmente, em relação à regulação posterior, isto é, à que ocorre depois da publicação. Em

---

Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.287.

<sup>96</sup> Em síntese, a Seção 230 do *Communications Decency Act* estabelece uma distinção entre aquele que veicula uma informação e o intermediário que opera a plataforma que o disponibiliza, responsabilizando o primeiro – e não o segundo – pelo seu conteúdo. Trata-se, assim, de uma espécie de imunidade dos intermediários, que, assim, podem veicular aquilo que um terceiro cria sem o risco de ser responsabilizado. Cf. 47 U.S.C. § 230, Section 230(c)(1) (“No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”); cf. AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.287 (“In a nutshell, if I defame Professor Mark Tushnet on Facebook, Twitter, or YouTube, those sites are generally not liable for my defamation. Only I am liable for it. Therefore, those companies can carry my speech without the fear of liability. Without CDA 230, speech platforms would likely have to engage in considerable prescreening of all speech on their sites or abandon their models altogether”); HEINS, M. The Brave New World of Social Media Censorship. **Harvard Law Review Forum**, v. 127, n. 8, 2014, pp. 328/329; KOSEFF, J. **The Twenty-Six Words that Created the Internet**. Ithaca: Cornell University Press, 2019, pp. 77 e ss.

<sup>97</sup> A regra geral de imunidade estabelecida, por exemplo, pela Seção 230 do CDA não é absoluta. Um exemplo disso é a legislação que protege o direito do autor nos Estados Unidos: a Seção 512 do Digital Millennium Copyright Act estabelece que intermediários – como plataformas na internet – não serão responsabilizados pelo conteúdo que os seus usuários veiculam, mesmo que ele viole direitos autorais, desde que eles aceitem retirar esse conteúdo após uma notificação não judicial. Caso isso não seja feito, o intermediário sai do que a literatura chama de “*porto seguro*” (“*safe harbor*”) e passa a poder responder pelo o que seu usuário fez. Cf. AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.290/2.293; e HEINS, M. The Brave New World of Social Media Censorship. **Harvard Law Review Forum**, v. 127, n. 8, 2014, pp. 328/329.

<sup>98</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.314/2.324.

síntese, as restrições prévias (i) abrangem uma quantidade muito maior de informação, porque quando esse tipo de ação é tomada posteriormente, apenas um volume limitado de dados é levada ao conhecimento da autoridade; (ii) mudam o ônus de se provar que uma comunicação atende àquilo que a lei exige, porque cabe àquele que comunica demonstrá-lo previamente ao governo; (iii) viabilizam métodos informais de controle administrativo, que podem estar sujeitos tanto a menos limites procedimentais que um julgamento jurisdicional quanto a um grau de ação menor pelo Poder Judiciário; (iv) podem ser aplicadas por um sistema que atua de forma não transparente, tirando a questão da discussão pública, o que facilita a regulação; (v) criam incentivos para a censura em um grau maior, porque geram uma estrutura institucional administrativa cujo trabalho é precisamente encontrar algo para ser censurado; (vi) tornam mais fácil para o governo encontrar aquilo que o prejudica, na medida em que cabe àquele que quer se manifestar trazê-lo ao conhecimento do Estado, dispensando uma postura mais ativa do censor.

As restrições prévias, nesse sentido, não são abandonadas por esse novo modelo de regulação da liberdade de expressão, mas têm o seu conjunto de ferramentas adequados à tecnologia digital, “*para atingir os mesmos custos e os mesmos efeitos de alteração de ônus da restrição prévia por técnicas que não envolvem, necessariamente, por um lado, a revisão administrativa ou burocrática; e, por outro, medidas judiciais*”<sup>99</sup>. Balkin cita cinco exemplos de técnicas que são utilizadas por Estados no mundo digital para isso: (i) a filtragem (“*filtering*”), em que dados que identificam determinadas comunicações ou comunicadores na internet são utilizados para bloqueá-los, impedindo, por exemplo, que páginas sejam acessadas; (ii) a apreensão de nomes de domínios (“*domain name seizures*”), em que o Estado assume o controle de páginas na internet; (iii) ações judiciais que induzem à filtragem ou à censura colateral (“*Injunctions Designed to Induce Filtering and/or Collateral Censorship*”), em que, ao contrário do método anterior, que enfocava na utilização de instrumentos judiciais contra os próprios responsáveis pela disponibilização do conteúdo – como no caso *Pentagon Papers*<sup>100</sup> –, ações são ajuizadas contra as companhias que são responsáveis pela infraestrutura que viabiliza essa disponibilização na internet, pressionando-os – sob pena de responsabilização cível e

---

<sup>99</sup> “[...] *to achieve many of the same cost and burden-shifting effects of prior restraint through techniques that do not necessarily involve administrative or bureaucratic review on the one hand, or judicial injunctions on the other*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.318.

<sup>100</sup> *Cf. supra*, nota n.º 67.

criminal – a atuar naquilo que percorre suas plataformas; (iv) restrições prévias direcionadas à infraestrutura digital local (“*Prior Restraints Directed Against the Digital Infrastructure*”), em que, à vista da impossibilidade de se atingir estruturas localizadas fora do país, o Estado toma atitudes na infraestrutura local, impedindo, por exemplo, que sites de busca vinculem o material no território nacional; e (v) restrições prévias privadas direcionadas à infraestrutura digital local (“*Private Prior Restraint*” *Directed Against the Digital Infrastructure*”), em que incentivos – como a atribuição legislativa de multas – são criadas para que entes privados, como plataformas que viabilizam doações, deixem de atuar com agentes que, localizados no exterior, sejam responsáveis pela disponibilização de informação nociva.

O terceiro eixo<sup>101</sup> enfoca as formas de cooperação em que o Estado atua com os particulares para restringir a liberdade de expressão. Mesmo no modelo anterior, os governos não atuavam sozinhos: parte importante dessa atividade regulatória era a construção de estímulos capazes de produzir ganhos recíprocos entre agentes públicos e privados, viabilizando cooperações como, por exemplo, no caso do acesso a fontes governamentais em troca de posicionamentos favoráveis na imprensa, ou, então, na indução para que particulares identificassem e isolassem veículos considerados contrários aos posicionamentos oficiais.

Novamente, modificações na infraestrutura da liberdade de expressão levaram os governos a adaptarem as ferramentas utilizadas nesse tipo de regulação, como Balkin indica em três exemplos.

O primeiro exemplo refere-se ao acesso às informações de particulares: como na era digital uma parte importante da atividade de empresas que atuam na internet envolve a coleta de dados pessoais<sup>102</sup>, a cooperação público-privado nesse tipo de regulação frequentemente se dá pela construção de estímulos – ou mesmo de obrigatoriedades – para que esses dados sejam entregues ao governo. Assim, por exemplo, “*o Estado pode oferecer às companhias privadas imunidade em troca de cooperação, de*

---

<sup>101</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.324/2.329.

<sup>102</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.325.

*acesso tecnológico e de policiar e bloquear manifestações que o governo entenda ser danosas ou perigosas*”<sup>103</sup>.

O segundo exemplo diz respeito à censura colateral: desde que cooperem com os objetivos regulatórios do governo, intermediários responsáveis pela estrutura que viabiliza as manifestações na era digital, notadamente plataformas virtuais, são beneficiados com imunidades quanto à responsabilização civil e criminal. Nos Estados Unidos, por exemplo, o *Stop Online Piracy Act*, uma legislação destinada a combater a violação de direitos autorais na internet, garante a plataformas de pagamento na internet, como o PayPal, a ausência de responsabilização, desde que, uma vez notificados, deixem de fazer negócio com suspeitos de disponibilizar conteúdo protegido por direitos autorais.

O terceiro exemplo refere-se ao uso de métodos que Balkin denomina “*soft power*”: o Estado se utiliza de meios não formais para, mobilizando o seu poder de influência, coibir a liberdade de expressão, restringindo a circulação de informações sensíveis. No caso *Wikileaks*, por exemplo, após o vazamento de uma carta do Departamento de Estado dos Estados Unidos ao Amazon, que hospedava a página do movimento – na qual se insinuava que o *Wikileaks* era o responsável pela circulação de informações sigilosas que ameaçavam a segurança nacional – e uma ligação de um Senador à companhia, “*várias partes da infraestrutura digital começaram a se negar a prestar serviço ao Wikileaks*”<sup>104</sup>.

Juntos, esses três eixos permitem distinguir um quarto eixo, relativo à visibilidade da ação regulatória pré-digital – “*old-school*” – em comparação à que ocorre no mundo digital – “*new-school*”. Segundo Balkin, marcadas pelo momento tecnológico que lhe é específico – isto é, por sua correspondente infraestrutura da liberdade de expressão –, as técnicas “*old-school*” voltaram-se ao controle de “*corpos, espaços e tecnologias pré-digitais de distribuição em massa*”<sup>105</sup>, o que se traduziu em mecanismos de controle exercidos após a disseminação da comunicação – como ações judiciais indenizatórias ou mesmo criminais contra quem disponibiliza informações

---

<sup>103</sup> “[...] *the state can offer private companies immunity in exchange for cooperation, for technological access, and for policing and blocking speech that the government finds harmful or dangerous*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.325.

<sup>104</sup> “[...] *various parts of the digital infrastructure began denying service to WikiLeaks*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.328.

<sup>105</sup> “*Old-school regulation tries to control bodies, spaces, and predigital technologies of mass distribution*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.340.

governamentais. Como essa forma de regulação tem, em si, uma eficácia limitada – porque não impede a distribuição da informação, que, nesses casos, já ocorreu –, ela é frequentemente acoplada, pelos governos, a instrumentos de aumento de visibilidade – como, por exemplo, ações na própria imprensa – que divulgam a existência desses mecanismos repressivos, dissuadindo pessoas a adotarem condutas semelhantes no futuro. Aqui, a visibilidade é uma qualidade que os Estados buscam como uma forma de reforçar e garantir, aos métodos de controle tradicional, um caráter preventivo.

Algo diferente ocorre no modelo “*new-school*”: como a infraestrutura da liberdade de expressão disponível significou um maior número de pessoas com lugar de fala, mas um menor número de intermediários – isto é, de companhias capazes de manejar a estrutura técnica necessária para fazer uma plataforma funcionar –, ficou mais fácil para o governo exercer métodos regulatórios que atuem antes da disseminação do conteúdo, como, por exemplo, a censura colateral que induz essas empresas a bloquear certas manifestações mesmo antes de elas serem disponibilizadas. Com isso – e ao contrário do que acontecia antes –, busca-se não a visibilidade, mas exatamente o contrário: uma invisibilidade que induza o usuário a disponibilizar cada vez mais informações a serem bloqueadas por essas plataformas – algo útil seja para as companhias, que têm um modelo de negócio baseado nessa disponibilização, seja para os governos que, de posse de informações que não seriam fornecidas pelos usuários caso soubessem que são vigiados, podem servir tanto para o controle da liberdade de expressão quanto, por exemplo, para a perseguição penal. A visibilidade que era uma qualidade na “*old-school*” se torna algo prejudicial às técnicas da “*new-school*”.

O quadro abaixo, em paralelo ao anterior, sintetiza a mudança regulatória de que fala Balkin, fazendo-o a partir dos eixos em que dividi acima a transição:

**Quadro 2 – Mudança regulatória: as regulações da liberdade de expressão**

	<i>Old-School</i>	<i>New-School</i>
<b><u>Primeiro eixo</u></b> Espaço da ação regulatória	Base material e humana utilizada pela mídia tradicional para viabilizar a sua própria comunicação (sistemas de concessões para estruturas de transmissão de dados e de licenciamento de jornalistas, p. ex.).	Base material e humana que intermediários utilizam para viabilizar a comunicação de terceiros (censura colateral de plataformas da internet via responsabilização civil, p. ex.).

<p><b><u>Segundo eixo</u></b></p> <p>Tempo da ação regulatória</p>	<p>Restrições preferencialmente prévias, mas nem sempre é possível fazê-lo; via de regra, os sucessos nesse âmbito dizem respeito à ação burocrático-administrativa e judicial prévia à publicação.</p>	<p>Restrições preferencialmente prévias e acopladas, com sucesso, às características tecnológicas digitais, via técnicas de filtragem, apreensão de domínios, ações judiciais, e direcionadas à infraestrutura digital privada e local.</p>
<p><b><u>Terceiro eixo</u></b></p> <p>Cooperação na ação regulatória</p>	<p>Mecanismos repressivos e retributivos (“<i>carrots and sticks</i>”), como sanções a particulares (“<i>blacklists</i>”) e cooptação da mídia em troca de favores, como o acesso a fontes.</p>	<p>Mecanismos repressivos e retributivos (“<i>carrots and sticks</i>”), como troca de imunidade por acesso a dados, estabelecimento da responsabilidade civil como um meio de censura colateral e meios não formais de pressão (“<i>soft power</i>”).</p>
<p><b><u>Quarto eixo</u></b></p> <p>Visibilidade da ação regulatória</p>	<p>Permanecer <u>visível</u>, como forma de contrabalançar a fragilidade de métodos repressivos prévios, mais eficientes mas menos disponíveis à regulação que ocorre na infraestrutura pré-digital.</p>	<p>Permanecer <u>invisível</u>, porque os meios repressivos prévios são mais efetivos nos intermediários digitais e, com isso, quanto menor a visibilidade, mais fácil obter mais informações.</p>

O argumento de Balkin de que transformações na infraestrutura da liberdade de expressão se traduziram em transformações no modelo regulatório de liberdade de expressão pode ser resumido em uma esquematização que relacione, de um lado, *mudança social*, e, de outro, *mudança regulatória*. Essa sistematização, nos termos em que fiz abaixo, entrecruza o primeiro e o segundo quadros desta seção, produzindo um terceiro capaz de relacioná-los e que deve ser lido a partir de uma advertência trazida por Balkin no artigo<sup>106</sup>: as infraestruturas de liberdade de expressão pré-digital e digital não atuam em uma lógica de exclusão, mas de complementariedade, porque a emergência de uma não depende de que a outra deixe de existir. Na realidade, tanto a infraestrutura pré-digital – relacionada à mídia impressa e televisiva – quanto a digital – afeta à internet – coexistem atualmente, o que, por consequência, significa que ambos os modelos regulatórios – “old-

<sup>106</sup> BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.340 e ss.



*school*” e “*new-school*” – vigem como formas utilizadas por governos para controlar aquilo que as pessoas dizem.

**Quadro 3 – Mudança social e mudança regulatória**

<b><u>Da mudança social ...</u></b>		
	<b>Infraestrutura pré-digital</b>	<b>Infraestrutura digital</b>
<b>Quanto ao nível de tecnologia:</b>	Associada às mídias impressas (jornais e revistas, p. ex.) e às cadeias televisivas.	Associada à internet, em específico ao que viabiliza plataformas como Facebook, Twitter e YouTube.
<b>Quanto àquele que se manifesta:</b>	Disponibilização de espaços de manifestação para aqueles que detém a infraestrutura (“ <i>publishers</i> ”).	Disponibilização de espaços de manifestação para os usuários de serviços daqueles que detém a infraestrutura (“ <i>platforms</i> ”).
<b>Quanto aos seus principais detentores:</b>	Conglomerados de mídias impressas e televisivas ( <i>The New York Times</i> e <i>CNN</i> , p. ex.)	Conglomerados de tecnologia na internet (Google, Facebook e Twitter, p. ex.).
<b><u>... à mudança regulatória</u></b>		
	<b><i>Old-School</i></b>	<b><i>New-School</i></b>
<b><u>Primeiro eixo</u></b>  Espaço da ação regulatória	Base material e humana utilizada pela mídia tradicional para viabilizar a sua própria comunicação (sistemas de concessões para estruturas de transmissão de dados e de licenciamento de jornalistas, p. ex.).	Base material e humana que intermediários utilizam para viabilizar a comunicação de terceiros (censura colateral de plataformas da internet via responsabilização civil, p. ex.).
<b><u>Segundo eixo</u></b>  Tempo da ação regulatória	Restrições preferencialmente prévias, mas que nem sempre conseguiam fazê-lo; via de regra, os sucessos nesse âmbito dizem respeito à ação burocrático-administrativa e judicial prévia à publicação.	Restrições preferencialmente prévias e acopladas, com sucesso, às características tecnológicas digitais, via técnicas de filtragem, apreensão de domínios, ações judiciais, e direcionadas à infraestrutura digital privada e local.

<p><b><u>Terceiro eixo</u></b></p> <p>Cooperação na ação regulatória</p>	<p>Mecanismos repressivos e retributivos (“<i>carrots and sticks</i>”), via estabelecimento de sanções a particulares (“<i>blacklists</i>”) e cooptação da mídia, em troca de favores, como acesso a fontes.</p>	<p>Mecanismos repressivos e retributivos (“<i>carrots and sticks</i>”), como troca de imunidade por acesso a dados, estabelecimento da responsabilidade civil como um meio de censura colateral e meios não formais de pressão (“<i>soft power</i>”).</p>
<p><b><u>Quarto eixo</u></b></p> <p>Visibilidade da ação regulatória</p>	<p>Permanecer <u>visível</u>, como forma de contrabalançar a fragilidade de métodos repressivos prévios, mais eficientes mas menos disponíveis à regulação que ocorre na infraestrutura pré-digital.</p>	<p>Permanecer <u>invisível</u>, porque os meios repressivos prévios são mais efetivos nos intermediários digitais e, com isso, quanto menor a visibilidade, mais fácil obter mais informações.</p>

### 3. A MODERAÇÃO DE CONTEÚDO NAS PLATAFORMAS DA INTERNET: UMA JUSTIFICATIVA

#### 3.1. Introdução

Não é um acaso que o texto de Balkin que reconstruí acima é uma referência na análise dos impactos do mundo digital na regulação da liberdade de expressão<sup>107</sup>. Como argumentei na seção anterior, seu mérito é construir paralelismos entre a mudança social – encapsulada no conceito de infraestrutura da liberdade de expressão – e a mudança regulatória – a transição do “*old-school*” para o “*new-school*” – e, com isso, explicar as transformações pelas quais passa a regulação da liberdade de expressão com os avanços tecnológicos da era digital.

Existe, entretanto, uma restrição que Balkin estabelece à sua abordagem, mencionada no início do artigo:

*“Eu uso o termo ‘regulação da liberdade de expressão’, que diz respeito primariamente à regulação do Estado, à ação do Estado tomada ou capaz de cooptar entes privados, ou à*

<sup>107</sup> Cf. *supra*, nota n.º 58.

*regulação do Estado que aproveita do controle privado da infraestrutura para atingir fins estatais*”<sup>108</sup>.

Dessa forma, o diagnóstico feito por Balkin tem como pressuposto a ideia de que, para entender como a regulação da liberdade de expressão é transformada no mundo digital, é suficiente compreender como a atuação dos Estados nesse âmbito mudou. É por isso que ele elabora, em síntese, uma taxonomia da atuação estatal, distinguindo-a de acordo com a infraestrutura da liberdade de expressão vigente. Assim, por exemplo, na infraestrutura pré-digital – à qual corresponde a regulação do tipo “*old-school*” –, predominavam atuações *diretas* do Estado: como ocorreu nos casos *Pentagon Papers* e *Sullivan*, regular a liberdade de expressão se traduzia em estabelecer proibições a que algo fosse divulgado, fazendo-o, por exemplo, via ações judiciais direcionadas contra veículos de imprensa. Na infraestrutura digital – em que criados novos espaços de manifestação, notadamente as plataformas na internet e à qual corresponde a regulação do tipo “*new-school*” – predominam, de modo diferente, atuações *indiretas*: o Estado cria mecanismos de pressão, como a responsabilização civil, que induzem atores privados com domínio sobre estruturas tecnológicas a atuar como seus agentes, restringindo a divulgação de conteúdos tidos como indesejáveis. A partir de transições como essas – entre, no caso, atuação *direta* e *indireta* do Estado –, Balkin oferece, dentro da restrição que estabeleceu à ideia de “*regulação da liberdade de expressão*”, uma resposta à pergunta com a qual iniciei este capítulo: transformações tecnológicas impactaram na regulação da liberdade de expressão à medida que, com a criação de novos espaços de expressão, as técnicas pelas quais o Estado age também se transformaram, adaptando-se a uma nova realidade tecnológica para que permanecessem efetivas.

Evidentemente, o Balkin de 2014 não ignora a emergência de novos atores privados no campo da liberdade de expressão – em específico, as companhias por trás das plataformas da internet –, nem o poder que a infraestrutura do mundo digital conferiu a eles. Entretanto, no contexto de sua abordagem, o papel que é atribuído a esses novos agentes é o de instrumentos de um “*Estado que aproveita do controle privado da infraestrutura para atingir fins estatais*”. Por isso, seu artigo não se propôs a investigar o

---

<sup>108</sup> “I use the term, ‘speech regulation’ concerns primarily state regulation, state action that partners with or co-opts private parties, or state regulation that leverages private control of infrastructure to achieve state ends. Practices of ‘censorship,’ by contrast, need have no connection to the state”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.299.

impacto da mudança tecnológica na atuação independente dessas companhias, mas sim o poder que o Estado exerce sobre elas. Esse enfoque fica perceptível, por exemplo, no movimento do texto que, por um lado, reconhece a existência de “*práticas de censura [...] que não precisam ter uma conexão com o Estado*”<sup>109</sup>, mas, por outro, insiste que, para analisar as transformações regulatórias desencadeadas pelo mundo digital, é suficiente identificar e sistematizar fenômenos como a censura colateral e o *soft Law* no modelo “*new-school*”. Ambos seriam capazes, por si só, de explicar as restrições impostas pelas plataformas àquilo que seus usuários podem disponibilizar – interpretando-as, nesse sentido, como os resultados visíveis da ação indutora estatal que, para proibir a divulgação de determinados conteúdos, prefere fazê-lo agindo indiretamente sobre esses intermediários do que diretamente contra aqueles que os disponibilizam<sup>110</sup>.

No que segue deste capítulo, sustento a necessidade de se fazer uma incursão pelo caminho que o Balkin de 2014 não percorreu. Argumento, nesse sentido, que é preciso fazer uma análise de como os mesmos avanços tecnológicos que desencadearam adaptações nas ferramentas regulatórias do Estado – fazendo-as transitar do modelo “*old-school*” para o “*new-school*” – transformaram o papel desempenhado por companhias privadas na regulação da liberdade de expressão. Faço isso a partir de uma perspectiva baseada inicialmente em três casos que, reconstruídos na seção seguinte, ilustram que, ao contrário do que afirma o Balkin de *Old-School/New-School Speech Regulation*, essa transição entre infraestruturas de liberdade de expressão não necessariamente transforma essas empresas em instrumentos do Estado, mas frequentemente as coloca na posição de decidir, de forma independente, controvérsias quanto à liberdade de expressão. Argumentarei, nesse âmbito, e a partir de uma reconstrução do mecanismo decisório criado para atender a essa demanda – a moderação de conteúdo –, que os impactos do mundo digital na regulação da liberdade de expressão não se traduziram apenas em mudanças nas técnicas utilizadas pelo Estado. Pelo contrário: eles também significaram a criação de novos espaços de decisão, materializados em nichos de poder privados que, localizados em companhias como o Facebook e o Google, regulam a liberdade de expressão ao determinar, com independência, aquilo que as pessoas podem ou não podem dizer em

---

<sup>109</sup> “*Practices of ‘censorship,’ by contrast, need have no connection to the state*”. BALKIN, J. *Old-School/New-School Speech Regulation*. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.299.

<sup>110</sup> BALKIN, J. *Old-School/New-School Speech Regulation*. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.320 e ss.

importantes – se não os mais importantes – espaços de manifestação atual: as plataformas da internet.

Por isso, sustentarei que um recorte que restringe a análise da regulação da liberdade de expressão, direta ou indiretamente, à ação estatal é problemático: ele retira de seu próprio alcance um conjunto de situações em que questões centrais quanto àquilo que as pessoas podem dizer são discutidas e decididas, com relativa independência, em nichos de poder privado criados justamente pela transição entre infraestruturas detectada por Balkin. Deixa-se, assim, escapar aquilo de verdadeiramente revolucionário que o mundo digital trouxe à regulação da liberdade de expressão: não a existência de novos *espaços de manifestação*, em que o Estado se vê obrigado a adaptar suas técnicas regulatórias, mas o surgimento de novos *espaços de decisão*, cuja independência demonstra tanto a insuficiência de análises como a do Balkin de 2014 – restrita, como reconstruí acima, à ação estatal e ao seu impacto nos agentes privados – quanto a necessidade de complementá-la – levando a sério a moderação de conteúdo como uma forma de regular a liberdade de expressão, como passa a fazer o Balkin de 2018 e me proponho a abordá-la nesta dissertação.

Sob uma perspectiva estrutural, dedico a próxima seção (3.2) a oferecer uma contraposição à tese de Balkin. Farei isso reconstruindo três casos em que companhias que detêm plataformas na internet – e estão inseridas, portanto, em contextos típicos da infraestrutura digital – são instadas a tomar decisões quanto àquilo que as pessoas podem ou não podem dizer em espaços de expressão virtuais, como o Facebook, o Twitter e o YouTube. Os casos foram escolhidos como contraexemplos à tese do autor, porque expõem situações nas quais o resultado desse processo decisório – ou seja, a decisão quanto à proibição ou a permissão de algo nessas plataformas – é objeto de crítica, e não de respaldo, por parte dos Estados envolvidos nas controvérsias. Ofereço, assim, indicativos de que nem sempre a ação regulatória quanto à liberdade de expressão que ocorre nesses ambientes é uma extensão da ação estatal, mas que ela também pode ser o resultado de uma política própria da plataforma, que tanto restringe, ela mesma, aquilo que os seus usuários podem dizer, quanto enfrenta, por si, controvérsias quanto ao direito à liberdade de expressão que lhe são opostas.

Em seguida (3.3), reconstruo esse mecanismo de decisão independente sob o denominador comum da moderação de conteúdo (3.3.1 e 3.3.2), explicando-o como uma via decisória privada criada para atender à demanda que é imposta, a partir da era digital, às companhias responsáveis pela administração dessas plataformas. Farei isso

problematizando a abordagem do Balkin de 2014 (3.3.3), no ponto em que ele restringe uma análise quanto à liberdade de expressão à atuação do Estado, rejeitando, em seguida, uma possível defesa da sua posição (3.3.4). Em seguida, identifico na transição feita pelo Balkin de 2018, no texto *Free Speech is a Triangle* (3.3.5), uma saída para o problema, concretizada no reconhecimento de que a era digital ocasionou novas dinâmicas de poder cuja complexidade só pode ser enfrentada por uma perspectiva pluralista da regulação da liberdade de expressão.

## 3.2 *Três casos concretos*

### 3.2.1 Introdução

Como reconstruí acima, o Balkin de 2014 se propõe a analisar os impactos das transformações tecnológicas na regulação da liberdade de expressão a partir das mudanças que identifica na atuação estatal. Com isso, sua investigação toca os agentes privados apenas na medida em que eles agem como emissários do Estado: restrições como as impostas por plataformas àquilo que seus usuários podem dizer são interpretadas, nesse contexto, como o resultado de meios de pressão formais e informais pelos quais Estados agem para impedir a disseminação de conteúdo indesejável, alinhando a atuação das plataformas àquilo que lhes interessa.

Para sustentar uma investigação em um sentido diferente – voltada às mudanças identificadas não na ação estatal, mas nas companhias que protagonizam essa nova infraestrutura da liberdade de expressão – percorro, a seguir, o caminho inverso. Primeiro (3.2.2.1 a 3.2.2.3), ofereço contraexemplos à tese de Balkin: três casos que expõem como a atuação de Estados – via, por exemplo, censura colateral ou exercício do *soft power* – explica apenas uma parte das restrições à liberdade de expressão que ocorrem em plataformas da internet. Por meio deles, fica claro que as mesmas transformações tecnológicas que criaram novos espaços de manifestação – as plataformas da internet – e, por isso, exigiram adaptações no modelo regulatório dos Estados, também desencadearam situações em que as companhias que administram esses ambientes se veem premidas a tomar decisões quanto àquilo que seus usuários publicam. Nesse sentido, e como os casos abaixo esclarecem, esses agentes privados, ao se verem rotineiramente chamados a atender a essa demanda, não só tomam decisões que são percebidas e controvertidas como julgamentos que afetam o direito à liberdade de expressão – tanto quanto quaisquer outras

restrições provenientes do Estado – como o fazem de acordo com políticas internas que, por vezes, indicam estarem eles desalinhados – e não alinhados – com os Estados. Assim, espero expor tanto a inaplicabilidade de ideias como as de censura colateral e *soft power* a, pelo menos, algumas situações de restrição à liberdade de expressão típicas da infraestrutura digital; quanto a existência de um âmbito independente de atuação decisória que, localizado junto a agentes privados como o Facebook e o Google, não é abrangido pelo enfoque conferido pelo Balkin de 2014 à questão.

A seguir (3.3), proponho uma investigação quanto a esse âmbito decisório, tanto o abordando sob o denominador comum da moderação de conteúdo das plataformas da internet, quanto o justificando como um tema de pesquisa. Argumento, nesse sentido, que para compreender toda a extensão dos impactos da era digital na regulação da liberdade de expressão, é preciso dar um passo além do que fez o Balkin de 2014 – na direção do Balkin de 2018 –, e, assim, entender como funcionam as estruturas decisórias privadas que, criadas a partir da infraestrutura digital, definem, com relativa independência, aquilo que as pessoas podem ou não podem dizer na internet.

### 3.2.2 Três Casos

#### 3.2.2.1 Caso n.º 1: O Horror da Guerra

O primeiro caso começa com uma postagem feita no Facebook pelo jornalista norueguês Tom Egeland, em setembro de 2016<sup>111</sup>. Egeland publicou na plataforma um artigo, em que discutia o impacto que a fotografia tem na história das guerras. Como era de se esperar, ele incluiu no trabalho a imagem “*O Horror da Guerra*”, uma fotografia em que crianças fogem de um ataque de *napalm* durante a Guerra do Vietnã. A imagem é especialmente impactante porque uma das crianças, Kim Phúc, aparece nua, sofrendo com as queimaduras decorrentes do ataque:

---

<sup>111</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 9 (Kindle, n. p.).

**Figura 2 – O Horror da Guerra**



Fonte: Time<sup>112</sup>

Logo após ser publicada, a imagem foi excluída pelo Facebook, porque violava os padrões de comunidade da plataforma – provavelmente os associados à proibição da exposição de violência e nudez infantil<sup>113</sup>. Egeland repostou a imagem duas vezes, acompanhada de críticas à postura da plataforma: segundo ele, “*a minha publicação não foi uma tentativa de desafiar o puritanismo americano do Facebook*” mas uma manifestação quanto à possibilidade de se “*refletir sobre o poder da fotografia*”<sup>114</sup>. Na primeira vez em que veiculada novamente, a imagem foi mais uma vez deletada, impondo-se ao jornalista uma suspensão que o proibia de se manifestar na plataforma por vinte e quatro horas; na segunda, a punição foi ampliada para três dias<sup>115</sup>.

---

<sup>112</sup> UT, N. The Terror of War. **Time**, The Most Influential Images of All Time. Disponível em <<http://100photos.time.com/photos/nick-ut-terror-war>>, acesso em 13.6.2019.

<sup>113</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 6 (Kindle, n. p.).

<sup>114</sup> EGELAND, T. (Tom Egeland). Sem título. Oslo, 26 de agosto de 2016. **Facebook**: tom.egeland. Disponível em <<https://www.facebook.com/tom.egeland/posts/10153691636277750>>, acesso em 13.6.2019.

<sup>115</sup> SCOTT, M., ISAAC, M. Facebook Restores Iconic Vietnam War Photo It Censored for Nudity. **The New York Times**, 9.9.2016, disponível em <<https://www.nytimes.com/2016/09/10/technology/facebook-vietnam-war-photo-nudity.html>>, acesso em 13.6.2019.



Como relata Gillespie<sup>116</sup> o caso começou a ganhar maiores proporções quando o jornal de maior circulação da Noruega, em que Egeland trabalhava, fez uma postagem criticando a remoção. Novamente compartilhada a imagem no Facebook – agora pelo jornal –, a plataforma reagiu de novo. Por meio de um e-mail, os editores foram informados que “Nós [a moderação de conteúdo do Facebook] *estabelecemos limitações na visualização de nudez para limitar a exposição das diferentes pessoas que utilizam a nossa plataforma a conteúdo sensível*”; por isso, “*quaisquer fotografias de pessoas expondo nudez total [...] serão removidas*”, com exceção de “*fotos de mulheres ativamente envolvidas na amamentação*” ou “*imagens produzidas digitalmente e postadas para fins educacionais ou satíricos, ou fotografias de produções artísticas reais*”, razões pelas quais se exigiu ao jornal “*que removesse ou colocasse pixels nesta foto*”<sup>117</sup>. Antes que a publicação pudesse responder, a imagem foi novamente deletada.

A exclusão da publicação de um jornal de grande circulação no país iniciou um movimento que mobilizou tanto a imprensa quanto as autoridades do governo norueguês contra o que decidiu o Facebook. A publicação que teve a manifestação deletada fez da situação a capa de sua edição diária, em que não só expôs o contexto, mas reproduziu a fotografia e publicou uma carta aberta a Mark Zuckerberg, criticando-o pela forma como o Facebook lida com a liberdade de expressão:

*“Preste atenção, Mark, isso é sério. Primeiro você cria regras que não distinguem pornografia infantil de fotografias de guerra conhecidas. Então você aplica essas regras sem dar espaço para um bom julgamento. Finalmente, você censura até mesmo as críticas*

---

<sup>116</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 9 e ss (Kindle, n. p.).

<sup>117</sup> “*We place limitations on the display of nudity to limit the exposure of the different people using our platform to sensitive content. Any photographs of people displaying fully nude genitalia or buttocks, or fully nude female breasts, will be removed. Photos of women actively engaged in breast feeding or exposing reconstructed nipples for awareness are allowed. We also make allowances for digitally produced content posted for educational, humorous, or satirical purposes, and for photographs of real world art. We understand that these limitations will sometimes affect content shared for legitimate reasons, including awareness campaigns or artistic projects, and we apologize for the inconvenience. Therefore I ask you to either remove or pixelize this picture*”. HANSEN, E. Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture. **Aftenposten**, 8.9.2016, disponível em <<https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture>>, acesso em 13.6.2019.

e as discussões quanto à decisão – e você pune a pessoa que ousa dar voz a essa crítica”<sup>118</sup>.

Figura 3 – Capa do jornal Aftenposten



Fonte: Aftenposten<sup>119</sup>

A crítica que segue é clara: o Facebook se tornou “a plataforma que é a líder-mundial para que informações sejam divulgadas, para debater e para o contato social entre as pessoas” e isso fez dele “o editor mais poderoso do mundo”. Banir, entretanto, uma imagem que “contribui para um debate mais aberto e mais crítico”,

---

<sup>118</sup> “Listen, Mark, this is serious. First you create rules that don’t distinguish between child pornography and famous war photographs. Then you practice these rules without allowing space for good judgement. Finally you even censor criticism against and a discussion about the decision – and you punish the person who dares to voice criticism”. HANSEN, E. Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture. **Aftenposten**, 8.9.2016, disponível em <<https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture>>, acesso em 13.6.2019. Sobre a relação entre a imagem e um suposto teor pornográfico, cf. IBRAHIM, Y. Facebook and the Napalm Girl: Reframing the Iconic as Pornographic. **Social Media + Society**, pp. 1/10, 2017, disponível em <<https://journals.sagepub.com/doi/pdf/10.1177/2056305117743140>>, acesso em 13.6.2019.

<sup>119</sup> HANSEN, E. Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture. **Aftenposten**, 8.9.2016, disponível em <<https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture>>, acesso em 13.6.2019.

significa que “o meio de comunicação mais importante do mundo está limitando a liberdade no lugar de tentar estendê-la, e isso ocasionalmente ocorre de um jeito autoritário”, já que “se você tomar a liberdade de desafiar as regras do Facebook, você vai encontrar – como nós temos visto – a censura”<sup>120</sup>. Outros veículos de imprensa também reagiram classificando a exclusão como uma forma de censura<sup>121</sup> e de restrição à liberdade de expressão<sup>122</sup>.

As autoridades da Noruega, por sua vez, endossaram as manifestações da imprensa, repostando, em contas do Facebook, a reportagem da publicação local, acompanhada de mensagens em que criticavam a postura da plataforma. O embate é bem exemplificado pelo que ocorreu com a Primeira Ministra do país, Erna Solberg. Após a imagem ser banida da página do jornal, ela a publicou com uma mensagem que ecoa as críticas feitas pela imprensa do país à plataforma:

---

<sup>120</sup> “*But, dear Mark, you are the world’s most powerful editor [...] They contributed to a more open, more critical debate [...] the world’s most important medium is limiting freedom instead of trying to extend it, and that this occasionally happens in an authoritarian way [...] If you take the liberty to challenge Facebook’s rules, you will be met – as we have seen – with censorship*”. HANSEN, E. Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture. **Aftenposten**, 8.9.2016, disponível em <<https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture>>, acesso em 13.6.2019.

<sup>121</sup> O termo censura foi utilizado à época para descrever a conduta do Facebook em relação à imagem; *cf.*, p. ex., KLEINMAN, Z. Fury over Facebook ‘Napalm girl’ censorship. **BBC News**, 9.9.2016, disponível em <<https://www.bbc.com/news/technology-37318031>>, acesso em 13.6.2019; OHLHEISER, A. Facebook backs down, will no longer censor the iconic ‘Napalm Girl’ war photo. **Washington Post**, 9.9.2016, disponível em <[https://www.washingtonpost.com/news/the-intersect/wp/2016/09/09/abusing-your-power-mark-zuckerberg-slammed-after-facebook-censors-vietnam-war-photo/?utm\\_term=.2176e89e5381](https://www.washingtonpost.com/news/the-intersect/wp/2016/09/09/abusing-your-power-mark-zuckerberg-slammed-after-facebook-censors-vietnam-war-photo/?utm_term=.2176e89e5381)>, acesso em 13.6.2019; e SIMS, A. Facebook to Reinstate Censored Image of ‘Napalm Girl’ after Mark Zuckerberg accused of ‘abusing power’. **The Independent**, 9.9.2016, disponível em <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-to-reinstate-censored-image-of-napalm-girl-after-mark-zuckerberg-accused-of-abusing-power-a7235021.html>>, acesso em 13.6.2019.

<sup>122</sup> *Cf.*, por exemplo, ecoando a crítica do Aftenposten, LUI, K. Mark Zuckerberg Has Been Blasted Over Facebook’s Censorship of the ‘Napalm Girl’ Photo. **Time**, 9.9.2016, disponível em <<https://time.com/4484905/facebook-censorship-napalm-girl-newspaper-editor-norway/>>, acesso em 13.6.2019.

“Eu aprecio o trabalho do Facebook e outras mídias para barrar imagens e conteúdo que exponham abuso e violência. Mas o Facebook entendeu errado ao censurar imagens como essas. Ele contribui para restringir a liberdade de expressão”.<sup>123</sup>

#### Figura 4 – Publicações da Primeira-Ministra norueguesa no Facebook, antes e depois da ação da moderação de conteúdo da plataforma



Fonte: Aftenposten<sup>124</sup>

Tanto as manifestações da imprensa, quanto as dos políticos – incluindo a mensagem da Primeira Ministra, que, em seguida, republicou a foto com um quadrado preto encobrendo a criança, como reproduzido acima – foram sistematicamente deletadas pelo Facebook. Após ser questionado por alguns veículos de mídia e ser acusada de ter

<sup>123</sup> “I appreciate the work of Facebook and other media to stop pictures and content showing abuse and violence,” the PM wrote in an accompanying message. “But Facebook gets it wrong when it censors pictures like these. It contributes to restricting the freedom of speech.”. PRICE, R. Facebook just censored the prime minister of Norway for posting one of the most famous photos ever. **Business Insider**, 9.9.2016, disponível em <<https://www.businessinsider.com/facebook-censors-norway-prime-minister-erna-solberg-napalm-girl-vietnam-photo-aftenposten-2016-9>>, acesso em 19.6.2019.

<sup>124</sup> A primeira-ministra publicou a imagem autocensurada, à esquerda, após ter a original, à direita, excluída pelo Facebook. FAERAS, A., NORDBY, K. OLSEN, T. GJERDE, R. Erna Solberg svarer sensur av historisk Vietnam-fotografi med å legge ut svartsladdede versjoner av flere kjente og historiske bilder. **Aftenposten**, disponível em <<https://www.aftenposten.no/norge/i/qEjdL/Slik-svarer-Erna-Solbergs-pa-Facebook-sensur>>, acesso em 13.6.2019.

abusado de seu poder<sup>125</sup>, a plataforma manteve sua posição, justificando que “[a moderação de conteúdo] *reconhece que a foto é icônica, mas é difícil criar uma distinção entre permitir uma fotografia de uma criança nua em uma situação e não em outras*”, ressaltando que “*nós tentamos encontrar um balanço correto entre permitir às pessoas se expressarem enquanto mantemos um ambiente seguro e respeitoso para a nossa comunidade global*”<sup>126</sup>.

A situação veio a termo apenas algumas semanas depois, quando pressões direcionadas à plataforma fizeram o Vice-Presidente do Facebook, Justin Osofsky – viesse a público esclarecer que o “*o Facebook tem padrões de comunidade para manter as pessoas seguras e criar uma experiência positiva em nossa plataforma*”<sup>127</sup>. Assim, “*na vasta maioria dos casos, essas políticas funcionam bem*”; “*Às vezes, entretanto, a importância global e histórica de uma foto como ‘O Horror da Guerra’ superam a importância de se manter a nudez fora do Facebook*”. Comprometendo-se a “*melhorar nossas políticas para garantir que elas promovam a livre expressão e mantenham a comunidade segura*”, a plataforma decidiu abrir uma exceção para a fotografia pelo seu caráter “*icônico*”<sup>128</sup> e permitir publicações com a imagem – reestabelecendo as postagens

---

<sup>125</sup> Cf., por exemplo, BLAIR, O. Mark Zuckerberg accused of ‘abusing power’ after Facebook deletes iconic image of ‘napalm girl’. **The Independent**, 9.6.2016, disponível em <<https://www.independent.co.uk/news/people/mark-zuckerberg-facebook-accused-abusing-power-napalm-girl-vietnam-war-image-a7233431.html>>, acesso em 19.6.2019.

<sup>126</sup> “*While we recognize that this photo is iconic, it's difficult to create a distinction between allowing a photograph of a nude child in one instance and not others. We try to find the right balance between enabling people to express themselves while maintaining a safe and respectful experience for our global community. Our solutions won't always be perfect, but we will continue to try to improve our policies and the ways in which we apply them*”. LEVIN, S., WONG, J., HARDING, L. Facebook backs down from 'napalm girl' censorship and reinstates photo. **The Guardian**, 9.9.2016, disponível em <<https://www.theguardian.com/technology/2016/sep/09/facebook-reinstates-napalm-girl-photo>>, acesso em 13.6.2019.

<sup>127</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 11 (Kindle, n. p.).

<sup>128</sup> “*Facebook has community standards to keep people safe and create a positive experience on our platform [...] In the vast majority of cases, these policies work well [...] Sometimes, however, the global and historical importance of a photo like ‘Terror of War’ outweighs the importance of keeping nudity off Facebook*”. OSOFSKY, J. (Justin Osofsky). I want to share some thoughts about expression and community standards on Facebook. Sem localização, 9 de setembro de 2016. **Facebook**: josofsky. Disponível em <<https://www.facebook.com/josofsky/posts/10157347245570231>>, acesso em 13.6.2019.

que haviam sido deletadas tanto da página do jornalista, quanto do periódico e dos demais membros do governo norueguês.

### 3.2.2.2 Caso n.º 2: Alex Jones e o Infowars

O cerne do segundo caso são as manifestações feitas por um usuário específico, Alex Jones, em contas pessoais e institucionais de plataformas como o Facebook, o Twitter e o YouTube. Jones é o fundador do site Infowars, conhecido pela divulgação de informações falsas e teorias conspiratórias enfocadas majoritariamente no ataque a minorias. Ele teve exposição ao publicar, depois de um ataque que matou 26 crianças na escola Sandy Hook em 2012, uma série de vídeos no YouTube sustentando que o evento tinha sido armado por famílias e pelo governo para retirar as armas de cidadãos americanos<sup>129</sup> – um argumento semelhante ao que ele usaria seis anos depois, quanto a um tiroteio em uma escola em Parkland, Florida<sup>130</sup>. Dentre outras controvérsias, Jones fez publicações sustentando que vacinas causam autismo<sup>131</sup>, que um genocídio branco existiu<sup>132</sup> e admitiu que ajudou a espalhar, nas eleições de 2016 nos Estados Unidos, o boato de que Hillary Clinton e outros membros da cúpula do Partido Democrata administravam um esquema de pornografia infantil e tráfico de pessoas disfarçado de pizzaria, em Washington, DC – na época, a polícia prendeu uma pessoa que foi até o

---

<sup>129</sup> WILLIAMSON, E. How Alex Jones and Infowars Helped a Florida Man Torment Sandy Hook Families. **The New York Times**, 29.3.2019, disponível em <<https://www.nytimes.com/2019/03/29/us/politics/alex-jones-infowars-sandy-hook.html>>, acesso em 19.6.2019.

<sup>130</sup> MORRIS, D. YouTube Pulls Alex Jones Video Saying Student Anti-Gun Activists Were Actors. **Fortune**, 24.2.2018, disponível em <<http://fortune.com/2018/02/24/youtube-pulls-alex-jones-infowars-video/>>, acesso em 19.6.2019.

<sup>131</sup> ZADROZNY, B. Vaccine misinformation and Infowars: Researchers wary of Facebook's embrace of 'Groups'. **NBC News**, 16.2.2019, disponível em <<https://www.nbcnews.com/tech/tech-news/vaccine-misinformation-infowars-researchers-wary-facebook-s-embrace-groups-n972286>>, acesso em 19.6.2019.

<sup>132</sup> DEARDEN, L. New Zealand attack: How nonsensical white genocide conspiracy theory cited by alleged gunman is spreading poison around the world. **The Independent**, 16.3.2019, disponível em <<https://www.independent.co.uk/news/world/australasia/new-zealand-christchurch-mosque-attack-white-genocide-conspiracy-theory-a8824671.html>>, acesso em 19.6.2019.

estabelecimento e disparou um rifle semiautomático enquanto investigava as “denúncias”<sup>133</sup>.

As postagens de Jones e do Infowars foram sujeitas a uma constante de atenção das companhias por trás de plataformas como o Facebook, Twitter e YouTube. Frequentemente, ele era objeto de sanções relacionadas à possibilidade de utilizar a plataforma para propagar aquilo que veiculava em seu site, em específico quanto a notícias que sabia que eram falsas. Em julho de 2018, por exemplo, o Facebook determinou uma suspensão da página pessoal de Jones, proibindo-o de publicar na plataforma por 30 dias, após receber “denúncias de quatro vídeos diferentes” e avaliar que “o conteúdo é contrário aos nossos Padrões de Comunidade”<sup>134</sup>.

Na mesma semana, o YouTube retirou os vídeos do canal de Jones, vedando transmissões ao vivo por 90 dias, por violações às “políticas quanto ao discurso de ódio e a colocar crianças em perigo” – algo parecido ao que tinha feito em fevereiro do mesmo ano, quando excluiu uma gravação em que Jones acusava uma vítima de um tiroteio ocorrido em uma escola primária de ser um ator<sup>135</sup>. A plataforma informou, em um comunicado, que “nós aplicamos nossas políticas consistentemente, de acordo com o conteúdo dos vídeos, independentemente de quem os veicula ou do canal”<sup>136</sup>. Na mesma linha, o Twitter impediu que a conta de Jones fizesse novas manifestações na por sete dias, sustentando, de início, que a existência da página não violava os seus termos e condições<sup>137</sup>.

---

<sup>133</sup> ROSENBERG, E. Alex Jones Apologizes for Promoting ‘Pizzagate’ Hoax. **The New York Times**, 25.3.2017, disponível em <<https://www.nytimes.com/2017/03/25/business/alex-jones-pizzagate-apology-comet-ping-pong.html>>, acesso em 19.6.2019.

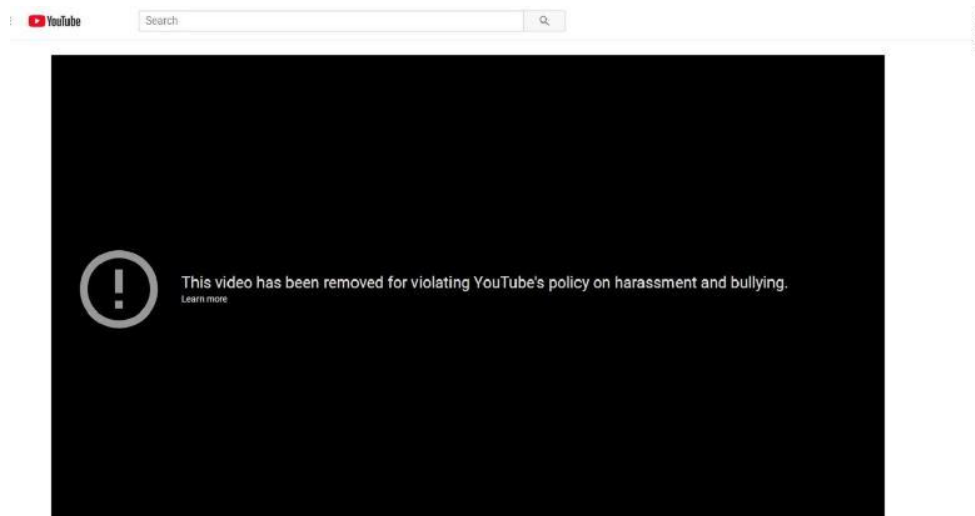
<sup>134</sup> VINCENT, J. Alex Jones hit with 30-day Facebook suspension for bullying and hate speech. **The Verge**, 27.7.2018, disponível em <<https://www.theverge.com/2018/7/27/17621334/alex-jones-facebook-suspension-30-day>>, acesso em 20.6.2019.

<sup>135</sup> MURPHY, P. InfoWars' main YouTube channel is two strikes away from being banned. **CNN**, 24.2.2018, disponível em <<https://edition.cnn.com/2018/02/23/us/infowars-youtube-videos-trnd/index.html>>, acesso em 19.6.2019.

<sup>136</sup> ROOSE, K. Facebook and YouTube give Alex Jones a Wrist Slap. **The New York Times**, 27.7.2018, disponível em <<https://www.nytimes.com/2018/07/27/technology/alex-jones-facebook-youtube.html>>, acesso em 19.6.2019.

<sup>137</sup> TAN, J. Twitter cracks down on right-wing media personality Alex Jones. **CNBC**, 15.8.2018, disponível em <<https://www.cnn.com/2018/08/15/twitter-suspends-alex-jones.html>>, acesso em 19.6.2019.

**Figura 5 – Mensagem no YouTube informando a remoção de um vídeo do Infowars**



Fonte: YouTube / CNN<sup>138</sup>

Após um primeiro semestre de 2018 particularmente ativo de Jones, entretanto, as companhias por trás dessas plataformas começaram a tomar medidas mais enérgicas contra as publicações do Infowars, até que tanto a página do site, quanto a pessoal do ativista, foram sendo, uma a uma, permanentemente banidas das plataformas mais importantes da internet.

O YouTube desativou as páginas associadas a Alex Jones em agosto de 2018 – usuários que tentavam acessá-las eram redirecionados a uma mensagem padrão, com a informação de que “*essa conta foi encerrada por violar os Padrões de Comunidade do YouTube*”. Em resposta a um questionamento da imprensa a respeito, o Google – proprietário do YouTube – anunciou que “*todos os usuários concordam em obedecer os nossos Termos de Serviço e Padrões de Comunidade quando se inscrevem para usar o YouTube*”; assim, “*Quando usuários violam repetidamente essas políticas, como as nossas políticas contra o discurso de ódio e assédio, ou os nossos termos que proíbem atos que contornem os nossos meios de executá-las [enforcement], nós encerramos as suas contas*”<sup>139</sup>. O Twitter fez o mesmo em setembro, baseado em “*novas denúncias de Tweets*

---

<sup>138</sup> MURPHY, P. InfoWars' main YouTube channel is two strikes away from being banned. **CNN**, 24.2.2018, disponível em <<https://edition.cnn.com/2018/02/23/us/infowars-youtube-videos-trnd/index.html>>, acesso em 19.6.2019.

<sup>139</sup> SALINAS, S. YouTube removes Alex Jones' page, following bans from Apple and Facebook. **CNBC**, 6.8.2018, disponível em <<https://www.cnbc.com/2018/08/06/youtube-removes-alex-jones-account-following-earlier-bans.html>>, acesso em 19.6.2019.



e vídeos postados ontem [5.9.2018] que violavam nossa política sobre comportamento abusivo, adicionalmente às violações anteriores da conta”<sup>140</sup>.

**Figura 6 – Mensagens de bloqueio no Twitter das contas de Alex Jones e do Infowars**



Fonte: Twitter<sup>141</sup>

A exclusão de Alex Jones e do Infowars das plataformas da internet gerou uma escalada de controvérsias a esse respeito, em específico por sua associação à extrema-direita dos Estados Unidos – o CEO da Apple, Tim Cook, por exemplo, veio a público na

---

<sup>140</sup> O PayPal, utilizado pelo Infowars para receber doações, encerrou o contrato com o ativista no mesmo mês. Não há notícia, entretanto, tal como no que ocorreu no *Wikileaks* (cf. *supra*, nota n.º 84), que isso tenha sido consequência de uma pressão governamental: tudo indica que foi um movimento da própria plataforma. Cf. LIAO, S. Twitter permanently suspends Infowars and Alex Jones. **The Verge**, 6.9.2018, disponível em <<https://www.theverge.com/2018/9/6/17829188/twitter-permanently-suspension-infowars-alex-jones>>, acesso em 19.6.2019; FUNG, B. PayPal bans Alex Jones, saying Infowars 'promoted hate or discriminatory intolerance'. **The Washington Post**, 29.9.2018, disponível em <[https://www.washingtonpost.com/technology/2018/09/21/paypal-bans-alex-jones-saying-infowars-promoted-hate-or-discriminatory-intolerance/?utm\\_term=.6f6931d5660d](https://www.washingtonpost.com/technology/2018/09/21/paypal-bans-alex-jones-saying-infowars-promoted-hate-or-discriminatory-intolerance/?utm_term=.6f6931d5660d)>, acesso em 19.6.2019.

<sup>141</sup> INFOWARS. (Infowars). Página no Twitter. Sem localização, sem data. **Twitter**: infowars. Disponível em <<https://twitter.com/account/infowars>>, acesso em 19.6.2019; JONES, A. (Alex Jones). Página no Twitter. Sem localização, sem data. **Twitter**: realalexjones. Disponível em <<https://twitter.com/account/realalexjones>>, acesso em 19.6.2019.

mesma época, afirmar que a retirada dos podcasts operados por Jones e pela Infowars não era o resultado de um ataque coordenado com as outras empresas de tecnologia<sup>142</sup>.

O último movimento de exclusão – a partir do qual a questão ganhou a maior proporção – foi a decisão, tomada pelo Facebook em 2019, de banir permanentemente da rede social não apenas as páginas do ativista, mas também qualquer publicação que elogiasse ou desse suporte a ideias semelhantes às que ele publicava. O impulso para isso começou em março, quando o Facebook decidiu reverter uma política estabelecida desde o início da plataforma e banir conteúdos que *“elogiem, apoiem ou representem nacionalismo ou separatismo branco”*, porque teria ficado claro que *“esses conceitos estão profundamente vinculados com grupos de ódio organizado e não têm lugar nos nossos serviços”*. A nota oficial, com o título *“Opondo-se contra o ódio”*, fornece um raro retrato do raciocínio por trás de uma mudança de política de conteúdo: segundo o Facebook, já existia uma vedação a publicações que implicassem em tratamentos de ódio a *“pessoas baseados em características como a raça, a etnicidade ou a religião”*. Essa proibição se aplicava ao supremacismo branco, mas não ao separatismo branco, porque o último era vinculado pelas políticas de conteúdo da plataforma ao separatismo em geral – *“situações tais como o orgulho americano ou o separatismo basco, que são uma parte importante da identidade das pessoas”* – o qual, via de regra, é permitido. Após dialogar com a comunidade e constatar que o separatismo branco estava indissociavelmente ligado ao supremacismo branco, o Facebook reviu a orientação conferida aos moderadores de conteúdo e banuiu ambos da plataforma: *“as pessoas ainda poderão demonstrar orgulho sobre a sua origem étnica, mas não vamos tolerar exaltações ou apoio ao nacionalismo e separatismo brancos”*<sup>143</sup>.

---

<sup>142</sup> STATT, N. Tim Cook says Apple didn't coordinate Alex Jones ban with other tech companies. **The Verge**, 2.10.2018, disponível em <<https://www.theverge.com/2018/10/2/17930534/apple-ceo-tim-cook-alex-jones-infowars-ban-independent-decision>>, acesso em 19.6.2019.

<sup>143</sup> *“Today we're announcing a ban on praise, support and representation of white nationalism and white separatism on Facebook and Instagram, which we'll start enforcing next week. It's clear that these concepts are deeply linked to organized hate groups and have no place on our services”; “We didn't originally apply the same rationale to expressions of white nationalism and white separatism because we were thinking about broader concepts of nationalism and separatism — things like American pride and Basque separatism, which are an important part of people's identity”*. FACEBOOK. Standing Against Hate. **Facebook Newsroom**, 27.3.2019, disponível em <<https://newsroom.fb.com/news/2019/03/standing-against-hate/>>, acesso em 19.6.2019.

Pouco mais de um mês após o anúncio, o Facebook decidiu excluir permanentemente as páginas do Infowars e de Alex Jones – em meio a um anúncio mais amplo, que também baniu outras figuras da extrema direita, como Paul Joseph Watson, Paul Nehlen Milo Yiannopoulos, Laura Loomer e Louis Farrakhan. Segundo um porta-voz da plataforma, “*nós sempre banimos indivíduos e organizações que promovem ou se engajam em violência e ódio, independentemente da ideologia*”, justificando que “*o processo para avaliar violadores em potencial é extenso e foi ele que nos levou à decisão de remover essas contas hoje*”<sup>144</sup>.

Uma possível explicação para o maior impacto conferido ao evento foi o fato de que o Facebook não apenas baniu as páginas de Jones, mas proibiu qualquer usuário ou grupo que use a plataforma de disseminar conteúdo proveniente do Infowars. Logo após a notícia vir a público, formou-se uma controvérsia quanto ao poder que o Facebook exerce ao fazer esse tipo de exclusão e em que medida ela se traduz em uma restrição à liberdade de expressão: de um lado, houve quem considerasse que “*faz tempo que é óbvio que as redes sociais precisam fazer mais para controlar as páginas de pessoas que espalham o ódio em suas plataformas*”, associando-se o uso, por exemplo, do Facebook a ataques a refugiados na Alemanha, genocídios no Sri Lanka e a um massacre em mesquitas na Nova Zelândia<sup>145</sup>. De outro, e traçando um paralelo com as ações que Larry Flint enfrentou pelas publicações pornográficas que patrocinava, sobrevieram posições no sentido de que “*restringir linguagem ofensiva ou danosa para um bem maior é algo que vai muito bem até que você se torna dependente de uma definição de ‘bem maior’ com a qual você não concorda*”<sup>146</sup>.

Assim como no caso anterior, a discussão foi da mídia às autoridades governamentais. Após o Facebook banir Jones, o Presidente dos Estados Unidos se manifestou a respeito: primeiro, ele compartilhou um vídeo do Infowars, que lhe era

---

<sup>144</sup> ROETTIGERS, J. Facebook Bans Alex Jones, Milo Yiannopoulos, Other Far-Right Figures. **Variety**, 2.10.2019, disponível em <<https://variety.com/2019/digital/news/facebook-alex-jones-ban-1203203444/>>, acesso em 19.6.2019.

<sup>145</sup> ALAMO, K. Facebook handled Alex Jones just right. **CNN**, 5.5.2019, disponível em <<https://edition.cnn.com/2019/05/05/opinions/facebook-banish-alex-jones-alaimo/index.html>>, acesso em 19.6.2019; e GRAHAM, J. Did social media do the right thing by banning Alex Jones? **USA Today**, 6.8.2018, disponível em <<https://www.usatoday.com/story/tech/talkingtech/2018/08/06/did-facebook-youtube-and-apple-do-right-thing-banning-alex-jones/916763002/>>, acesso em 19.6.2019.

<sup>146</sup> LOUIS, E. We need to talk about Alex Jones. **CNN**, 2.5.2019, disponível em <<https://edition.cnn.com/2018/08/06/opinions/infowars-opinion-roundup/index.html>>, acesso em 19.6.2019.

favorável<sup>147</sup>. Depois, por meio de uma série de publicações no Twitter, Trump afirmou que continuava a monitorar a censura que cidadãos americanos sofriam em redes sociais pelas plataformas, indicando que as exclusões se opunham ao direito à liberdade de expressão como estabelecido nos Estados Unidos<sup>148</sup>:

*“Eu continuo a monitorar a censura de CIDADÃOS AMERICANOS em plataformas de mídias sociais. Estes são os Estados Unidos da América – e nós temos o que é chamado de LIBERDADE DE EXPRESSÃO! Nós estamos monitorando e vigiando, de perto”*<sup>149</sup>.

*“As maravilhosas Diamond e Silk têm sido tratadas de uma maneira tão horrível pelo Facebook. Elas trabalharam duro e que tem sido feito com elas é muito triste – e nós estamos investigando a respeito. Está ficando cada vez pior para os Conservadores nas mídias sociais”*<sup>150</sup>.

*“Como é possível que James Woods (e muitos outros), uma forte mas responsável Voz Conservadora, ter sido banida do Twitter? Mídias Sociais & Mídias de Fake News, juntas com o seu parceiro, o Partido Democrata, não têm ideia dos problemas que eles estão causando para eles mesmos. MUITO INJUSTO”*<sup>151</sup>.

---

<sup>147</sup> TRUMP, D. (Donald Trump). Sem título. Sem localização, 4 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124645001533771776>>, acesso em 19.6.2019.

<sup>148</sup> HAYES, C. Trump attacks social media sites after James Woods, Alex Jones other extremist conservatives banned. **USA Today**, 4.5.2019, disponível em <<https://www.usatoday.com/story/news/politics/2019/05/04/donald-trump-attacks-twitter-facebook-james-woods-banned/1102418001/>>, acesso em 19.6.2019.

<sup>149</sup> *“I am continuing to monitor the censorship of AMERICAN CITIZENS on social media platforms. This is the United States of America — and we have what’s known as FREEDOM OF SPEECH! We are monitoring and watching, closely!!”*. TRUMP, D. (Donald Trump). Sem título. Sem localização, 3 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124447302544965634>>, acesso em 19.6.2019.

<sup>150</sup> *“The wonderful Diamond and Silk have been treated so horribly by Facebook. They work so hard and what has been done to them is very sad - and we’re looking into. It’s getting worse and worse for Conservatives on social media”*. Diamond e Silk são referências a Lynnette Hardaway e Rochelle Richardson, ativistas que veiculam em mídias sociais mensagens de apoio a Trump. TRUMP, D. (Donald Trump). Sem título. Sem localização, 3 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124454393749934080>>, acesso em 19.6.2019.

<sup>151</sup> *“How can it be possible that James Woods (and many others), a strong but responsible Conservative Voice, is banned from Twitter? Social Media & Fake News Media, together with their partner, the Democrat Party, have no idea the problems they are causing for themselves. VERY UNFAIR”*. TRUMP, D. (Donald

Como parte de uma política contra o que Trump denominou a censura praticada por plataformas na internet, a Casa Branca publicou uma mensagem no Twitter, em que diz que “*a Administração Trump está lutando pela liberdade de expressão online*” e, por isso, criou um formulário na internet, em que as pessoas podem relatar casos de enviesamento na conduta adotada por plataformas sociais<sup>152</sup>. A página indicada na mensagem permite que o usuário preencha um campo em branco, que é acompanhado da seguinte orientação:

*“PLATAFORMAS DE MÍDIAS SOCIAIS devem fazer evoluir a LIBERDADE DE EXPRESSÃO. Mesmo assim, muitos Americanos têm visto suas contas serem suspensas, banidas ou denunciadas de forma fraudulenta por ‘violações’ pouco claras de políticas de usuários.*

*Não importa quais são os seus posicionamentos, se você suspeitar que o enviesamento político gerou alguma ação do tipo contra você, compartilhe sua história com o Presidente Trump”<sup>153</sup>.*

Mesmo com esse conjunto de ações, as contas de Alex Jones – e dos outros ativistas da extrema direita nos Estados Unidos – permanecem bloqueadas no Facebook, Twitter e YouTube.

### 3.2.2.3 Caso n.º 3: Holocausto negado

---

Trump). Sem título. Sem localização, 4 de maio de 2019. Twitter: realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124743267873116160>>, acesso em 19.6.2019.

<sup>152</sup> WHITE HOUSE. Sem título. Sem localização, 15 de maio de 2019. **Twitter:** WhiteHouse. Disponível em <<https://twitter.com/WhiteHouse/status/1128765001223663617>>, acesso em 19.6.2019.

<sup>153</sup> “*SOCIAL MEDIA PLATFORMS should advance FREEDOM OF SPEECH. Yet too many Americans have seen their accounts suspended, banned, or fraudulently reported for unclear “violations” of user policies. No matter your views, if you suspect political bias caused such an action to be taken against you, share your story with President Trump*”. O formulário estava disponível, em acesso inicial datado de 15.5.2019, na página <<https://whitehouse.typeform.com/to/Jti9QH>>; novo acesso realizado em 19.6.2019 resultou em mensagem segundo a qual “*o formulário está fechado*”. Cf. STACK, L. Trump Wants Your Tales of Social Media Censorship. And Your Contact Info. **The New York Times**, 15.5.2019, disponível em <<https://www.nytimes.com/2019/05/15/us/donald-trump-twitter-facebook-youtube.html>>, acesso em 19.6.2019.

O terceiro caso diz respeito ao posicionamento do Facebook quanto a um tema bastante comum às discussões sobre o direito à liberdade de expressão: a negação do Holocausto<sup>154</sup>. A questão veio à tona em uma entrevista em que Mark Zuckerberg, CEO do Facebook, foi questionado por uma repórter quanto às políticas da plataforma para exclusão de páginas criadas por seus usuários. Em um determinado momento da conversa, a jornalista o perguntou quanto à existência de páginas no Facebook que negavam a existência do Holocausto:

*“Kare Swisher: Ok. ‘Sandy Hook não aconteceu’ não é um debate. É falso. Você não pode simplesmente tirar isso do ar?”*

*Mark Zuckerberg: Eu concordo que é falso.*

[...]

*Eu sou judeu, e há um conjunto de pessoas que negam que houve um Holocausto. Eu acho isso profundamente ofensivo. Mas no final do dia, eu não acredito que a nossa plataforma deve tirar isso do ar porque eu acho que há coisas que diferentes pessoas compreendem errado. Eu não acho que eles intencionalmente entendem isso errado, mas eu penso-.*

*Kare Swisher: No caso de que nega o Holocausto, eles podem estar, mas siga em frente.*

*Mark Zuckerberg: É difícil impugnar e entender a intenção [de alguém]. Eu penso apenas, que por mais horríveis que sejam esses exemplos, eu acho que a realidade é que eu também entendo coisas errado quando eu falo publicamente. E eu tenho certeza que você também. E eu tenho certeza que vários líderes e figuras públicas que nós respeitamos também, e eu não penso que é a coisa certa dizer ‘Nós vamos tirar alguém da plataforma se entender as coisas errado, mesmo múltiplas vezes’. O que nós vamos fazer é que vamos dizer: Ok, você tem a sua página, e se você não está tentando organizar [atos para] prejudicar alguém, ou atacando alguém, então você pode por esse conteúdo na sua página, mesmo que as pessoas discordem disso ou pensem que isso é ofensivo’. Mas isso não significa que nós temos que fazer isso ser distribuído amplamente no Feed de Notícias. Eu acho que nós, na realidade, pelo contrário-”<sup>155</sup>.*

---

<sup>154</sup> Para um resumo de diferentes posições a respeito, cf. LASSON, K. Holocaust Denial and the First Amendment: The Quest for Truth in a Free Society. **George Mason Law Review**, v. 6, n. 35, pp. 35/86, 1997; via de regra, o problema da negação do Holocausto no âmbito da liberdade de expressão é tratado sob a perspectiva de inseri-lo ou não como uma forma de discurso de ódio; cf., por exemplo, KAHN, R. Holocaust Denial and Hate Speech in HENNEBEL, L., HOCHMANN, T. **Genocides Denials and the Law**. New York: Oxford University Press, 2011, pp. 77/108.

<sup>155</sup> *“Kare Swisher: Okay. “Sandy Hook didn’t happen” is not a debate. It is false. You can’t just take that down? Mark Zuckerberg: I agree that it is false. [...] I’m Jewish, and there’s a set of people who deny that the Holocaust happened. I find that deeply offensive. But at the end of the day, I don’t believe that our platform should take that down because I think there are things that different people get wrong. I don’t think*

A entrevista teve repercussão<sup>156</sup>, porque acabou trazendo acidentalmente novamente à tona a decisão dos moderadores de conteúdo do Facebook de permitir, na plataforma, manifestações de que o Holocausto não existiu. Após publicação da entrevista e em razão das críticas que sofreu, Zuckerberg encaminhou uma mensagem à repórter, mantendo a posição da plataforma de não excluir esse tipo de postagem, mas esclarecendo que “*eu pessoalmente acho a negativa do Holocausto profundamente ofensiva, e eu absolutamente não pretendi defender a intenção de pessoas que negam isso*”, mas “*Nosso objetivo com as fake news não é impedir as pessoas de dizerem algo falso – mas parar com que as fake news e as desinformações se espalhem por nosso serviço*”<sup>157</sup>.

Naturalmente, a situação levantou novamente o debate quanto a se manifestações que negam a existência do Holocausto devem ser permitidas ou não. A controvérsia – que, na origem, remonta às discussões que levaram tanto à aprovação

---

*that they're intentionally getting it wrong, but I think- Kare Swisher: In the case of the Holocaust deniers, they might be, but go ahead. Mark Zuckerberg: It's hard to impugn intent and to understand the intent. I just think, as abhorrent as some of those examples are, I think the reality is also that I get things wrong when I speak publicly. I'm sure you do. I'm sure a lot of leaders and public figures we respect do too, and I just don't think that it is the right thing to say, "We're going to take someone off the platform if they get things wrong, even multiple times." What we will do is we'll say, "Okay, you have your page, and if you're not trying to organize harm against someone, or attacking someone, then you can put up that content on your page, even if people might disagree with it or find it offensive." But that doesn't mean that we have a responsibility to make it widely distributed in News Feed. I think we, actually, to the contrary-".* SWISHER, K. Zuckerberg: The Recode interview. **Vox**, 8.10.2018, disponível em <<https://www.vox.com/2018/7/18/17575156/mark-zuckerberg-interview-facebook-recode-kara-swisher>>, acesso em 19.6.2019.

<sup>156</sup> Cf., por exemplo, HERN, A. Mark Zuckerberg's remarks on Holocaust denial 'irresponsible'. **The Guardian**, 19.7.2018, disponível em <<https://www.theguardian.com/technology/2018/jul/19/mark-zuckerberg-holocaust-denial-facebook-remarks-offensive>>, acesso em 20.6.2019.

<sup>157</sup> “*I personally find Holocaust denial deeply offensive, and I absolutely didn't intend to defend the intent of people who deny that*”, mas que “*Our goal with fake news is not to prevent anyone from saying something untrue — but to stop fake news and misinformation spreading across our services*”. SWISHER, K. Mark Zuckerberg clarifies: ‘*I personally find Holocaust denial deeply offensive, and I absolutely didn't intend to defend the intent of people who deny that*.’. **Vox**, 18.7.2018, disponível em <<https://www.vox.com/2018/7/18/17588116/mark-zuckerberg-clarifies-holocaust-denial-offensive>>, acesso em 19.6.2019.

quanto à rejeição de legislações restritivas nesse sentido<sup>158</sup> – tem lugar nas plataformas da internet desde o momento em que popularizadas, no início de 2010. No Facebook, por exemplo, a polêmica remonta a 2011, quando 21 sobreviventes do Holocausto enviaram uma carta “*exigindo que a negativa do Holocausto não seja mais permitida como parte da política de ‘liberdade de expressão’ do Facebook*”<sup>159</sup> – um pedido que, como o renascimento acima da polêmica indicou, foi negado. O Twitter, mesmo que de maneira implícita, também sempre permitiu esse tipo de manifestação na plataforma, em que surgiu, inclusive, um *hashtag* capaz de vinculá-las – “*holohoax*” [“*o boato do holocausto*”], política semelhante à do YouTube<sup>160</sup>.

---

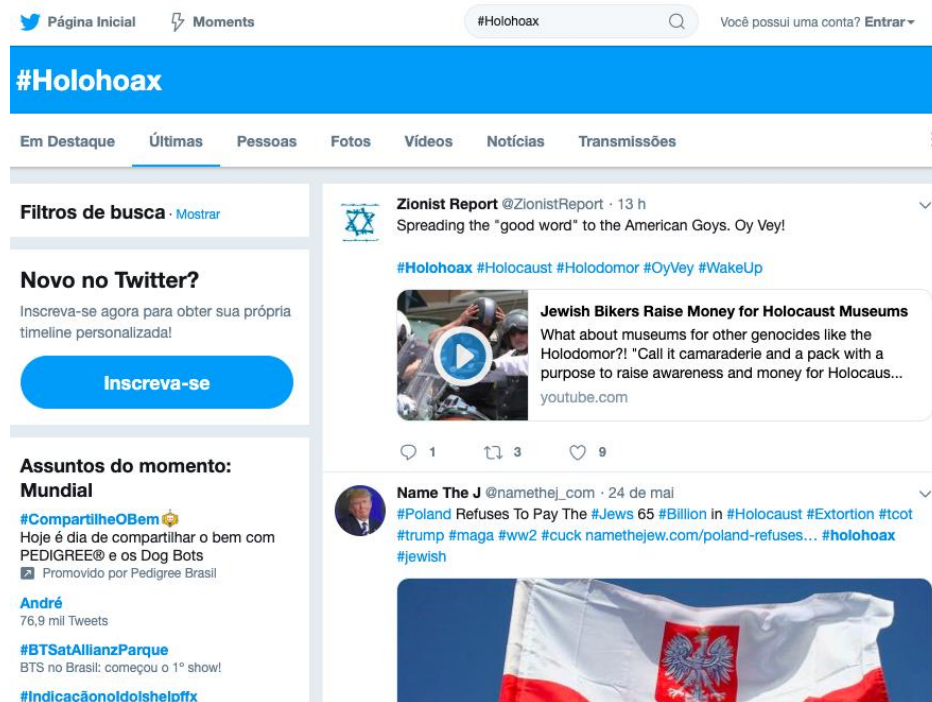
<sup>158</sup> Para uma análise das legislações quanto à negação do Holocausto, cf. HENNEBEL, L., HOCHMANN, T. **Genocides Denials and the Law**. New York: Oxford University Press, 2011, pp. 255/267.

<sup>159</sup> “[...] *asking that Holocaust denial no longer be allowed as part of facebook's "free speech" policy* [...]”. LIPMAN, J. Facebook: Holocaust denial pages can stay. **The Jewish Chronicle**, 28.7.2011, disponível em <<https://www.thejc.com/news/world/facebook-holocaust-denial-pages-can-stay-on-site-1.25833>>, acesso em 19.6.2019.

<sup>160</sup> FINKELSTEIN, D. Don't change laws to mop up after YouTube. **The Times**, 29.5.2018, disponível em <<https://www.thetimes.co.uk/article/don-t-change-laws-to-mop-up-after-youtube-c9lhrkwrq>>, acesso em 19.6.2019. Enquanto este trabalho era escrito, o YouTube decidiu implantar uma alteração na sua política de conteúdo e passou a proibir, na plataforma, conteúdos que neguem o Holocausto. Cf., por exemplo, STEWART, E. YouTube finally banned content from neo-Nazis, Holocaust deniers, and Sandy Hook skeptics. **Vox**, 5.6.2019, disponível em <<https://www.vox.com/recode/2019/6/5/18653666/youtube-bans-white-supremacists-holocaust-carlos-maza>>, acesso em 19.6.2019.



Figura 7 – Exemplos de resultados do hashtag “holohoax”



Fonte: Twitter<sup>161</sup>

Assim, e à semelhança do que ocorreu quanto às demais plataformas e ao próprio Facebook em momentos anteriores, a entrevista gerou a reação de diversos veículos de mídia associados ao sionismo, que questionaram a decisão da plataforma<sup>162</sup>. De um lado, o Facebook declarou, por meio de seu Vice-Presidente de Políticas, que “*não vai remover mentiras ou conteúdos que são imprecisos – não importa se se trata de negar o Holocausto, o massacre armênio ou o fato de que o governo sírio matou centenas de milhares do seu próprio povo*”, justificando que, a despeito da política da plataforma de vedar o discurso de ódio contra o antissemitismo, isso não se aplicaria ao caso porque “*nós acreditamos firmemente que as pessoas devem ser capazes de dizer algo que é errado ou impreciso no Facebook, mesmo se são ofensivas*”<sup>163 164</sup>. De outro, movimentos sionistas

<sup>161</sup> Página obtida a partir da inserção do verbete “holohoax” na busca do Twitter em 25 de maio de 2019; a página é acessível pelo endereço <<https://twitter.com/search-home>>.

<sup>162</sup> SALES, B. Facebook says it will allow Holocaust denial. **The Jerusalem Post**, 4.5.2019, disponível em <<https://www.jpost.com/Diaspora/Antisemitism/Facebook-says-it-will-allow-Holocaust-denial-588651>>, acesso em 19.6.2019.

<sup>163</sup> “[...] *not remove lies or content that is inaccurate — whether it’s denying the Holocaust, the Armenian massacre, or the fact that the Syrian government has killed hundreds of thousands of its own people. This is because we do believe that people should be able to say things on Facebook that are wrong or inaccurate,*

argumentaram que negar o Holocausto seria uma forma de antissemitismo, porque implicava dizer que os judeus teriam manipulado os governos e a opinião pública nesse sentido e, por isso, deveria ser proibido<sup>165</sup>.

A controvérsia atual, entretanto, tem sua especificidade porque foi a primeira vez que o problema da negativa do Holocausto nas redes sociais veio à tona com mais força após a aprovação, na Alemanha, de uma legislação especificamente voltada ao controle do conteúdo postado em plataformas na internet – a “*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken – NetzDG*”, conhecida como “*Lei do Facebook*”<sup>166</sup>. A norma exige que plataformas da internet com mais de 2 milhões de usuários registrados na removam conteúdo evidentemente ilegal 24 horas após a notificação<sup>167</sup>.

---

*even when they are offensive.*”. SHAMSIAN, J. Facebook's head of policy it would allow 'denying the Holocaust' in the weeks before banning high-profile anti-Semitic conspiracy theorists. **Business Insider**, 3.5.2019, disponível em <<https://jewishinsider.com/2019/05/daily-kickoff-facebook-removes-inflammatory-figures-but-will-allow-holocaust-denial-sfrc-confirms-david-schenker-nyu-tel-aviv-singled-out/>>, acesso em 19.6.2019.

<sup>164</sup> O caso da negação do Holocausto e a insistência do Facebook em não excluir postagens que veiculem informações falsas foi objeto recentemente de duas polêmicas na plataforma. Na primeira, a empresa decidiu que não iria excluir um vídeo em que imagens de Nancy Pelosi, Presidente da Câmara dos Deputados dos Estados Unidos, foram manipuladas para que ela parecesse bêbada; na segunda, usuários resolveram desafiar a política da empresa compartilhando, no Instagram – uma plataforma detida pelo Facebook, com políticas de conteúdo semelhantes – um vídeo manipulado do CEO do Facebook, Mark Zuckerberg, fazendo afirmações falsas. Cf. KANG, C. Nancy Pelosi Criticizes Facebook for Handling of Altered Videos. **The New York Times**, 29.5.2019, disponível em <<https://www.nytimes.com/2019/05/29/technology/facebook-pelosi-video.html>>, acesso em 20.6.2019; e METZ, C. A Fake Zuckerberg Video Challenges Facebook's Rules. **The New York Times**, 11.6.2019, disponível em <<https://www.nytimes.com/2019/06/11/technology/fake-zuckerberg-video-facebook.html>>, acesso em 20.6.2019.

<sup>165</sup> ADL. Quantifying Hate: A Year of Anti-Semitism on Twitter. Disponível em <<https://www.adl.org/resources/reports/quantifying-hate-a-year-of-anti-semitism-on-twitter#holocaust-denial>>, acesso em 20.6.2019.

<sup>166</sup> ALEMANHA. *Netzwerkdurchsetzungsgesetz vom 1. September 2017 (BGBl. I S. 3352)*, disponível em <<https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>>, acesso em 20.6.2019.

<sup>167</sup> § 3.º, (2) “*Das Verfahren muss gewährleisten, dass der Anbieter des sozialen Netzwerks*” [...] 2. “*einen offensichtlich rechtswidrigen Inhalt innerhalb von 24 Stunden nach Eingang der Beschwerde entfernt oder den Zugang zu ihm sperrt; [...]*” (§ 3.º, (2) “*O procedimento deve garantir que o mantenedor da rede social:*” [...] 2. “*remova ou bloqueie o acesso a conteúdo que seja manifestamente ilegal dentro de 24 horas após o recebimento da reclamação*”).

O problema surgiu, a partir do contexto acima, porque negar o Holocausto é ilegal na Alemanha<sup>168</sup> – o que significa que manifestações nesse sentido se enquadrariam dentre as que devem ser removidas da plataforma. Isso opôs, de um lado, o Facebook – que estabeleceu, como política, não retirar conteúdo que negasse o Holocausto – e, de outro, o governo alemão, que exigia isso como parte de sua própria legislação. A situação ficou mais sensível após uma pesquisa do *The Times* mostrar que “*cartuns que mostram judeus como baratas com narizes aquilinos, links para sites que vendem livros sobre o ‘holohoax’ banidos pelas principais lojas e fanpages de negadores do holocausto condenados*”<sup>169</sup> foram denunciados no Facebook e permaneciam ativos na plataforma, que via, de regra, disponibiliza o mesmo conteúdo para todos os países em que atua.

Quanto a esse específico contexto – e em uma reação à entrevista de Zuckerberg – a Ministra da Justiça da Alemanha deixou clara a posição do governo a respeito, associando a negativa do holocausto ao antissemitismo: “*Não deve haver lugar para o antissemitismo*”, afirmou, “*e isso inclui ataques verbais e físicos aos judeus, assim como a negativa do Holocausto*”<sup>170</sup>; um porta-voz disse, ainda, que os padrões utilizados pela plataforma nos Estados Unidos não poderiam ser aplicados na Alemanha<sup>171</sup>. Essas declarações estão alinhadas com posições que, como citei acima, veem na negação do holocausto um tipo de antissemitismo, um debate que ensejou posições, por exemplo, favoráveis ao Facebook, no sentido de que essa é uma vedação que tem “*efeitos negativos na liberdade de expressão*”<sup>172</sup>.

---

<sup>168</sup> ALEMANHA. Strafgesetzbuch in der Fassung der Bekanntmachung vom 13. November 1998 (BGBl. I S. 3322), § 130 Volksverhetzung, disponível em <[https://www.gesetze-im-internet.de/stgb/\\_130.html](https://www.gesetze-im-internet.de/stgb/_130.html)>.

<sup>169</sup> GIBBONS, K. Antisemitic hate posts allowed by Facebook. **The Times**, 27.7.2018, disponível em <<https://www.thetimes.co.uk/article/antisemitic-hate-posts-including-holocaust-denial-allowed-by-face-book-7crdvw2x>>, acesso em 19.6.2019.

<sup>170</sup> “*There must be no place for antisemitism. This includes verbal and physical attacks on Jews as well as the denial of the Holocaust.*”. SCHUMACHER, E. Facebook refuses to censor Holocaust denial as social media sites struggle with German laws. **Deutsche Welle**, 27.7.2018, disponível em <<https://www.dw.com/en/facebook-refuses-to-censor-holocaust-denial-as-social-media-sites-struggle-with-german-laws/a-44855519>>, acesso em 19.6.2019.

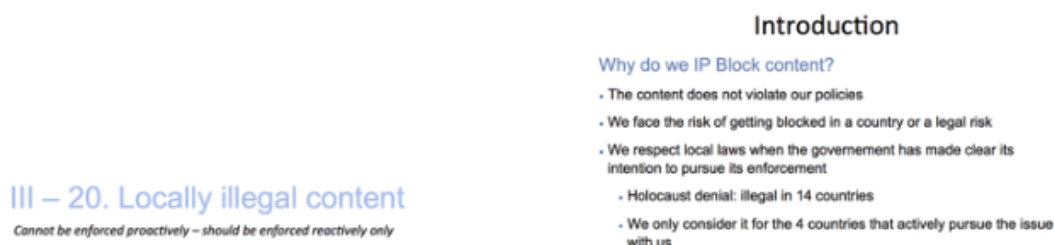
<sup>171</sup> DELCKER, J. Germany to Zuckerberg: There won't be Holocaust denial on German Facebook. **Politico**, 19.7.2018, disponível em <<https://www.politico.eu/article/germany-to-zuckerberg-there-wont-be-holocaust-denial-on-german-facebook/>>, acesso em 19.6.2019.

<sup>172</sup> “*Criminalization and bans also have a negative effect on free speech*”. ROSE, F. Zuckerberg was right about how to handle Holocaust deniers. **The Washington Post**, 30.7.2018, disponível em

A solução encontrada pela plataforma – e inicialmente detectada em 2017, após documentos utilizados para treinar os responsáveis por analisar o conteúdo publicado na plataforma vazarem – foi estabelecer geobloqueios, isto é, restrições a determinados conteúdos de acordo com a origem do acesso à plataforma. A publicação desses documentos – denominados, à época, de “*Facebook Files*”<sup>173</sup> – forneceu um raro acesso ao modo como a companhia lida com as legislações locais.

Em uma série de páginas de uma apresentação em Power Point iniciada com o título “*Conteúdos localmente ilegais*”<sup>174</sup>, o Facebook explica aos responsáveis por aplicar a política de conteúdo como lidar com postagens que neguem o Holocausto. De início, o geobloqueio é explicado pela plataforma como algo justificável quando o “*conteúdo não viola as nossas políticas*”, mas “*nós sofremos ou risco de sermos bloqueados em um país ou um risco jurídico*”, ressaltando-se que “*nós respeitamos as leis locais quando o governo deixa clara a sua intenção de buscar aplica-lo*”.

**Figura 8 – Páginas de PowerPoint do material utilizado para treinar os moderadores**



---

<[https://www.washingtonpost.com/news/worldpost/wp/2018/07/30/facebook/?noredirect=on&utm\\_term=.39567f50c7f5](https://www.washingtonpost.com/news/worldpost/wp/2018/07/30/facebook/?noredirect=on&utm_term=.39567f50c7f5)>, acesso em 19.6.2019.

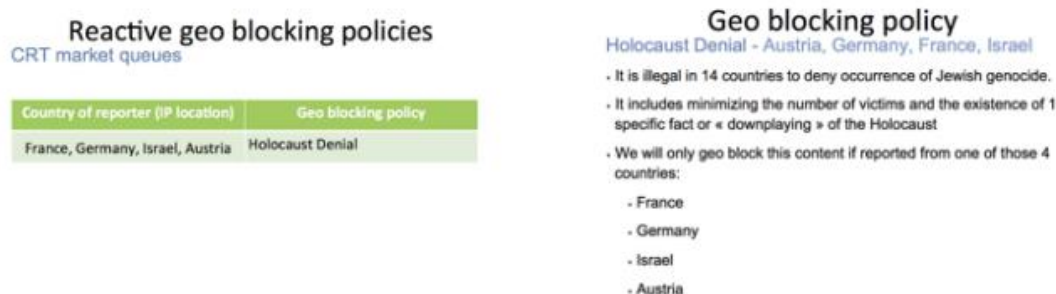
<sup>173</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 120 (Kindle, n. p.).

<sup>174</sup> Os documentos de treinamento dos moderadores de conteúdo do Facebook foram reunidos em um vazamento denominado “*Facebook Files*”. As imagens de *Power Point* que seguem foram retiradas da reportagem realizada pelo *The Guardian*. THE GUARDIAN. How Facebook handles Holocaust denial. **The Guardian**, 24.5.2017, disponível em <<https://www.theguardian.com/news/gallery/2017/may/24/how-facebook-handles-holocaust-denial>>, acesso em 19.6.2019.

Fonte: The Guardian<sup>175</sup>

A página seguinte traz como exemplo justamente a negação do Holocausto. Segundo o Facebook explica no material, apesar de ser ilegal fazê-lo em quatorze países, a plataforma faz o geobloqueio apenas em quatro – França, Alemanha, Israel e Áustria – por considera-los locais em que a proibição é efetivamente aplicada ao ponto em que descumpri-la pode gerar riscos à sua atuação:

**Figura 9 – Páginas de PowerPoint do material utilizado para treinar os moderadores**



Fonte: The Guardian<sup>176</sup>

Nos casos em que há geobloqueio, portanto, o conteúdo não é retirado da plataforma: ele apenas é tornado invisível para pessoas que o acessem de determinadas localizações. Com isso, a política de permissibilidade do Facebook quanto a publicações que neguem o Holocausto permanece vigente; ela apenas é atenuada para quem estiver se conectando à internet de um computador localizado, por exemplo, na Alemanha.

### 3.3 A Moderação de conteúdo das plataformas da internet

#### 3.3.1 Introdução

Como mencionei antes, reconstruí os casos acima com um objetivo: oferecer evidências de que a abordagem feita por Balkin no texto do início do capítulo tem um

<sup>175</sup> Cf. *supra*, nota n.º 21.

<sup>176</sup> Cf. *supra*, nota n.º 21.

problema. Em específico, sustento que a realidade da regulação da liberdade de expressão no mundo digital é mais complexa do que a que ele se propõe a analisar, porque não se resume às ferramentas que o Estado utiliza para, por exemplo, impedir que alguém divulgue uma informação. Meu argumento, nesse sentido, é de que essa transição tecnológica produziu novos espaços de manifestação administrados por atores privados – em específico, as empresas por trás de plataformas na internet como o Facebook, o Twitter e o YouTube –, que, ao constituir novos espaços de decisão, têm o poder de definir, com independência, o que pode ser dito neles, regulando a liberdade de expressão. Por isso, a única forma de se ter um quadro suficientemente completo da regulação da liberdade de expressão nessa nova realidade é expandindo os horizontes dessa análise, levando-a a abranger também o modo como esses agentes privados tomaram para si a responsabilidade de decidir aquilo que as pessoas podem ou não dizer.

Os casos acima, nesse sentido, foram selecionados como contraexemplos: eles são situações que utilizo para refutar a abordagem de Balkin, porque expõem retesamentos entre a relação que existe entre Estado e plataforma. Em específico, eles exemplificam que, ao contrário do que afirma o autor, nem sempre a atuação de agentes privados no sentido de permitir ou restringir manifestações em espaços típicos da infraestrutura digital decorre de uma ação de Estado. Na verdade, por vezes a decisão de uma plataforma da internet quanto àquilo que um usuário pode ou não dizer é o resultado de políticas de conteúdo próprias e independentes, criadas por esses atores privados a partir da necessidade de enfrentar controvérsias quanto à liberdade de expressão que são replicadas nesses novos espaços por eles criados. Assim, decisões a respeito do que é permitido ou proibido dizer nesses ambientes tanto decorrem, em alguns casos, da ação regulatória estatal – que pressiona esses atores privados a agirem em seu nome – algo que Balkin explica; quanto, em outros casos, da ação independente das próprias plataformas, que aplicam, aos seus usuários, políticas de conteúdo que elas mesmo criaram – algo que abordagens como a de Balkin não conseguem explicar. É nesse último flanco que este trabalho se concentra.

Para expor isso com mais detalhes, proponho unir a reconstrução que fiz no início deste capítulo (2) com os casos que apresentei na seção anterior (3.2). Farei isso em três etapas: na primeira (3.3.2), unifico os casos que reconstruí em três pontos de convergência que os perpassam, de modo a, deixando de lado as especificidades de cada um, extrair uma narrativa comum quanto a como agentes privados em ambientes típicos da infraestrutura digital da liberdade de expressão agiram nessas situações. Na segunda

(3.3.3), exponho essa narrativa como um contraexemplo a abordagens como a de Balkin, problematizando-a no específico ponto em que demonstram a existência de decisões provenientes das plataformas da internet que, ao mesmo tempo em que regulam a liberdade de expressão, não são remissíveis ao Estado. Terceiro (3.3.4), exponho e refuto uma defesa possível de Balkin a essa abordagem e, em conclusão (3.3.5), sustento a necessidade de se empreender um esforço de pesquisa em um sentido diferente do inicialmente traçado por ele, algo que vai ao encontro da solução por ele proposta em um texto seguinte, construindo uma ponte entre Balkin de 2014 e o de 2018.

### 3.3.2 Por uma narrativa comum

Como referi acima, os casos que reconstruí na seção anterior foram selecionados como contraexemplos à abordagem de Balkin, isto é, como situações que indicam problemas na forma como ele se dispôs a analisar a regulação da liberdade de expressão na era digital. Para que isso fique mais claro, proponho, nesta seção, que as três situações acima sejam unificadas em uma narrativa única, construída a partir dos pontos de convergência entre esses casos, para, em seguida, opô-la àquilo que sustenta Balkin.

O primeiro ponto em comum entre esses três casos é a forma como eles começam. Todos eles tratam de controvérsias que emergem, inicialmente, a partir da veiculação de determinados conteúdos em plataformas da internet por seus usuários. Em particular, eles são ocasionados por manifestações que ocorreram em ambientes detidos por empresas de tecnologia – a exemplo do Facebook e do Google –, que mantêm espaços na internet destinados não a expor as suas próprias posições, mas que são dedicados a veicular aquilo que os seus usuários decidem publicar. Assim, no primeiro caso, a discussão ocorre em torno da publicação da imagem “*O Horror da Guerra*” por Tom Egeland no Facebook; no segundo, em razão dos vídeos que Alex Jones veiculou no YouTube – e, depois, no Twitter e no Facebook –, em particular o que nega a veracidade dos tiroteios em escolas dos Estados Unidos; e, no terceiro, é causada pelas manifestações de usuários do Facebook, que afirmam que o Holocausto não aconteceu.

Nesse sentido, e sob uma perspectiva de Balkin, esses casos são típicos acontecimentos inseridos na infraestrutura digital da liberdade de expressão<sup>177</sup>. Todos eles dizem respeito a manifestações que, *quanto ao nível de tecnologia*, estão vinculadas à que

---

<sup>177</sup> Cf. *supra*, Quadro n.º 3.

viabiliza a internet; que, *quanto àquela que se manifesta*, são confeccionadas pelos usuários e não pelos administradores desses espaços; e que, *quanto aos detentores desses espaços*, ocorrem em ambientes privados detidos por conglomerados de tecnologia como o Google, o Facebook, e o Twitter.

O segundo ponto em comum entre esses casos diz respeito ao momento imediatamente posterior às manifestações que mencionei no item anterior. Em síntese: uma vez veiculadas essas manifestações nas plataformas, surgiram, nos três casos, discussões quanto a se aquilo que disseram esses usuários deveria ou não permanecer na internet. Essas polêmicas, entretanto, não foram direcionadas apenas às pessoas responsáveis pelas publicações, mas tiveram outro alvo direto: as empresas que, por estarem por trás dessas plataformas, detêm o controle sobre a infraestrutura digital que as viabiliza e que, nessa posição, foram instadas a decidir, nos três casos, se a publicação de seus usuários deveria ou não permanecer ativa.

Assim, se os três casos acima começaram a partir de um ato específico de um usuário, eles se desdobraram em discussões não apenas em relação à conduta do particular, mas sim quanto ao modo como as plataformas em que veiculadas essas manifestações deveriam agir. No primeiro caso, a publicação de uma fotografia da Guerra do Vietnã no Facebook gerou uma discussão quanto a se a plataforma deve ou não permitir esse tipo de imagem em seu ambiente; no segundo, em que um tiroteio em uma escola pública foi noticiado como uma farsa, houve um debate quanto a se Facebook, Twitter e YouTube deveriam banir alguém que dissemina notícias falsas; por fim, no terceiro, manifestações de que o Holocausto não teria existido geraram uma discussão sobre em que medida o Facebook deve ou não vedar esse tipo de publicação na plataforma.

Esse direcionamento fica particularmente claro no segundo e no terceiro casos: o banimento de Alex Jones e do Infowars, nesse sentido, era uma exigência que há muito era endereçada às plataformas da internet, sob o argumento de que elas teriam sido instrumentalizadas para disseminar discursos prejudiciais a minorias e falsas notícias produzidas pelo canal<sup>178</sup>. O Facebook, por exemplo, fez isso depois de ser acusado de não

---

<sup>178</sup> Cf., por exemplo, PEARCE, M., PIERSON, D. Tech companies gave massive platforms to conspiracy theorists like Alex Jones. Is the crackdown finally here? **Los Angeles Times**, 7.8.2018, disponível em <<https://www.latimes.com/nation/la-na-alex-jones-20180807-story.html>>, acesso em 20.6.2019.



ter políticas consistentes para decidir o que deve permanecer ou não na plataforma<sup>179</sup>. Algo muito semelhante aconteceu no caso das páginas que negavam o Holocausto: assim que a entrevista de Zuckerberg trouxe a posição permissiva do Facebook a respeito à tonalidade novamente, a plataforma – e não os seus usuários – foi diretamente acusada de ser responsável pela disseminação de “*desinformações de ódio quanto ao Holocausto*”:

*“Essa posição [do Facebook quanto a quem nega o Holocausto] é tão bizarra, que é difícil saber por onde começar. Primeiramente, certamente espalhar desinformações de ódio quanto ao Holocausto – pensadas para enganar as massas e minar a consciência social do preconceito antissemita histórico – constitui uma ‘tentativa de organizar danos’. Além disso, mesmo se nós aceitarmos o argumento questionável de Zuckerberg que alguns negam o Holocausto mais por ignorância do que por malícia, isso não absolve o Facebook da responsabilidade por, de maneira acrítica, hospedar e disseminar esse conteúdo”<sup>180</sup>.*

As empresas por trás dessas plataformas, nesse sentido, por deterem a infraestrutura da liberdade de expressão utilizada para veicular dessas manifestações – e exercerem, sobre ela, o poder de decidir o que pode permanecer ali – são instadas a tomar uma decisão – seja por uma vítima do caso Sandy Hook que se sente ofendido quando Alex Jones o acusa de ser um ator<sup>181</sup> ou por uma associação de judeus que é atacada pela existência de uma página que nega o Holocausto<sup>182</sup> – quanto a se aquele conteúdo deve ou não ficar na internet. Concretamente, elas foram provocadas, nos casos, a utilizar o domínio que têm sobre essa nova infraestrutura da liberdade de expressão para decidir se Tom Egeland pode publicar uma imagem que expõe o sofrimento de uma criança nua

---

<sup>179</sup> ROOSE, K. Facebook Banned Infowars. Now What? **The New York Times**, 10.8.2018, disponível em <<https://www.nytimes.com/2018/08/10/technology/facebook-banned-infowars-now-what.html>>, acesso em 19.6.2019.

<sup>180</sup> “*This position is so bizarre, it’s hard to know where to begin. For one, surely spreading hateful misinformation about the Holocaust—designed to mislead the masses and undermine societal awareness of historic anti-Semitic prejudice—constitutes “trying to organize harm.” Moreover, even if we accept Zuckerberg’s questionable claim that some deny the Holocaust out of ignorance rather than malice, this does not absolve Facebook of responsibility for uncritically hosting and spreading that content*”. ROSENBERG, Y. Mark Zuckerberg Is Doubly Wrong About Holocaust Denial. **The Atlantic**, 19.7.2018, disponível em <<https://www.theatlantic.com/ideas/archive/2018/07/the-solution-to-holocaust-deniers-on-facebook/565562/>>, acesso em 19.6.2019,

<sup>181</sup> Cf. *supra*, Cap. I.1, 3.2.2.2.

<sup>182</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

apenas porque ela é uma fotografia de importância histórica; se Alex Jones pode oferecer versões alternativas ao tiroteio que ocorreu na escola Sandy Hook, mesmo que elas sejam falsas; ou, por fim, se devemos permitir que as pessoas neguem o Holocausto<sup>183</sup>.

O terceiro ponto em comum entre esses casos é uma consequência do anterior. Em síntese: em todos eles, após pressões nesse sentido, decisões quanto a se um usuário de uma plataforma na internet poderia ou não veicular uma determinada manifestação foram tomadas pelos atores privados que administram esses espaços, adotando, para isso, normatizações internas criadas por eles para esse fim.

Como expus no item anterior, os casos acima exemplificam situações em que a veiculação de conteúdos em infraestruturas típicas do mundo digital ocasionou divergências por parte de outros usuários e terceiros, que as direcionaram não apenas a quem disponibilizou o conteúdo, mas também às companhias que administram tais espaços, que, por controlarem essas estruturas, foram instadas a decidir a respeito.

Esse terceiro ponto de convergência decorre disso: em primeiro lugar, nos três casos, as companhias por trás dessas plataformas tomaram para si essa responsabilidade, produzindo e aplicando decisões quanto ao conteúdo que deveria ser permitido nos ambientes que administravam. No primeiro caso, a decisão foi de proibir a veiculação da imagem “*O Horror da Guerra*” – o que gerou a exclusão de publicações de jornais e autoridades; no segundo, de impedir Alex Jones e o Infowars de se manifestarem em plataformas da internet – razão pela qual seus perfis foram banidos; no terceiro, decidiu-se que manifestações que negam o Holocausto seriam permitidas no Facebook – e, por isso, páginas a respeito permaneceriam visíveis, exceto em países como Alemanha e Israel.

Segundo, essas decisões foram tomadas – e mesmo modificadas – a partir de códigos internos, estabelecidos por essas plataformas para fixar o que os seus usuários poderiam dizer por meio delas. Assim, no caso em que discutida a publicação de Tom Egeland, o movimento do Facebook de inicialmente vedar a publicação e, em seguida, autorizá-la, é expresso a partir de uma mudança de suas políticas internas – que são chamadas, nas manifestações da plataforma, de “*padrões de comunidade*”. A fala do Vice-

---

<sup>183</sup> Essa não é a única via pela é dada vazão a divergências quanto àquilo que é dito em plataformas na internet: como o próprio exemplo da negativa do Holocausto demonstra, grupos sociais organizados mobilizam diferentes instrumentos para fazer prevalecer suas posições, como, por exemplo, a edição de leis como a alemã, que acabou por vedar a permanência desse tipo de manifestação em plataformas do país; *cf. supra*, Cap. I.1, 3.2.2.3.

Presidente da plataforma quanto a isso é especialmente esclarecedora, porque ele pontua tanto a existência de regras anteriores ao caso – *“o Facebook tem padrões de comunidade para manter as pessoas seguras e criar uma experiência positiva em nossa plataforma”* – quanto a alteração dessas políticas, decorrente da situação enfrentada pela empresa quanto à publicação da imagem por Egeland – materializada tanto na constatação de que *“Às vezes, entretanto, a importância global e histórica de uma foto como ‘O Horror da Guerra’ superam a importância de se manter a nudez fora do Facebook”* quanto no compromisso de *“melhorar nossas políticas para garantir que elas promovam a livre expressão e mantenham a comunidade segura”*<sup>184</sup>.

A mesma situação ocorre no segundo caso: cada uma das plataformas que banuiu Alex Jones e o Infowars deixou explícito que o fez porque eles infringiram as suas respectivas políticas de conteúdo. O YouTube afirmou que *“todos os usuários concordam em obedecer os nossos Termos de Serviço e Padrões de Comunidade quando se inscrevem para usar o YouTube”*; assim, *“Quando usuários violam repetidamente essas políticas, como as nossas políticas contra o discurso de ódio e assédio, ou os nossos termos que proíbem atos que contornem os nossos meios de executá-las [enforcement], nós encerramos as suas contas”*<sup>185</sup>; o Twitter, por sua vez, excluiu as contas à vista de *“novas denúncias de Tweets e vídeos postados ontem [5.9.2018] que violavam nossa política sobre comportamento abusivo, adicionalmente às violações anteriores da conta”*<sup>186</sup>. O Facebook, em um anúncio à parte – e que reconstruí com mais detalhes acima – é especialmente minucioso a esse respeito: a plataforma explicou, antes do banimento, que havia mudado a sua política de conteúdo interno, para vedar também manifestações de nacionalismo e separatismo branco. Como as manifestações de Jones e do Infowars se enquadravam nesses novos critérios, ambos foram banidos da plataforma.

No terceiro caso, também há uma decisão – a de permitir manifestações que neguem o Holocausto no Facebook – que é posta como uma consequência de uma política interna da plataforma, aplicada, no caso, para tomar decisões quanto à permanência de páginas criadas por seus usuários. Assim, descrevendo um capítulo dessa política, seu Vice-Presidente, disse que *“não vai remover mentiras ou conteúdos que são imprecisos – não importa se se trata de negar o Holocausto, o massacre armênio ou o fato de que o*

---

<sup>184</sup> Cf. *supra*, notas n.º 127 e 128.

<sup>185</sup> Cf. *supra*, nota n.º 139.

<sup>186</sup> Cf. *supra*, nota n.º 140.

governo sírio matou centenas de milhares do seu próprio povo”, justificando que, a despeito da vedação, no Facebook, ao discurso de ódio, isso não se aplicaria ao caso porque “nós acreditamos firmemente que as pessoas devem ser capazes de dizer algo que é errado ou impreciso no Facebook, mesmo se são ofensivas”<sup>187</sup>.

Em conclusão: todos os casos acima se reúnem em torno de uma narrativa comum, composta por três pontos de convergência. Os três se iniciam com a veiculação de manifestações por usuários plataformas da internet, que o fazem, portanto – e de acordo com a classificação de Balkin –, em típicos ambientes da infraestrutura digital da liberdade de expressão. A partir dessas manifestações, surgiram divergências quanto a se aquilo que os usuários desses serviços publicaram devia ou não permanecer ativo, que foram direcionadas não apenas aos autores das publicações, mas também às companhias por trás das plataformas que, por deterem a infraestrutura que as viabiliza, foram instadas a se posicionarem a respeito da permanência ou não dessas manifestações em seus respectivos ambientes. Por fim, nos três casos essas companhias tomaram decisões, permitindo ou proibindo determinadas manifestações nas plataformas segundo códigos internos construídos por elas para esse fim. O quadro abaixo sintetiza essa sequência:

**Quadro 4 – Narrativa comum**

	<b><u>Fase 1</u></b> Publicação	<b><u>Fase 2</u></b> Controvérsia	<b><u>Fase 3</u></b> Decisão
	O usuário veicula um conteúdo próprio na internet por meio da plataforma.	Usuários e terceiros controvertem a respeito se aquilo que o usuário publicou deve ou não permanecer na internet, instando as plataformas a decidirem.	A partir de critérios próprios criados para esse fim, as plataformas decidem manter ou excluir a publicação de seu usuário.
<b>Caso 1</b>	Tom Egeland publica a imagem “ <i>O Horror da Guerra</i> ” no Facebook.	A moderação de conteúdo do Facebook, usuários e o governo da Noruega divergem quanto à manutenção da imagem.	O Facebook inicialmente decide excluir a imagem; depois, a partir de uma mudança de política interna, resolve mantê-la.
<b>Caso 2</b>	Alex Jones publica, em diversas plataformas, informações falsas.	Conglomerados de mídias impressas e televisivas ( <i>The New York Times</i> e <i>CNN</i> , p. ex.)	Alex Jones é banido do Facebook, Twitter e YouTube.

<sup>187</sup> Cf. *supra*, nota n.º 163.

<b>Caso 3</b>	Usuários veiculam no Facebook mensagens que negam a existência do Holocausto.	Grupos sionistas, ativistas da liberdade de expressão, usuários e a plataforma, por seu CEO, controvertem quanto à permanência de tais mensagens.	O Facebook decide manter, na plataforma, publicações que negam o Holocausto.
---------------	---	---	--

### 3.3.3 O problema de Balkin

Como mencionei antes, reuni os casos acima como contraexemplos à abordagem de Balkin em *Old-School / New-School Speech Regulation*, expondo que a realidade do mundo digital é mais complexa do que a que ele se propõe a analisar. Nesse sentido, minha intenção, ao descrevê-los, é expor como a atuação das empresas por trás das plataformas da internet, restringindo ou permitindo determinados conteúdos, pode ocorrer de forma paralela – e, por vezes, divergente – à dos Estados. Isso sugere que decisões a respeito do que as pessoas podem ou não dizer nesses ambientes também resultam da aplicação independente de políticas próprias desses atores privados – que, por isso, nem sempre agem como meros delegatários do Estado, algo que justifica promover um estudo à parte quanto à sua atuação.

Para explicar melhor isso, é preciso retomar, por um momento, o argumento de Balkin. Segundo ele, mudanças sociais associadas à evolução tecnológica teriam propiciado uma transformação na infraestrutura da liberdade de expressão, fazendo-a transitar de um modelo pré-digital para uma era digital. De acordo com Balkin, para diagnosticar como isso impactou na regulação da liberdade de expressão, bastaria analisar como as ferramentas utilizadas pelo Estado se transformaram nesse período. Ele propõe, por isso, uma espécie de taxonomia da atuação regulatória estatal, que associa tanto à infraestrutura pré-digital quanto à infraestrutura digital modelos específicos de intervenção do Estado na liberdade de expressão, respectivamente denominados de “*old-school*” e “*new-school*”.

Nesse quadro, o papel dos agentes privados estaria restrito ao de delegatários da ação estatal: a partir de um conjunto de mecanismos que vai da censura colateral ao *soft power*, o Estado seria capaz de influenciar as companhias por trás de plataformas como o Facebook, o Twitter e o YouTube, para que elas agissem em seu nome e a seu favor, restringindo, por exemplo, a disponibilização de conteúdos entendidos como indesejáveis aos posicionamentos oficiais. Por isso, a atuação de atores privados na regulação da liberdade de expressão na era digital seria, na verdade, o de executar uma

ação estatal por delegação, à qual o comportamento do particular seria remissível. Disso, Balkin extrai o que seria a melhor estratégia para lidar com essa transição de infraestruturas: como é o Estado que sempre regula a liberdade de expressão – não importa se diretamente, proibindo algo, ou indiretamente, induzindo alguém a proibir algo –, então é na sua atuação que devemos buscar respostas que expliquem o que mudou, por exemplo, com o surgimento de plataformas como o Facebook na internet.

Reconstruí, como disse, os casos acima como contraexemplos a essa abordagem. Todos têm uma narrativa comum: a veiculação de uma manifestação em uma ou várias plataformas da internet, ambientes típicos da infraestrutura da era digital; a polêmica que disso seguiu e foi endereçada às companhias responsáveis por essas plataformas, instando-as a decidir se algo deveria permanecer na internet; e a tomada de decisão, por essas empresas, que proibiram ou permitiram determinadas manifestações de acordo com códigos internos criados para esse fim.

Até esse ponto, a abordagem de Balkin e a narrativa desses casos não apresentam divergências aparentes. Para o autor, é perfeitamente concebível que agentes privados atuem restringindo aquilo que as pessoas podem dizer nas plataformas da internet, de acordo com o que determina o Estado. Balkin, inclusive, vê isso como algo característico da era digital<sup>188</sup>: como nessa nova infraestrutura os intermediários são identificados com muito mais facilidade que os indivíduos que efetivamente veiculam o conteúdo, faz mais sentido que a censura estatal atue por meio dessas plataformas, alinhando o que fazem ao que exige o Estado. Nem mesmo a existência de políticas internas de conteúdo nessas companhias é um fato suficiente, por si só, para afetar a abordagem de Balkin: basta que essas políticas estejam alinhadas àquilo que exige o Estado para que a sua explicação – e a sua abordagem – continuem fazendo sentido.

O problema, que fica exposto pelos casos acima, é que esse alinhamento entre, de um lado, a exigência de um Estado – que faz valer uma posição por meio de um agente privado; e, de outro, a atuação de uma plataforma – que serve de um instrumento da vontade do Estado – não é o que acontece na realidade.

Não há dúvidas que, às vezes, essa convergência acontece: há situações em que plataformas da internet restringem as manifestações de seus usuários a partir do que

---

<sup>188</sup> Essa possibilidade aparece, em específico, no momento em que Balkin explica o mecanismo da censura colateral, pelo qual estímulos são criados pelo Estado para que particulares cooperem atuando de forma restritiva quanto àquilo que é veiculado por meio de suas plataformas na internet. *Cf. supra*, nota n.º 96.

um Estado determina. Assim, no terceiro caso que reconstruí acima, com a edição de uma legislação que proibiu manifestações contrárias à existência do Holocausto na internet, o Facebook decidiu bloquear localmente o acesso a esses conteúdos, aplicando o que determinou o governo alemão. Essa, entretanto, é uma situação dentre as várias possíveis na era digital. Como os casos acima demonstram, faz parte dessa nova realidade de infraestrutura da liberdade de expressão conjunturas em que a exigência dos Estados e a política das plataformas divergem e nas quais, mesmo assim, essas companhias tomam decisões que permitem ou proíbem os seus usuários de dizerem algo. O alinhamento Estado-plataforma, nesse sentido, não é um pressuposto para a tomada de decisão do agente privado. Isso sugere que, mais do que emissárias do Estado, essas companhias têm as suas próprias políticas de conteúdo, que, ao menos em algumas situações, são aplicadas independentemente do que lhes é exigido por atores estatais. É nesse sentido que os casos acima são contraexemplos à abordagem de Balkin: eles expõem circunstâncias em que agentes privados tomam decisões em ambientes típicos da infraestrutura digital, permitindo ou proibindo que seus usuários digam algo, mesmo que suas políticas de conteúdo sejam – por especificidades de cada uma das situações – claramente divergentes das posições estatais a esse respeito. Eles demonstram, em síntese, que se é verdade que às vezes os Estados atuam por meio dessas plataformas, não é toda decisão tomada por elas que pode ser entendida a partir de uma remissão à ação estatal.

Assim, no primeiro caso, em que discutida a publicação de uma imagem no Facebook, é possível depreender das especificidades do contexto que a exclusão da postagem do jornalista Tom Egeland foi feita não em razão das posições, por exemplo, do governo da Noruega – local em que a controvérsia se desenvolveu –, mas sim motivada por uma política própria da plataforma, que proíbe tanto a nudez quanto imagens que, pelo seu conteúdo, causam impacto em usuários mais sensíveis. Há pelo menos dois indicativos disso: primeiro, na divergência entre o Facebook – que defendeu a impossibilidade de publicação da imagem – e o governo da Noruega – que sustentava o contrário, algo exposto inclusive na postura das autoridades governamentais em manifestações na própria plataforma, no sentido de que a publicação do jornalista local era um ato de liberdade de expressão e que, como tal, não deveria sofrer restrições. Na situação que teve maior repercussão, a própria Primeira Ministra do país, Erna Solberg, veio a público defender o direito do jornalista de veicular a fotografia, criticando o Facebook por contribuir “*para*

*restringir a liberdade de expressão” ao “censurar imagens como essas”*<sup>189</sup>. Nesse âmbito, o Facebook não só discordou do que dizia o Estado local, como deletou a imagem original e as manifestações da Primeira Ministra e das demais autoridades do país.

Segundo, pela circunstância de que a decisão do Facebook de excluir a imagem – e as postagens seguintes, que a veicularam em protesto –, foi extensamente fundamentada pela plataforma em suas próprias políticas de conteúdo e não em uma determinação alheia, que teria sido apenas por ela operacionalizada. Isso fica claro tanto no momento em que a exclusão foi feita – à época, a moderação de conteúdo do Facebook disse que estabelecia, como política, *“limitações na visualização de nudez para limitar a exposição das diferentes pessoas que utilizam a nossa plataforma a conteúdo sensível”*; quanto no momento em que ela foi revertida – um movimento explicado pelo Facebook como a criação de uma exceção na política anterior, porque se ponderou que *“a importância global e histórica de uma foto como ‘O Horror da Guerra’”* supera *“a importância de se manter a nudez fora do Facebook”*. Proibir ou permitir a imagem na plataforma, portanto, não é algo que decorreu da forma como o Facebook funciona como um delegatário de um Estado com o que está alinhado – mesmo porque, no caso, esse alinhamento não existia – mas sim do resultado da aplicação de uma política própria da empresa, por ela criada e modificada.

No segundo caso, em que discutida a exclusão, dentre outras figuras, de Alex Jones e do Infowars de comunidades na internet, esse desalinho entre política de Estado e política de plataforma – que, novamente, não impediu que decisões fossem tomadas – fica ainda mais evidente. Como na situação anterior, essa é uma divergência de dois lados: de um, o progressivo movimento de restrição de plataformas como Facebook, Twitter e YouTube, que começou com a suspensão de funcionalidades das contas de Jones, até o seu banimento; de outro, uma política de Estado, especificamente defendida pelos Estados Unidos, no sentido de que figuras da extrema-direita exerciam o seu direito à liberdade de expressão quando faziam publicações nas plataformas da internet. Nesse sentido, o Presidente dos Estados Unidos opôs-se à *“censura de CIDADÃOS*

---

<sup>189</sup> *“I appreciate the work of Facebook and other media to stop pictures and content showing abuse and violence,” the PM wrote in an accompanying message. ‘But Facebook gets it wrong when it censors pictures like these. It contributes to restricting the freedom of speech.’”*. PRICE, R. Facebook just censored the prime minister of Norway for posting one of the most famous photos ever. **Business Insider**, 9.9.2016, disponível em <https://www.businessinsider.com/facebook-censors-norway-prime-minister-erna-solberg-napalm-girl-vietnam-photo-aftenposten-2016-9>, acesso em 19.6.2019.



*AMERICANOS em plataformas de mídias sociais*”, porque “*estes são os Estados Unidos da América – e nós temos o que é chamado de LIBERDADE DE EXPRESSÃO*”. A situação escalou a um ponto de divergência entre plataformas e Estado que, como relatei acima, a Casa Branca institucionalizou uma política de recepção de denúncias contra empresas que censuravam americanos, que teriam visto suas “*contas serem suspensas, banidas ou denunciadas de forma fraudulenta por ‘violações’ pouco claras de políticas de usuários*”.

O caso é especialmente interessante porque, mesmo com esse tipo de pressão, até hoje as contas dos ativistas permanecem bloqueadas, um movimento que, à semelhança da situação com o jornalista da Noruega, também foi explicado a partir da política de conteúdo dessas empresas – algo explícito, por exemplo, na forma como a retirada de Jones do Facebook ocorreu após uma mudança em algumas cláusulas desse código interno, incluindo o separatismo branco entre as manifestações vedadas aos usuários. Aqui também não parece que as plataformas estejam operando restrições a partir da influência estatal, mas, antes, tomando decisões próprias, que vão de encontro com a política de Estado vigente quanto às manifestações que devem ser permitidas na internet. Não faz o menor sentido, nesse âmbito, ver os agentes privados que baniram Alex Jones das plataformas da internet que administram como operadores de restrições estatais: antes, eles são executores de uma política própria e paralela, capaz de definir aquilo que as pessoas podem ou não dizer ali.

Por fim, o terceiro caso traz um exemplo interessante, porque expõe a complexidade da relação que existe entre Estados e plataformas – e que, novamente, não é captada por uma percepção que os vê sempre alinhados, como a de Balkin. A situação reconstruída acima – a postura do Facebook, replicada em outras plataformas, de não excluir postagens que negam o Holocausto – demonstra não só que a relação Estado e plataforma não se resume a uma instrumentalização da segunda pelo primeiro, como também que, no mesmo caso, relações distintas entre as políticas de conteúdo entre esses atores podem coexistir. Isso porque, por um lado – e como expus acima –, a situação oferece evidências no sentido de que aquilo que Balkin descreve efetivamente acontece. Pressionada pela Alemanha – e por três outros países que também estabeleceram legislações rígidas a respeito da negatividade do Holocausto –, o Facebook passou a restringir o acesso a determinadas publicações, a depender da localização do usuário. Trata-se de um exemplo típico do que Balkin expõe em seu artigo: a partir de mecanismos de indução próprios da atuação estatal – como a imposição de multas –, países como a Alemanha

impelem agentes privados a adotarem políticas estatais quanto à liberdade de expressão, que se refletem em restrições quanto àquilo que seus usuários podem dizer. Tem-se, assim, uma ação privada remissível a uma ação estatal: se o Facebook proíbe alguém de negar o Holocausto – ou, pelo menos, impede que isso seja visto por algumas pessoas –, ele faz isso a partir de uma proibição anterior, vinda do Estado, que regula, nesse caso, a liberdade de expressão.

Por outro lado, a solução fornecida pelo Facebook – não simplesmente bloquear essas manifestações na plataforma, mas utilizar a ferramenta geobloqueio, pela qual apenas as pessoas localizadas nesses países não terão acesso ao conteúdo que nega o Holocausto – funciona também como uma evidência desse paralelismo regulatório entre Estados e plataformas. Isso porque, visto sob a perspectiva de uma plataforma que funciona no mundo – e não apenas em um país –, o geobloqueio é uma exceção local que prova a regra – a vigência da política de conteúdo do Facebook, segundo a qual pessoas que negam o Holocausto devem poder dizer isso na internet. Ele torna ainda mais claro que essas empresas têm políticas próprias – às vezes excepcionadas pelo que exigem os países –, que vigem a partir de decisões internas quanto àquilo que deve ser permitido ou proibido nas plataformas. É nesse sentido que o caso é um interessante exemplo da complexidade da relação que existe entre plataformas e governos: visto sob a perspectiva do geobloqueio feito pelo Facebook a esse tipo de publicação, ele é um exemplo de como plataformas às vezes cedem àquilo que determinam os Estados; visto, entretanto, sob a perspectiva de que essas páginas continuam existindo e permanecem acessíveis de outro países, o caso é um exemplo de como essas plataformas também instituem e aplicam políticas próprias e que, muitas vezes, não estão alinhadas – nem são, portanto, consequências – das políticas dos Estados. Por isso, entender o que fazem essas empresas quando tomam decisões sobre o que as pessoas podem dizer como uma consequência do que exigem os Estados, como faz Balkin, é uma ferramenta útil para explicar casos como o da Alemanha e de Israel – em que há o bloqueio em razão do que determinam esses países –, mas insuficiente para compreender o que ocorre no resto do mundo – em que não há esse bloqueio, em obediência a uma política da própria plataforma.

Em resumo: empresas e Estados nem sempre estão de acordo quanto àquilo que os usuários de plataformas na internet podem dizer. Por isso, mesmo que às vezes a atuação das plataformas em relação ao conteúdo que é disponibilizado por meio delas decorra de uma ação estatal prévia, é um erro entender, como Balkin faz, que essa é a única forma pela qual a liberdade de expressão é regulada no mundo digital. Pelo

contrário: como os exemplos acima demonstram, plataformas da internet têm políticas próprias quanto àquilo que pode ser veiculado nesses ambientes, aplicando-as de forma relativamente independente à ação estatal, de modo a definir aquilo que as pessoas podem dizer na internet.

### 3.3.4 A defesa de Balkin

Balkin poderia oferecer uma defesa de sua abordagem de pesquisa contra os argumentos que fiz acima. Ele poderia sustentar que a restrição que estabeleceu, cingindo a análise dos impactos do mundo digital na regulação da liberdade de expressão àquilo que o Estado faz não aconteceu porque ele desconsiderou os casos acima. Pelo contrário: ele os considerou, mas não os abrangeu em sua pesquisa porque, como essas restrições não são remissíveis à ação do Estado, elas não são uma forma de regulação da liberdade de expressão.

Assim, a restrição de Balkin pareceria não só justificável, como reverberaria uma compreensão comum a respeito da liberdade de expressão. Ela refletiria, por exemplo, a perspectiva presente na jurisprudência da Suprema Corte dos Estados Unidos, que afirma que a Primeira Emenda – a cláusula que tutela, na Constituição dos Estados Unidos, a liberdade de expressão – é aplicável apenas contra restrições provenientes da ação governamental. Segundo essa visão, conhecida como “*State Action Requirement*”<sup>190</sup>, a proteção à liberdade de expressão só pode ser invocada quando aquele que a ameaça é o Estado – ou alguém que age em nome dele<sup>191</sup>. Portanto, e segundo a defesa de Balkin por esse entendimento, para que haja regulação da liberdade de expressão, é preciso que, na posição de regulador, esteja ou o Estado ou alguém que age em seu nome. Por isso, não interessaria saber o que agentes privados, como as plataformas da internet, fazem independentemente da ação estatal. De duas, uma: ou a restrição imposta pelas plataformas àquilo que dizem seus usuários é remissível ao Estado – e, por isso, é uma forma de regular a liberdade de expressão que Balkin abrange, via, por exemplo, censura colateral; ou, se

---

<sup>190</sup> BRANNON, V. Free speech and the regulation of social media content. **Congressional Research Service**, 27.3.2019, p. 5, disponível em <<https://fas.org/sgp/crs/misc/R45650.pdf>>, acesso em 19.6.2019.

<sup>191</sup> Cf., a esse respeito, por exemplo, DUNN, C. Applying the Constitution to Private Actors. **New York Civil Liberties Union**, 28.4.2009, disponível em <<https://www.nyclu.org/en/publications/column-applying-constitution-private-actors-new-york-law-journal>>, acesso em 19.6.2019.

não é remissível ao Estado – como nos casos acima –, não é uma forma de regular a liberdade de expressão, e por isso, não precisa ser abrangida esse tipo de abordagem.

Esse é um argumento que precisa ser rebatido. Isso porque o que eu proponho, neste trabalho, é uma pesquisa quanto à regulação da liberdade de expressão que não se baliza pelo “*State Action Requirement*”. Isto é: que não se restrinja àquilo que o Estado – ou alguém que age em seu nome – faz, mas que se concentre em uma situação específica, gerada pela era digital, em que particulares exercem o papel de decidir aquilo que outros particulares podem ou não podem dizer em um espaço que é privado. Assim, se essa fosse uma restrição válida, e a regulação da liberdade de expressão só existisse em contextos nos quais há ações de Estado, então circunstâncias em que particulares – usuários de plataformas – se veem sujeitos a restrições de outros particulares – empresas de tecnologia – não fariam parte de uma pesquisa do tipo que proponho nesta dissertação. Desenvolvê-la, por isso, exige uma resposta para a seguinte pergunta: por que uma pesquisa sobre liberdade de expressão deve se preocupar com as decisões que as empresas por trás das plataformas da internet tomam a respeito do que os seus usuários podem dizer, se essa é uma relação entre particulares?

O problema desse tipo de defesa é que ela vê os impactos decorrentes da era digital para a regulação da liberdade de expressão como uma simples transição de espaços de manifestação. Para o Balkin de 2014, a era digital seria suficientemente traduzida para a regulação da liberdade de expressão pela percepção de que aquilo que antes era falado na praça ou no jornal agora é dito nas plataformas – e, por isso, coube ao Estado que, no modelo “*old-school*”, “*controla o acesso a espaços públicos de reunião e protesto, e monopoliza, regula, apreende ou destrói as capacidades e as tecnologias para publicação e transmissão*”<sup>192</sup> tornar-se capaz de mirar, no “*new-school*”, “*os intermediários e instituições auxiliares que são cruciais para a expressão na Internet*”<sup>193</sup>.

---

<sup>192</sup> “*The state arrests, detains, or deports people; it controls access to public spaces for assembly and protest; and it monopolizes, regulates, seizes, or destroys capacities and technologies for publication and transmission like printing presses, broadcast facilities, movie projectors, videotapes, handbills, and books*” BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.306.

<sup>193</sup> “*They are often aimed at the intermediaries and supporting institutions that are crucial to Internet speech. The targets of newschool speech regulation range from ISPs and broadband providers to domain name registrars, hosting services, search engines, advertisers, and credit card companies*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.306.

Esse raciocínio, entretanto, reduz arbitrariamente as consequências do movimento de transição de infraestruturas, detectado pelo próprio Balkin, recortando um dos vários aspectos dessa transformação de modo que seja possível adequá-la a uma caracterização particular da regulação da liberdade de expressão, em específico a que a limita a uma relação em que, necessariamente, esteja presente a ação do Estado. Contextualizada em quadro mais amplo, entretanto, essa transição mostra toda a sua complexidade na medida em que fica perceptível que ela não significou apenas uma mudança do lugar em que as pessoas falam. Agora, nas plataformas, as pessoas se manifestam em um ambiente essencialmente diferente do anterior, porque, ao contrário das praças, esse é um espaço privado, tanto construído como controlado por atores empresariais. Os “*intermediários*”<sup>194</sup> a que Balkin se refere não são propriamente intermediários – isto é, entrepostos pelos quais a comunicação é transmitida e eventualmente bloqueada pela influência do Estado –, mas estruturas corporativas que exercem poder sobre os espaços em que as pessoas se manifestam, decidindo, em última análise, o que elas podem ou não dizer ali. O Estado, nesse contexto, não é mais o único agente capaz de controlar o lugar em que as pessoas falam e, nesse sentido, as “*práticas de censura [...] que não precisam ter uma conexão com o Estado*”<sup>195</sup> a que faz referência Balkin não são fatos isolados, mas os indícios de uma nova sistemática de regulação da liberdade de expressão, materializada em um mecanismo de produção de decisões empresariais quanto àquilo que as pessoas podem dizer em espaços privados – a moderação de conteúdo das plataformas da internet.

Ele existe porque as controvérsias que já se faziam presentes nas praças e nos jornais quanto àquilo que se falava ali não simplesmente desaparecem. Elas acompanham tanto esse movimento das praças às plataformas, quanto essa nova configuração de poder, indo a esses novos espaços com uma diferença importante: enquanto as manifestações das pessoas eram realizadas em espaços pré-digitais – isto é, públicos – controvérsias que surgiam a esse respeito eram naturalmente submetidos ao Estado, na medida em que ele controlava esses espaços, tomando a forma, por exemplo, de ações judiciais em que se discutia o que, como e onde as pessoas poderiam se manifestar.

---

<sup>194</sup> Cf., a esse respeito, GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 48 e ss (item “*Platforms are not like traditional intermediaries*”; Kindle, n. p.)

<sup>195</sup> “*Practices of "censorship," by contrast, need have no connection to the state*”. BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.299.

Debatia-se, nesse campo, o direito à liberdade de expressão, invocado como uma defesa no Estado à ação do Estado em ambientes estatalmente controlados.

Com era digital, entretanto, essas manifestações se deslocaram para espaços essencialmente privados, administrados pelas empresas que, por trás dessas plataformas, detêm a infraestrutura necessária para o seu funcionamento e que definem tanto a sua arquitetura quanto, em última análise, o que permanece ou não ali. A transição de espaços de manifestação, nesse sentido, leva a uma transição de poder e, conseqüentemente, a uma nova dinâmica decisória, em que esses novos controladores dos ambientes nos quais as pessoas se manifestam são chamados a decidir essas controvérsias, passando a enfrentar, eles mesmos, tanto a invocação do direito da liberdade de expressão – que, diante desse novo contexto, surge não como uma oposição à ação de um Estado, mas como uma oposição à ação de uma empresa – quanto a necessidade de fornecer posicionamentos a respeito de divergências típicas do debate quanto à liberdade de expressão, como, por exemplo, a que diz respeito a se definir em que tipo de situação o interesse público em se ter acesso a uma manifestação supera o dever de suprimi-la, porque ofensiva a alguém<sup>196</sup>.

Isso é algo visível nos casos que reconstruí acima. No primeiro, não foi só a imagem de guerra que foi ao Facebook: ela trouxe, consigo, a discussão – que remonta o momento em que a imagem foi publicada pela primeira vez<sup>197</sup> – quanto a se veicular uma fotografia em que uma criança nua foge de um ataque de *napalm* no Vietnã é algo que se justifica ou não. Nesse sentido, e como a narrativa que consolidei acima expõe, uma vez disponibilizada a imagem, a controvérsia que surgiu não foi submetida ao Estado, mas à plataforma, que, como um agente privado que exerce poder sobre aquele espaço, tomou a decisão de excluí-la porque ela violava as políticas internas criadas para determinar o conteúdo que seria permitido naquele ambiente. Os termos dessa controvérsia são familiares: eles incorporam o direito à liberdade de expressão como parte da gramática

---

<sup>196</sup> Cf. a discussão no Twitter que resultou em uma nova política na plataforma, segundo a qual políticos ou candidatos a cargos públicos terão suas publicações visíveis mesmo que elas violem as políticas de empresa, porque “há certos casos em que pode haver um interesse público em ter acesso a alguns Tweets, mesmo que eles, de outro modo, sejam uma violação às nossas regras”. TWITTER SAFETY. Defining public interest on Twitter. **Twitter Blog**, 27.6.2019, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2019/publicinterest.html](https://blog.twitter.com/en_us/topics/company/2019/publicinterest.html)>, acesso em 23.7.2019

<sup>197</sup> Cf. o relato quanto à existência de uma controvérsia semelhante no momento em que a foto foi publicada pela mídia impressa, em GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 12 (Kindle, n. p.).

daqueles que ofereceram razões contra a decisão do Facebook – oposto, entretanto, não a uma ação de Estado, mas, antes, à conduta da plataforma de excluir a imagem. Como disse a Primeira-Ministra Erna Solberg à época, referindo-se à atuação do Facebook – e não do Estado – “*apagar postagens contendo imagens como essa é limitar a liberdade de expressão, a democracia e o direito de criticar e questionar, e ver os eventos passados como eles de fato foram, não como eles devem ser vistos por uma corporação gigante*”<sup>198</sup>. A crítica é clara: o Facebook, ao exercer poder sobre aquele espaço e banir a imagem, violou o direito à liberdade de expressão das pessoas, aqui invocado como uma salvaguarda não contra a ação do Estado, mas contra a ação da empresa, que, por trás de uma plataforma da internet, decidiu que aquela manifestação não seria admissível no ambiente que administra.

No segundo caso, o banimento de Alex Jones e de outras figuras do conservadorismo também ocasionou uma controvérsia centrada no papel regulador da liberdade de expressão que passam a ter essas empresas. Assim, excluídos os perfis, surgiu uma discussão que opôs, de um lado, quem visse nisso uma ação correta das plataformas, inclusive porque as consideravam como responsáveis pela disseminação desse tipo de conteúdo; e, de outro, quem entendeu a ação como uma forma de restringir a liberdade de expressão. Nesse último caso, em específico, fica evidente a forma como a gramática do direito à liberdade de expressão, como algo oposto a quem exerce o poder sobre um espaço de manifestação, foi transplantado para essa relação entre plataforma e usuário: para Trump, por exemplo, surgiu a necessidade, a partir do caso, de “*a monitorar a censura de CIDADÃOS AMERICANOS em plataformas de mídias sociais*”<sup>199</sup> e lutar “*pela liberdade de expressão online*”<sup>200</sup>, enfrentando não o Estado – por ele próprio representado, aliás – mas a plataforma.

No terceiro caso, também não foi só a manifestação que fez o caminho da praça à plataforma – ela foi acompanhada, como reconstruí acima, de uma controvérsia conhecida da regulação da liberdade de expressão, quanto à permissibilidade de

---

<sup>198</sup> “*To erase posts containing such images is to limit freedom of expression, democracy, the right to criticize and question, and view past events as they actually were, not as they have been deemed to be by one giant corporation*”. SOLBERG, E. Facebook had no right to edit history. **The Guardian**, 9.9.2016, disponível em <<https://www.theguardian.com/commentisfree/2016/sep/09/facebook-napalm-vietnamese-deleted-norway>>, acesso em 21.6.2019.

<sup>199</sup> Cf. *supra*, nota n.º 149.

<sup>200</sup> Cf. *supra*, nota n.º 152.

manifestações que negam o Holocausto. É interessante perceber, nesse sentido, que os mesmos atores que participam dos debates fora da plataforma – via de regra, associações vinculadas ao sionismo – atuaram na controvérsia, com a diferença que se antes as pressões eram endereçadas ao Estado – para que, por exemplo, estabelecesse legislações restritivas quanto a isso –, elas agora são direcionadas às plataformas, para que ou alterem suas políticas internas, ou as interpretem, de modo a banir desses ambientes pessoas que negam o Holocausto e suas manifestações. Assim, os sobreviventes do Holocausto enviaram uma carta “*exigindo que a negativa do Holocausto não seja mais permitida como parte da política de ‘liberdade de expressão’ do Facebook*”<sup>201</sup> – isto é, tanto trazendo essa controvérsia à plataforma, quanto a colocando em termos do direito à liberdade de expressão, oposto, novamente, não ao Estado, mas à plataforma, entendida como um agente que, mesmo sendo privado, é capaz de exercer poder sobre um espaço de manifestação relevante e, nesse sentido, decidir o que permanece ali ou não.

Portanto, o movimento da praça à plataforma não levou apenas as manifestações das pessoas do público ao privado. Ele trouxe, junto, e para as plataformas, as controvérsias quanto àquilo que se pode dizer<sup>202</sup> e, por consequência, a gramática do direito à liberdade de expressão, que, da mesma forma que oposto ao Estado que restringe a manifestação da praça, passa a ser oposto à empresa que restringe a manifestação na plataforma. O que muda não é apenas o local em que as pessoas se manifestam; muda, também, a dinâmica decisória sobre o que as pessoas podem dizer, que passa a refletir o fato de que o Estado não mais monopoliza o poder sobre os espaços de manifestação mas, antes, cede lugar às estruturas corporativas que, por trás das plataformas, fazem em seus espaços a sua própria lei. Assim, a existência de algo que se pode chamar de regulação da liberdade de expressão parece depender menos da atuação de uma entidade que se possa chamar de “*Estado*”, e mais do fato de que um determinado ator – estatal ou não – seja capaz de exercer poder sobre um espaço relevante de manifestação e, nesse sentido, restringir ou permitir aquilo que as pessoas podem dizer ali. A era digital faz emergir, nesse sentido, e mesmo que sob bases distintas – não a Primeira Emenda, mas políticas

---

<sup>201</sup> Cf. *supra*, nota n.º 159.

<sup>202</sup> Cf. *infra*, notas n.º 388, 389 e 390, em que exponho, de forma exemplificativa, paralelos entre controvérsias tratadas pela jurisprudência dos Estados Unidos quanto à liberdade de expressão e problemas surgidos na moderação de conteúdo quanto à definição do que é uma ameaça, às limitações que os direitos autorais estabelecem para as manifestações das pessoas e àquilo que se deve entender por uma incitação.



internas empresariais de conteúdo que a replicam –, uma forma de regulação da liberdade de expressão entre dois de seus novos atores: as plataformas e os usuários.

Em conclusão, reunidas, esta seção e a anterior indicam que a mudança social que Balkin encapsula como uma transição de infraestruturas de liberdade de expressão gerou cenários mais complexos do que aquele que ele se propõe a analisar. Que as mudanças tecnológicas que proporcionaram uma era digital ocasionaram uma revolução na infraestrutura da liberdade de expressão, de modo que o que antes era dito nas praças ou nos jornais, passou a ser expresso por meio de plataformas na internet, não há dúvidas: os próprios casos acima fornecem exemplos disso. O problema é ver como resultado dessa revolução apenas a criação de um novo campo de atuação do mesmo ator de sempre – o Estado – na regulação da liberdade de expressão. Mais do que o surgimento de novos *espaços de expressão*, os casos acima sugerem que essa revolução também significou o surgimento de novos *espaços de decisão*, em que os regimes jurídicos estatais passam a atuar em paralelo com políticas de conteúdo criados pelos atores privados que, ao exercer poder sobre a infraestrutura que detêm, transformam-se não em novos instrumentos do Estado, mas em novos atores decisórios da liberdade de expressão, regulando-a tanto na medida em que decidem aquilo que as pessoas podem dizer, quanto na que enfrentam, com isso, e contra si, a invocação do direito à liberdade de expressão.

É por isso que um recorte que restringe a regulação da liberdade de expressão – e os estudos a ela ligados –, direta ou indiretamente, à ação estatal é problemático, inobstante a defesa de Balkin<sup>203</sup>: ele retira do seu próprio alcance um conjunto de situações em que questões centrais quanto à liberdade de expressão são discutidas e decididas, com relativa independência, em nichos de poder privados criados justamente pela transição entre infraestruturas detectada por Balkin. Essa perspectiva erra, nesse sentido, ao tomar uma das formas de se proteger a liberdade de expressão – a Primeira Emenda – com a própria liberdade de expressão, um conceito mais geral<sup>204</sup>, e deixa, assim, escapar aquilo de verdadeiramente revolucionário que a era digital trouxe à regulação da liberdade de expressão: o surgimento de novos *espaços de decisão* cuja independência demonstra tanto a insuficiência de análises como a do Balkin de 2014 –

---

<sup>203</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, nota n.º 11.

<sup>204</sup> KLONICK, K. You'll Never Guess This One Crazy Thing Governs Online Speech. **Slate**, 24.8.2016, disponível em <<https://slate.com/technology/2016/08/free-speech-is-the-wrong-way-to-think-about-twitter-and-facebook.html>>, acesso em 19.6.2019.

restrita, como reconstruí acima, à ação estatal e ao seu impacto nos agentes privados – quanto a necessidade de complementá-la – levando a sério a moderação de conteúdo como uma forma, mesmo que incompatível com o “*State Action Requirement*”, de regular a liberdade de expressão.

### 3.3.5 A solução de Balkin

Como expus acima, o Balkin de 2014, que escreveu *Old-School/New-School Speech Regulation* tem um problema: ele elabora um quadro incapaz de abranger a complexidade que a era digital trouxe para a regulação da liberdade de expressão, porque restringe o conceito de regulação à atuação do Estado. Por isso, argumentei a necessidade de se caminhar na direção de uma análise diferente, que amplia a ideia de regulação para incluir também a relação de empresas que administram plataformas de internet e o conteúdo publicado por seus usuários. Nesta seção, concluo essa análise a partir de uma referência de um texto posterior – *Free Speech is a Triangle*<sup>205</sup> –, em que Balkin muda a sua posição inicial, que reconstruí acima, fazendo-o exatamente no sentido de abranger a atuação regulatória quanto à liberdade de expressão realizada por agentes privados como as plataformas da internet. Explico, nesse sentido, essa mudança de posição – que poderia parecer arbitrária, à primeira vista – com a abordagem que fiz acima, posicionando, nesse novo quadro proposto por Balkin, a linha que adoto neste trabalho.

Balkin argumenta – agora no texto de 2018 – que a regulação da liberdade de expressão transitou de um modelo diádico para um modelo pluralista. Um sistema de regulação da liberdade de expressão diádico é aquele em que “*há essencialmente dois atores: o Estado-nação de um lado, e aquele que se expressa de outro*”<sup>206</sup> – e que, nesse sentido, parece compatível com aquele ao qual ele fez referência no texto de 2014. Por sua vez, um sistema de regulação da liberdade de expressão pluralista é aquele em que atuam diversos outros atores, em específico, os agentes privados, que também agem como reguladores, tais como o Estado, em relação àquilo que as pessoas dizem.

---

<sup>205</sup> BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, pp. 2.011/2.056, 2018.

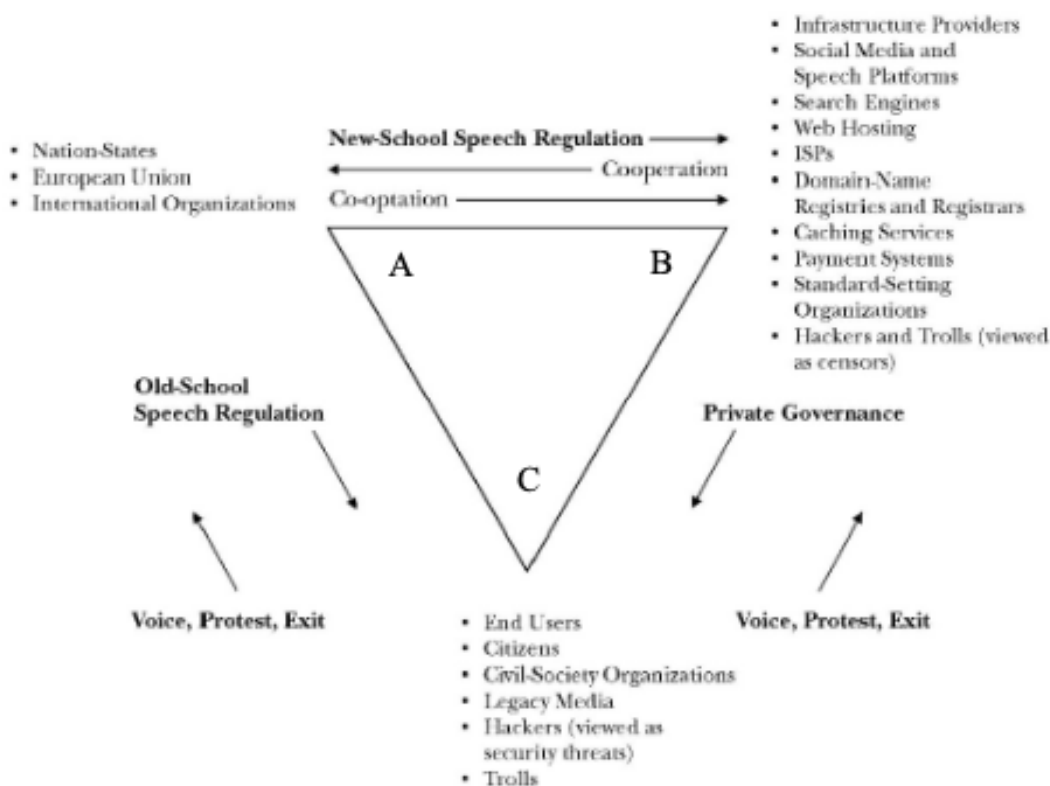
<sup>206</sup> “*In the dualist model, there are essentially two players: the nation-state on the one hand and the speaker on the other*”. BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, 2018, p. 2.013; *cf.*, no mesmo sentido BALKIN, J. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **UC Davis Law Review**, v. 51, 2018, p. 1.153.

O que faz Balkin, nesse sentido, é restringir a transição “*old-school*” / “*new-school*” para a atuação do Estado – direcionada, respectivamente, para o indivíduo que se expressa ou, então, para os intermediários da era digital, como as plataformas (eixos A-B e A-C, na figura abaixo) – e mobilizar um novo conceito, o de governança privada<sup>207</sup>, para abranger, dentro de uma nova ideia de regulação da liberdade de expressão – não atingida pela restrição que indiquei acima, que a igualava à atuação do Estado –, a atuação dos agentes privados diretamente endereçada às pessoas que se manifestam (eixo B-C). Nesse novo modelo, as companhias que administram os espaços das plataformas na internet deixam de ser caudatárias do Estado – isto é, instrumentos cujas políticas de conteúdo são automaticamente alinhadas às políticas de Estado, executando-as – e passam a constituir um novo vértice de regulação da liberdade de expressão – abaixo assinalado, na imagem de Balkin, com a letra B:

---

<sup>207</sup> No próximo capítulo, esclareço com mais detalhes o significado da ideia de governança privada; *cf.*, a esse respeito, o texto que inspirou a mudança de Balkin, KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.616 e ss; e *infra*, Cap. I.2, 3.2.2.

**Figura 10 – Modelo pluralista de regulação da liberdade de expressão**



**Fonte:** Balkin<sup>208</sup>

Esse quadro se contextualiza com o percurso deste capítulo na medida em que a realidade da moderação de conteúdo, nos termos em que a reconstruí acima, problematiza justamente um modelo diádico de regulação da liberdade de expressão – isto é, uma perspectiva centralizada na atuação estatal – porque expõe a sua incapacidade de abranger um conjunto de fenômenos regulatórios que emergem na era digital. Assim, compreender a regulação da liberdade de expressão na era digital – e, por consequência, a moderação de conteúdo – exige um modelo mais complexo que o do Balkin de 2014, capaz de abordar relações diferentes da estabelecida entre Estado e indivíduo. Figurativamente, isso significa que é preciso angularizar o modelo anterior, baseado em uma relação única entre ambos, acomodando-se um novo vértice – acima assinalado com a letra B –, ocupado por atores privados que, ao exercer poder sobre esses novos *espaços de*

<sup>208</sup> BALKIN, J. Free Speech is a Triangle. *Columbia Law Review*, v. 118, n. 7, 2018, p. 2.014. As letras A, B e C foram incluídas por mim, para facilitar a identificação no texto.

manifestação, fazem emergir *espaços de decisão* quanto àquilo que as pessoas – localizadas no vértice C – dizem. Nas palavras de Balkin:

*“Na era digital, indivíduos não enfrentam o modelo diádico familiar de regulação da liberdade de expressão. No modelo diádico, há dois atores centrais: o poder do Estado ameaça o direito do indivíduo de falar. Ao contrário disso, a era digital apresenta um modelo pluralista de controle da liberdade de expressão. No modelo pluralista, os indivíduos podem ser controlados, censurados, e vigiados tanto pelo Estado-nação, quanto pelos proprietários dos muitos diferentes tipos de infraestrutura privada, que operam através de fronteiras nacionais em múltiplas jurisdições. De fato, os maiores proprietários de infraestruturas privadas são tão poderosos que nós podemos considera-los mesmo como soberanos de propósitos especiais. Eles se engajam em lutas perpétuas pelo controle das redes digitais com os Estados nacionais que, a seu turno, tentam controlar e cooptar esses atores poderosos. A habilidade prática de falar na era digital é moldada pelo resultado dessas lutas por controle e cooptação”<sup>209</sup>.*

Ocorre que se a conclusão da seção anterior, por um lado, esclarece a complexidade dos impactos que o mundo digital teve na regulação da liberdade de expressão – refletida acima, nesse novo modelo de compreensão –, por outro, dela decorre um problema. Se a regulação da liberdade de expressão no mundo digital não fica a cargo apenas dos Estados – isto é, não se restringe aos eixos A-B e A-C –, mas também é feita, em paralelo, por políticas de conteúdo construídas por empresas que administram novos espaços de expressão – representadas pelo eixo B-C –, então as análises dedicadas apenas às transformações ocorridas nas ferramentas regulatórias estatais, que, via de regra, temos disponíveis, não são capazes de abranger todos os impactos que decorreram da era digital.

---

<sup>209</sup> *“In the digital age, individuals do not face the familiar dyadic model of speech regulation. In a dyadic model, there are two central actors: the power of the state threatens the individual’s right to speak. Instead, the digital age features a pluralist model of speech control. In the pluralist model individuals may be controlled, censored, and surveilled both by the nation state and by the owners of many different kinds of private infrastructure, who operate across national borders in multiple jurisdictions. In fact, the largest owners of private infrastructure are so powerful that we might even regard them as special-purpose sovereigns. They engage in perpetual struggles for control of digital networks with nation states, who, in turn, want to control and coopt these powerful players. The practical ability to speak in the digital age is shaped by the results of these struggles for control and cooptation”.* BALKIN, J. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **UC Davis Law Review**, v. 51, 2018, pp. 1.153.

Dito de outra forma: se os atores privados dessa transformação não são só delegatários da ação estatal – mas formadores independentes de políticas de conteúdo a partir das quais decisões quanto àquilo que as pessoas podem dizer são tomadas –, então entender apenas o que os Estados têm feito não é suficiente. É preciso ir além e compreender de que maneira essas companhias têm estruturado, em seu interior, mecanismos pelos quais decisões a respeito daquilo que as pessoas podem dizer são tomadas.

Isso é especialmente relevante porque, diferente do que acontece no Estado – em que os mecanismos regulatórios têm sido objeto de atenção há um certo tempo, produzindo abordagens sistematizadas como a do Balkin de 2014 –, o surgimento das plataformas da internet e, conseqüentemente, a existência de mecanismos privados que tomam decisões quanto àquilo que pode ser ali dito é relativamente recente. Além disso, como eles funcionam no interior de empresas – e não na administração pública, usualmente vinculada a protocolos quanto à transparência em suas condutas –, sabe-se muito pouco a respeito de como, quando e com que base essas decisões são tomadas. Isso se traduz na perplexidade manifestada pelos usuários diante de seus resultados – algo que aparece, por exemplo, no primeiro caso que reconstruí acima, em uma manifestação veiculada por Egeland logo após saber que sua imagem havia sido banida:

*“A censura do Facebook é problemática por várias razões. Nós não sabemos se ela ocorreu como o resultado de um algoritmo automático, ou se um ‘Editor’ anônimo deliberadamente removeu [a imagem] (algumas pessoas – incluindo o Facebook – indicaram que as fotos não são removidas automaticamente, mas que ‘alguém’ o faz, provavelmente não um editor em si, mas alguém que checa as fotos de acordo com a legislação do Facebook e não leva em consideração outras considerações). Nós não sabemos se o Facebook acredita que a nudez é sexualizada (e obviamente não é). [...] quer dizer, o Facebook deveria ter pedido que eu me manifestasse antes que eles deletassem algo da minha linha do tempo. É claro, eu concordo que o Facebook não deve ser uma arena para a pornografia. Mas há uma diferença significativa entre a imagem de Kim Phúc e o ‘David’ de Michelangelo, de um lado, e, de outro, pornografia ou nudez especulativa.”<sup>210</sup>*

---

<sup>210</sup> EGELAND, T. (Tom Egeland). Sem título. Oslo, 26 de agosto de 2016. **Facebook**: tom.egeland. Disponível em <<https://www.facebook.com/tom.egeland/posts/10153691636277750>>, acesso em 19.6.2019. A tradução do norueguês foi realizada com o recurso do *Google Tradutor*.

Mais: quando alguém se vê impedido de dizer algo pelo Estado, essa decisão, via de regra, está sujeita a algum mecanismo de revisão administrativa ou judicial, a ser acionada pelo interessado. Quando esse tipo de determinação vem de uma empresa, entretanto, é difícil vislumbrar, sem entender como funcionam esses mecanismos, quais são os caminhos adotados para reverter uma decisão como essa. Essa dificuldade também aparece na manifestação de Egeland:

*“Uma parte do problema com o Facebook é que não há ninguém para se chamar. Não há ninguém para se escrever. Sim, eu reclamei sobre a remoção da imagem – por um formulário completamente anônimo no ícone ‘configurações’ – e eu não recebi uma resposta. Eu não sei se alguém leu a reclamação como quando um jornal com um editor comete um erro e eu posso entrar em contato. Nesse caso, eu posso conseguir uma resposta. Essa instância não existe no Facebook. Eles não têm um editor – eles só tem as regras. A posição do Facebook como um meio de comunicação social e uma arena para o debate e reflexão é de um tipo que nós não podemos senão debater sua legislação e o governo questionável dessas regras. Talvez o Facebook mude e se desenvolva – para o melhor. Então, nós veremos se essa publicação – ou semelhantes – permanecerão de pé. Ou se uma alma sem nome do Facebook escolherá deletar tudo.”<sup>211</sup>*

Nesta dissertação, meu objetivo é enfrentar essa perplexidade. O tema deste trabalho, posicionado no eixo B-C da figura acima, é o funcionamento dos mecanismos decisórios que, localizados em empresas de tecnologia, tomam decisões quanto àquilo que as pessoas podem ou não dizer em plataformas na internet. Sob uma perspectiva ampla, minha intenção ao focar isso é fornecer uma contribuição à compreensão de como decisões a respeito da liberdade de expressão passaram a ser também tomadas, a partir de certas transformações tecnológicas, por agentes privados. Isso se justifica, por um lado, pela impossibilidade de se compreender a regulação da liberdade de expressão no mundo digital a partir de uma análise restrita às transformações na atuação do Estado; e, por outro, pela escassez de informações e de análises disponíveis quanto à forma de atuação desses agentes privados, sobretudo porque não estão vinculados – tal como o Estado – a certos protocolos de transparência quanto às decisões que tomam. Quem decidiu excluir a foto de

---

<sup>211</sup> EGELAND, T. (Tom Egeland). Sem título. Oslo, 26 de agosto de 2016. **Facebook**: tom.egeland. Disponível em <<https://www.facebook.com/tom.egeland/posts/10153691636277750>>, acesso em 19.6.2019. A tradução do norueguês foi realizada com o recurso do *Google Tradutor*.

Egeland? Qual foi o limiar que Jones ultrapassou para ser banido? De forma foi tomada a decisão de permitir que usuários de redes sociais neguem o Holocausto? Como, afinal de contas, plataformas da internet regulam a liberdade de expressão? Esta dissertação pretende ser uma resposta a essas perguntas e, nesse sentido, à circunstância de que decisões quanto àquilo que as pessoas podem ou não dizer têm sido tomadas por um mecanismo que, em grande parte, é desconhecido.

Como exponho na segunda parte deste trabalho, à parte da questão da opacidade dessa estrutura – que tratei, de forma indireta, neste capítulo, e à qual dedicarei integralmente o próximo –, trabalharei esse tema a partir de um problema específico: as transformações pelas quais passa a moderação de conteúdo, em particular quanto ao procedimento que resulta na tomada de decisão por essas empresas. Cada vez mais, as companhias responsáveis por essas decisões têm adotados procedimentos semelhantes aos dos Tribunais, instituindo mecanismos de transparência, contraditório e recorribilidade que não existiam antes, constituindo, dessa forma, uma espécie de replicação do devido processo legal em suas estruturas corporativas. É a partir da incompreensão dessa transformação, ante as limitações das ferramentas que temos disponíveis para analisá-la, que sustentarei, na segunda parte da dissertação, a necessidade de mobilizar a sociologia jurídica – e, em específico, a Teoria dos Sistemas – para melhor explicar tanto esse fenômeno decisório, quanto as transformações pelas quais ele tem passado.

#### 4. CONCLUSÃO

Mudanças tecnológicas impactam na forma como a liberdade que as pessoas têm de se expressar é regulada. Essas mudanças são encapsuladas, por Jack Balkin, como uma transição entre infraestruturas de liberdade de expressão – isto é, como uma mudança na estrutura socialmente construída e que, previamente a um resultado final comunicativo, viabiliza-o ao oferecer as condições materiais para a sua existência. Para o Balkin de 2014 – que representa uma forma comum de se abordar esse problema –, à mudança de uma infraestrutura de liberdade de expressão pré-digital para uma digital – associada à internet, em que disponibilizados, por empresas de tecnologia, espaços de manifestação para os seus usuários – corresponde uma mudança de técnicas regulatórias do Estado, que transitam de um modelo “*old-school*” – marcado, por exemplo, pela atuação direta sobre aqueles que se manifestam em sentido contrário ao de posições oficiais – para outro, chamado de “*new-*



*school*” – no qual se privilegia a ação estatal indireta, endereçada a intermediários que, nessa nova dinâmica, dominam a infraestrutura que viabiliza às pessoas se comunicarem.

O problema é que, ao restringir a análise dos impactos regulatórios da mudança tecnológica à atuação do Estado, Balkin deixou fora do seu escopo de abordagem parte das transformações que decorrem dessa transição: não apenas a criação de novos *espaços de manifestação*, essencialmente vinculados a plataformas na internet, a exemplo do Facebook, Twitter e YouTube; mas o surgimento de novos *espaços de decisão*, que, localizados no interior das estruturas corporativas das companhias que administram essas plataformas, produzem decisões que determinam aquilo que as pessoas podem ou não dizer. Nesse sentido, a reconstrução de três casos em que a veiculação de conteúdos nessas plataformas ocasionou controvérsias quanto a se essas publicações deveriam permanecer na internet demonstrou que as empresas por trás desses ambientes não apenas tomam decisões quanto àquilo que as pessoas podem dizer, como o fazem baseadas em códigos internos elaborados para esse fim. Mais: circunstâncias de cada uma dessas situações denotaram que essas decisões foram tomadas mesmo que contrariamente às posições manifestas por Estados associados às controvérsias, o que sugere que essas empresas não agem apenas como delegatárias da ação estatal mas, antes, comandam mecanismos decisórios paralelos e relativamente independentes, que regulam a liberdade de expressão – algo que justifica a transição, feita pelo Balkin de 2018, de um modelo diádico para uma perspectiva pluralista.

Por isso, não faz mais sentido entender os impactos da tecnologia na regulação da liberdade de expressão apenas sob a perspectiva de como as ferramentas dos Estados mudaram. Essa é uma abordagem insuficiente, porque deixa de lado a atuação independente de agentes privados que, como os casos expõem, instituíram regimes regulatórios em paralelo vocacionados à produção de decisões que determinam – tanto quanto os Estados – aquilo que as pessoas podem dizer em importantes – se não os mais importantes – ambientes de manifestação da atualidade: as plataformas da internet.

Entender a regulação da liberdade de expressão atualmente, portanto, traduz-se em também entender como, a partir da era digital, agentes privados passaram a estruturar mecanismos decisórios para exercer poder sobre as manifestações veiculadas por usuários nas plataformas que administram – e isso justifica a abordagem que proponho neste trabalho: compreender, a partir de um quadro sociológico robusto, como funcionam e se transformam esses mecanismos de moderação de conteúdo nas principais plataformas da internet.

O cerne deste capítulo, portanto, foi tanto apresentar o tema desta dissertação – a moderação de conteúdo nas plataformas da internet – quanto o justifica. No capítulo seguinte, meu objetivo é prosseguir nesse caminho: nele, meu foco será construir uma abordagem taxonômica da moderação de conteúdo, fornecendo conceitos e estruturando classificações que esclareçam tanto como esse mecanismo funciona, quanto quais são as interpretações que temos disponíveis a seu respeito.

## **CAPÍTULO 2 – DA OPACIDADE À CLAREZA: COMO FUNCIONAM OS MECANISMOS DE MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET**

### **1. INTRODUÇÃO**

A primeira parte deste trabalho tem dois objetivos. O primeiro, ao qual dedicado o capítulo anterior, é oferecer razões que justifiquem uma dissertação sobre a regulação da liberdade de expressão que tenha a moderação de conteúdo das plataformas da internet como tema. O segundo objetivo, ao qual dedicado este capítulo, é resolver um problema característico da moderação de conteúdo, que dificulta a compreensão de casos concretos e a elaboração de questionamentos e hipóteses mais complexos a seu respeito: a opacidade.

No capítulo anterior, o primeiro objetivo foi abordado inserindo a moderação de conteúdo em um quadro maior de mudanças sociais, a partir da constatação de que desenvolvimentos tecnológicos associados às telecomunicações transformaram o contexto que viabiliza às pessoas se manifestarem. Em particular, com a emergência da era digital e à medida que cada vez mais pessoas passaram a se manifestar na internet, ganharam protagonismo empresas de tecnologia como o Facebook, o Twitter e o Google, que disponibilizam e administram espaços virtuais de expressão para os seus usuários. Não foram só as manifestações, entretanto, que se deslocaram para esses novos ambientes: como os casos do capítulo anterior expõem, também foram às plataformas as divergências quanto àquilo que as pessoas podem dizer, concretizadas em controvérsias quanto à permissibilidade de publicações que contêm, por exemplo, nudez, informações falsas ou discurso de ódio. Essas questões passaram a ser submetidas às empresas que administram esses espaços, porque elas exercem poder nesses ambientes – definindo tanto a sua arquitetura, quanto, em última análise, aquilo que permanece ou não ali –, levando-as a construírem mecanismos que lhes permitem decidir, a partir de critérios internamente definidos, o que seus usuários podem dizer. Surgem, assim, mais do que novos espaços de manifestação, novos espaços de decisão, estruturas localizadas dentro de empresas de tecnologia e dedicadas a estabelecer tanto abstratamente, quanto concretamente, aquilo que as pessoas podem ou não dizer na internet.

Essa nova realidade desafia as perspectivas tradicionais da regulação da liberdade de expressão, porque é incompatível com a restrição, até então feita, que

circunscrevia a ideia de regulação a uma relação entre, de um lado, o Estado e, de outro, o particular. A emergência de novos nichos decisórios localizadas nas empresas de tecnologia – que controlam esses novos espaços de expressão da atualidade – exige um conceito diferente de regulação da liberdade de expressão, capaz de fazer frente a essa nova realidade, em que o Estado não mais monopoliza o poder nos locais em que as pessoas se manifestam. Por isso, a ideia de regulação da liberdade de expressão transita de uma perspectiva diádica – limitada à relação Estado/particular – para uma visão pluralista – aberta a uma análise das relações regulatórias estabelecidas entre particulares entre si, como a que ocorre entre usuários e plataformas. A partir desse contexto, o trabalho foi posicionado como uma dissertação que adota uma perspectiva pluralista da regulação da liberdade de expressão e que se concentra no funcionamento da moderação de conteúdo, isto é, do mecanismo que, localizado em empresas de tecnologia, produz decisões quanto àquilo que as pessoas podem dizer nas plataformas da internet.

Neste segundo capítulo, minha proposta é avançar ao segundo objetivo desta primeira parte, explicando como a moderação de conteúdo funciona. Diferente dos mecanismos regulatórios tradicionais da liberdade de expressão – como, por exemplo, a ida do Estado às Cortes para coibir a divulgação de uma informação, como nos casos *Pentagon Papers*<sup>212</sup> –, a moderação de conteúdo é uma atividade que se desenvolve no interior de empresas e que, por isso, não está sujeita aos mesmos protocolos de transparência que vinculam a administração pública. Ela é, por isso, uma forma de decidir essencialmente opaca: é difícil saber exatamente o que acontece quando uma plataforma determina, por exemplo, a retirada do conteúdo veiculado por um usuário. Mais do que uma característica, entretanto, isso é um problema, porque essa opacidade impede tanto que fenômenos como o caso Lipman e as três situações que reconstruí no capítulo anterior sejam compreendidos em toda a sua extensão, quanto que problemas e hipóteses de pesquisa – tais como as que desenvolvo no capítulo seguinte – sejam feitas com precisão. A opacidade gera, nesse sentido, uma desarticulação das informações, porque elas não são obtidas de forma consolidada – por exemplo, em um conjunto organizado de precedentes administrativos ou judiciais públicos –, mas esparsamente, por investigações jornalísticas e entrevistas com pessoas que trabalham ou trabalhavam nessas estruturas, elementos que precisam, portanto, ser estruturados antes de se prestarem a basear uma pesquisa.

---

<sup>212</sup> Cf. *supra*, nota n.º 67.

Para resolver esse problema, minha proposta é construir, a seguir, tanto conceitos que delimitem a moderação de conteúdo, quanto classificações que lancem luz à forma como as plataformas da internet tomam decisões a respeito daquilo que os seus usuários podem ou não dizer. Farei isso, inicialmente, aprofundando o tratamento casuístico que conferi à moderação de conteúdo, apreendendo-a por uma abordagem que extrai e sistematiza desses casos, da literatura disponível e de investigações jornalísticas, elementos quanto a como esse mecanismo funciona. Proponho, assim – e tal como Balkin fez em relação à regulação da liberdade de expressão nos Estados, subdividindo-a em um detalhamento que permitiu compreendê-la melhor – uma taxonomia da moderação de conteúdo das plataformas da internet.

A partir dessa taxonomia, irei sistematizar também as formas pelas quais a moderação de conteúdo é compreendida pela literatura que tem se dedicado a estudá-la, organizando esse debate em duas correntes distintas: a primeira, que a analisa como uma categoria dentro de quadros jurisdicionais específicos, como o da Primeira Emenda – de viés diádico; e a segunda, que a vê como uma forma de governança privada – de viés pluralista. Entender essas engrenagens do debate é algo importante não só porque ajuda a esclarecer a moderação de conteúdo, mas porque é nas deficiências identificadas em ambas as perspectivas que irei sustentar, na segunda parte deste trabalho, a necessidade de se abordar a moderação de conteúdo de uma maneira diferente.

Estruturalmente, inicio o capítulo reconstruindo dois textos de referência no tema, escritos por Tarleton Gillespie e Kate Klonick, em que a moderação de conteúdo é exposta a partir de entrevistas realizadas com os responsáveis por comandar esse trabalho nas empresas de tecnologia (2). Proponho, a partir deles, uma sistematização do seu funcionamento nas plataformas da internet, organizando-a, primeiro, por uma abordagem conceitual à moderação de conteúdo (2.2); e, segundo, entendendo-a tanto como uma forma de estabelecer regras abstratas de conduta (2.3), quanto como um procedimento adjudicatório (2.4). A partir disso, exponho as duas formas pelas quais a literatura enxerga a moderação de conteúdo (3) – como uma categoria dentro de contextos jurisdicionais específicos (3.1) e como uma forma de governança privada da liberdade de expressão (3.2). Encerro o capítulo (4) – e a primeira parte deste trabalho – resumindo as suas conclusões e abrindo caminho para a minha proposta na parte seguinte: problematizar a moderação de conteúdo a partir tanto da sistematização que segue, quanto das transformações pelas quais esse mecanismo tem passado recentemente.

## 2. A MODERAÇÃO DE CONTEÚDO DAS PLATAFORMAS DA INTERNET: UMA ABORDAGEM SISTEMÁTICA

### 2.1 Introdução

No capítulo anterior, a moderação de conteúdo das plataformas da internet foi exposta a partir de uma análise casuística de três situações, escolhidas com o objetivo de retesar a abordagem inicial de Balkin quanto à regulação da liberdade de expressão e, assim, evidenciar a necessidade de se percorrer um caminho diferente. Esse caminho, apresentado no capítulo anterior como o norte desta dissertação, é o de apreender a moderação de conteúdo como uma forma de regular a liberdade de expressão. Diferente, entretanto, dos mecanismos estatais, que estão submetidos a protocolos de transparência, a moderação de conteúdo é um modelo decisório essencialmente opaco. Isso é um problema em si, porque dificulta a compreensão de casos como os do capítulo anterior; e torna difícil construir problematizações e hipóteses mais complexas, como a que quero desenvolver na segunda parte da dissertação.

Neste capítulo, minha proposta é oferecer uma sistematização que resolva, ao menos parcialmente, o problema da opacidade que é característica da moderação de conteúdo. Meu objetivo, nesse sentido, é construir um cenário abstrato – e não casuístico, como no capítulo anterior – de como esse mecanismo funciona. Assim, e como fez Balkin quanto à regulação da liberdade de expressão nos Estados, dedico as próximas seções a tanto construir uma base conceitual, que delimite a moderação de conteúdo em suas diferentes vertentes; quanto a sistematizá-la em um quadro que posicione as diferentes formas pelas quais essas decisões são tomadas, esclarecendo-as.

Estruturalmente, organizo esta parte do trabalho em três etapas. Na primeira (2.2), ofereço dois conceitos fundamentais para os meus objetivos neste capítulo, delimitando-os a partir de definições quanto a (i) o que é uma plataforma; e (ii) o que é moderação de conteúdo. A partir disso, detalharei esse último ponto, explicando, primeiro (2.3), como a moderação de conteúdo funciona constituindo normas gerais de comportamento de usuários nas plataformas da internet; e, segundo (2.4), como a moderação de conteúdo funciona decidindo, concretamente, em que medida manifestações realizadas por determinados usuários devem ou não ser retiradas da plataforma.

### 2.2 Da promessa de democratização à moderação de conteúdo: dois conceitos

Meu objetivo, neste capítulo, é oferecer uma abordagem da moderação de conteúdo que a apreenda sob uma perspectiva abstrata, isto é, como uma estrutura decisória, deixando de lado, ao menos momentaneamente, a sua manifestação casuística em situações específicas como as que reconstruí no primeiro capítulo. Para tanto, o passo mais inicial possível é o de fornecer uma definição quanto ao que se deve entender por moderação de conteúdo nas plataformas da internet, delimitando, assim, o objeto dessa abordagem. Faço isso nesta seção, em dois passos: primeiro, apresento um conceito que define o que são plataformas na internet; para, em seguida, conceituar a moderação de conteúdo que ocorre nesses ambientes.

O que são plataformas de internet? Esse não é um termo estranho a este trabalho. No primeiro capítulo, ele foi utilizado para designar uma das vertentes da distinção, feita por Balkin, entre infraestruturas pré-digital e digital da liberdade de expressão. Em síntese, Balkin analisa os impactos da era digital – um conjunto de transformações tecnológicas associadas, fundamentalmente, às telecomunicações – para a regulação da liberdade de expressão compreendendo como da mudança social fez-se a mudança regulatória. Ele encapsula, para isso, a mudança social no conceito de infraestrutura da liberdade de expressão, que designa a estrutura social construída e que, previamente ao resultado final comunicativo, viabiliza-o ao oferecer as condições materiais e institucionais para a sua existência. A mudança social realizada pela era digital, nesse sentido, é traduzida como uma mudança de infraestruturas de liberdade de expressão, que transitam de um modelo pré-digital para um digital, marcado – dentre outras características<sup>213</sup> – pelo movimento que Balkin detecta dos “editores” para as “plataformas”<sup>214</sup>.

Nesse sentido particular, plataformas – como o Facebook – são instâncias construídas por Balkin como opostas aos “editores” – como o *The New York Times* –, porque constroem espaços destinados não à disponibilização de manifestações próprias, mas que se prestam a viabilizar que outras pessoas se manifestem – os seus usuários. Assim, e, respectivamente:

---

<sup>213</sup> Cf. *supra*, Quadro n.º 1.

<sup>214</sup> Cf. *supra*, nota n.º 81.

*“Na metade do século vinte, as companhias de mídia mais poderosas eram editoras que distribuíam conteúdo que eles criavam ou editavam: companhias de filmes, editoras de livros, jornais e radiodifusoras [...]”.*

*“Na segunda década do século vinte e um, as companhias de mídia mais poderosas são plataformas como o Google e o Facebook. Essas plataformas não são construídas primariamente para publicar o que o proprietário da plataforma cria. Em sentido contrário, elas criam oportunidades para usuários finais publicarem (Blogger, Tumblr, Twitter), mandarem e-mails e mensagens privadas (Gmail, Yahoo! Mail), fazer upload de conteúdos (YouTube e Pintarest) e compartilhar conteúdo (Facebook); e eles facilitam para usuários finais encontrar o conteúdo criado por outros (Google e Bing)”<sup>215</sup>.*

Mesmo que a ideia trazida por Balkin expresse o essencial – ou seja, que plataformas da internet são serviços viabilizados por empresas que disponibilizam espaços para as manifestações das pessoas na internet, intermediando-as – ela não é suficiente. Por um lado, como o autor aborda apenas incidentalmente as plataformas, ele não explica o que está por trás da existência desses novos intermediários – algo importante de ser compreendido, porque a promessa por trás da internet era exatamente a de construir uma comunidade mundial em que não existiriam intermediários para que as pessoas se manifestassem<sup>216</sup>. Por outro, Balkin une, nos exemplos que fornece, sob um mesmo conceito, serviços bem diferentes e incompatíveis com a sua definição. Assim, ele menciona no excerto acima tanto páginas como o Blogger, Tumblr, Twitter, Facebook e Pintarest – espaços na internet que se amoldam à sua definição, porque, de fato, disponibilizam espaços de manifestação para os seus usuários –, quanto serviços que realizam atividades bem diferentes disso, como é o caso de mecanismos de envio de mensagens de e-mail – o Gmail – e de busca na internet – o Google.

---

<sup>215</sup> *“In the middle of the twentieth century, the most powerful media companies were publishers who distributed content that they created or edited: motion picture companies, book publishers, newspapers, and broadcasters [...]”; “In the second decade of the twenty-first century, the most powerful media companies are platforms like Google and Facebook. These platforms are not primarily designed to publish what the platform owner creates. Instead they create opportunities to end users publish (Blogger, Tumblr, Twitter), send email and private messages (Gmail, Yahoo! Mail) upload content (YouTube, Pintarest) and share content (Facebook); and they made it easy for end users to find content created by others (Google, Bing)”.* BALKIN, J. Old-School/New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.299.

<sup>216</sup> Cf. nesta mesma seção, a reconstrução quanto a essa promessa e suas críticas a partir de um texto publicado à época por David Beer e Roger Burrows; *infra*, nota n.º 222.



Para precisar melhor esse conceito, especificando-o para uma abordagem que mobilize a ideia de plataforma em função de um trabalho sobre a moderação de conteúdo, Gillespie<sup>217</sup> afirma que, para entender *o que* são as plataformas, é preciso compreender *de onde as plataformas vêm* – e, em particular, *como* elas se relacionam com a promessa inicial de que a internet seria um espaço de livre manifestação para as pessoas.

Sob essa perspectiva, a internet nasceu, ao menos retoricamente, como um instrumento capaz de viabilizar às pessoas um espaço para se manifestarem, independentemente do crivo editorial daqueles que controlavam os meios de comunicação. Ela surge em um contexto no qual a difusão de informações em larga escala exigia, daquele que queria se manifestar, o acesso a estruturas materiais e humanas detidas por conglomerados de mídia – a exemplo de jornais, rádios e empresas de televisão. Esses agentes exerciam, por isso, poder sobre aquilo que veiculavam, agindo como “*guardiões*” (“*gatekeepers*”<sup>218</sup>) dessa infraestrutura e distribuindo, de forma desigual, a oportunidade e a dimensão da expressão das pessoas. Assim, e como, por exemplo, os casos *Pentagon Papers* expõem<sup>219</sup>, não importa o quanto a informação que alguém tem é importante, ela ainda precisava passar pelo crivo editorial das empresas que detinham a estrutura midiática – no caso, o *The New York Times* e o *Washington Post* – para vir a público.

Nesse sentido, se a internet foi concebida, desde o início, como um “*espaço colaborativo em que você pode se comunicar por meio do compartilhamento das informações*”<sup>220</sup>, ela teve um primeiro momento – atualmente designado sob o nome de

---

<sup>217</sup> GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 22 e ss (Kindle, n. p.).

<sup>218</sup> “*Access to the public would no longer be mediated by the publishers and broadcasters that played such powerful gatekeeper roles in the previous century*” (“*O acesso ao público não seria mais mediado pelos editores e empresas de transmissão que tiveram papéis tão poderosos de guardiões no século passado*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 23 (Kindle, n. p.). *Cf.*, a esse respeito, a referência que Zuckerberg faz quanto ao Facebook, posicionando-o como uma plataforma que supera esses “*gatekeepers*” em ZUCKERBERG, M. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>; *supra*, nota n.º 39 e *infra*, nota n.º 508.

<sup>219</sup> *Cf. supra*, nota n.º 67; em ambos os casos envolvendo os *Pentagon Papers* os documentos vazados vieram a público por veículos da mídia tradicional, os jornais *The Washington Post* (United States v. The Washington Post Company) e *The New York Times* (New York Times Company v. United States).

<sup>220</sup> A afirmação foi feita por Tim Berners-Lee, um dos fundadores da internet, em uma entrevista à BBC. “*The original idea of the web was that it should be a collaborative space where you can communicate*

Web 1.0 –, em que as conexões que viabilizava eram utilizadas para que determinados atores – que detinham os meios técnicos para tanto – disponibilizassem, aos indivíduos em geral, conteúdos por ele produzidos. Ela já funcionava, nesse momento inicial, como um instrumento de conexão entre as pessoas, mas sem modificar a lógica da produção de conteúdo, que continuava concentrada nos detentores da estrutura que era necessária para viabilizar a difusão da informação. Persistia uma firme linha entre a criação e o consumo de conteúdo, que gerava uma assimetria de papéis na produção da informação disponibilizada na internet semelhante à que se tinha em qualquer outra mídia<sup>221</sup>.

Essa realidade muda à medida que a Web 1.0 transita para a Web 2.0. David Beer e Roger Burrows<sup>222</sup> analisaram esse movimento sob a perspectiva sociológica, caracterizando-o como uma segunda geração de serviços baseados na estrutura na internet, essencialmente vinculada ao empoderamento do usuário comum, que passa a produzir, ele mesmo, as informações que são disponibilizadas na rede. Surge a Wikipedia<sup>223</sup>, uma enciclopédia na internet construída a partir da colaboração mútua de seus usuários, que criam e editam verbetes em uma estrutura participativa de produção do conhecimento; o Facebook, uma rede social constituída a partir das informações e manifestações de seus

---

*through sharing information*”. BBC. Web’s inventor gets a knighthood. **BBC**, 31.12.2003, disponível em <<http://news.bbc.co.uk/2/hi/technology/3357073.stm>>, acesso em 22.7.2019, *apud* ORLY, L. The Law of the Platform. **Minnesota Law Review**, San Diego Legal Studies Paper n.º 16-212, 2016, p. 96, nota n.º 24.

<sup>221</sup> “*The companies, services and technologies that survived, they argued, all had certain characteristics in common; they were collaborative in nature, interactive, dynamic, and the line between the creation and consumption of content in these environments was blurred*” (“*As companhias, serviços e tecnologias que sobreviveram, argumenta-se, todas têm certas características em comum; elas são por natureza colaborativas, interativas, dinâmicas e a linha entre a criação e o consume de conteúdo nesses ambientes é borrada*”). MANESS, J. M. Library 2.0 Theory: Web 2.0 and its implications for Libraries. **Webology**, v. 3, n. 2, 2006, n. p., *apud* BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>222</sup> BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>223</sup> “*Wikipedia [...] is a multilingual, web-based, free-content encyclopedia project supported by the Wikimedia Foundation and based on a model of openly editable content*” (“*A Wikipedia é um projeto de uma enciclopédia multi-linguística, baseada na web, de livre conteúdo apoiado pela Fundação Wikimedia e baseada em um modelo de conteúdos abertamente editáveis*”). WIKIPEDIA. **About**. Disponível em <<https://en.wikipedia.org/wiki/Wikipedia:About>>, acesso em 23.7.2019.

usuários, que criam perfis em uma espécie de comunidade virtual<sup>224</sup>; o Blogger<sup>225</sup>, ferramenta que viabiliza às pessoas produzirem as próprias páginas na internet, disponibilizando aquilo que pensam; e, um pouco mais tarde, serviços de nichos específicos, como os de compartilhamento de imagens (Tumblr<sup>226</sup>, Flickr<sup>227</sup>), vídeos (YouTube<sup>228</sup>) e pequenos textos (Twitter<sup>229</sup>).

O quadro abaixo, retirado do texto de Beer e Burrows, sintetiza esse movimento:

**Quadro 5 – “Algumas diferenças esquemáticas entre a Web 1.0 e a Web 2.0”**

<b>Dimensões da diferença</b>	<b>Web 1.0 (1993-2003)</b>	<b>Web 2.0 (2004 – além)</b>
<b>Modo...</b>	Ler	Escrever e contribuir
<b>Unidade primária de conteúdo...</b>	Página	Postagem / registro
<b>Estado...</b>	Estático	Dinâmico
<b>Visto por meio de...</b>	Navegador	Qualquer coisa
<b>Conteúdo criado por...</b>	Programador da Web	Todos
<b>Domínio dos...</b>	Web Designer e Nerds	Uma nova cultura de pesquisa pública?

<sup>224</sup> FACEBOOK. **Termos de Serviço**. Disponível em <<https://www.facebook.com/legal/terms>>, acesso em 23.7.2019.

<sup>225</sup> “O Blogger é um serviço gratuito destinado à comunicação, livre manifestação e liberdade de expressão”. BLOGGER. **Política de Conteúdo**. Disponível em <<https://www.blogger.com/content.g?hl=pt&bpli=1>>, acesso em 23.7.2019.

<sup>226</sup> “Nós deixamos a coisa muito, muito fácil para que as pessoas possam criar um blog e publicar o que elas bem entenderem”. TUMBLR. **O que é o Tumblr?** Disponível em <[https://www.tumblr.com/login?language=pt\\_BR](https://www.tumblr.com/login?language=pt_BR)>, acesso em 23.7.2019. Atualmente é possível publicar textos, vídeos e outros formatos no Tumblr, que se aproximou a um serviço de *blogging*.

<sup>227</sup> “Queremos ajudar as pessoas a disponibilizar suas fotos para as pessoas que são importantes para elas”. FLICKR. **Sobre o Flickr**. Disponível em <<https://www.flickr.com/about>>, acesso em 23.7.2019.

<sup>228</sup> “Nossa missão é dar a todos uma voz e revelar o mundo”. YOUTUBE. **Sobre**. Disponível em <<https://www.youtube.com/intl/pt-BR/yt/about/>>, acesso em 23.7.2019.

<sup>229</sup> “O Twitter é o lugar certo para saber mais sobre o que está acontecendo no mundo e sobre o que as pessoas estão falando agora”. TWITTER. **Sobre**. Disponível em <<https://about.twitter.com/pt.html>>, acesso em 23.7.2019.

O texto de Beer e Burrows, que é de 2007 – e é uma das primeiras abordagens à Web 2.0, porque contemporânea ao seu surgimento –, propõe três caminhos de pesquisa sociológica possível quanto a essa nova utilização da internet, sustentados naquilo de novo que esse momento traz. Primeiro, a circunstância de que “*a linha entre a criação e o consumo de conteúdo nesses ambientes é obscurecida*”<sup>231</sup>, de forma que “*os usuários são envolvidos no processo de produção e consumo, na medida em que eles geram e navegam pelo conteúdo online, na medida em que marcam [tag], escrevem em blogs, postam e compartilham*”<sup>232</sup>. Segundo – e aqui antecipando os problemas de privacidades presentes nessa nova realidade –, o fato de que, com a Web 2.0, “*o ato de produção, particularmente no caso de SNS [Social Networking Sites] cria uma situação em que informações privadas se tornam abertas e acessíveis a qualquer pessoa com acesso à internet*”<sup>233</sup>. Terceiro – ponto de maior interesse aqui – o surgimento de uma nova retórica de democratização, segundo a qual a Web 2.0 se traduziria em uma “*nova, colaborativa, participativa e abertura cultura, em que qualquer pessoa pode se envolver e todos têm o potencial de serem vistos ou ouvidos*”<sup>234</sup>. Novamente se antecipando aos

---

<sup>230</sup> BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p.. O quadro foi traduzido de forma livre.

<sup>231</sup> “[...] *‘the line between the creation and consumption of content in these environments was blurred’*”. MANESS, J. M. Library 2.0 Theory: Web 2.0 and its implications for Libraries. **Webology**, v. 3, n. 2, 2006, n. p., *apud* BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>232</sup> “[...] *users are involved in processes of production and consumption as they generate and browse online content, as they tag and blog, post and share*”. BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12 n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>233</sup> “*The act of production, particularly in the case of SNS, creates a situation in which private information becomes open and accessible to anyone with access to the Internet*”. BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>234</sup> “*This, we are led to believe, has led to a new collaborative, participatory or open culture, where anyone can get involved, and everyone has the potential to be seen or heard*”. BEER, D., BURROWS, R. Sociology

debates que viriam dali em diante, o texto não absorve acriticamente esse novo discurso, mas antes recomenda que ele seja adotado com cautela, porque, já naquela época – em que o Facebook tinha 18 milhões de usuários, menos de um centésimo do número atual – ficava clara a “*formação de novas hierarquias e divisões sociais; de poder das novas indústrias culturais; de problemas e subversões geradas pela cultura colaborativa; de padrões de participação social; da criação de novas ‘pessoas na moda’; de operações de novas estratégias de marketing viral; e de quem é ouvido acima da algazarra que é a Web 2.0*”<sup>235</sup>.

Esse terceiro caminho é importante, porque ele revela que, mesmo no início da Web 2.0, já eram perceptíveis as limitações em se cumprir a promessa de que a internet seria uma plataforma sem regulações e, principalmente, intermediários, capaz de derrubar as barreiras impostas pelas empresas tradicionais de mídia e eliminando os “*gatekeepers*”. Isso porque, à medida que as pessoas foram sendo inseridas nessa nova realidade – sob a classe de “*usuários*” –, ficou cada vez mais claro que não bastava um computador e um acesso à internet para que pudessem se expressar. Antes, era necessário estabelecer uma interface tecnológica que organizasse isso, isto é, um serviço que, prestado por uma empresa a os usuários, disponibilizasse as ferramentas pelas quais algo pudesse ser veiculado nesse espaço. O Facebook, inicialmente uma plataforma que reunia estudantes, é um bom exemplo disso: seu sucesso está baseado na criação de uma interface tecnológica em que as pessoas criavam perfis de si próprias em um ambiente virtual e ali organizavam comunidades de interesse comum, manifestando-se a partir de uma página intuitiva que, criada pela companhia, mediava a internet enquanto um conjunto de códigos complexos de computador, acessíveis apenas a quem detinha um conhecimento específico, e o usuário, que desejava se manifestar naquele ambiente mesmo sem ter conhecimento técnico. A transição entre um momento em que o conteúdo era criado por um “*Programador da Web*” para outro em que ele era produzido por “*Todos*” exigiu serviços que dispensassem as

---

and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>235</sup> “[...] *the formation of new hierarchies and social divisions; the power of the new culture industries; the problems and subversions afforded by the collaborative culture; patterns of social participation; the creation of new ‘in-crowds’; the operations of new viral marketing strategies; and, who it is that gets heard above the din that is Web 2.0*”. BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

pessoas de serem, elas mesmas, programadores de computador – e é essa a demanda que fez surgir as plataformas da internet. Elas aparecem, essencialmente, como “*serviços criados para ‘resolver’ alguns dos desafios de se navegar na internet aberta*”, em específico “*simplificando as ferramentas necessárias para postar, distribuir, compartilhar e comentar; elas conectavam os usuários a uma audiência maior, até mesmo global*”.

Dessa forma, “*apesar da retórica da democratização, a Web 2.0 é um mercado comercial e regulamentado*”<sup>236</sup>, constituída por estruturas criadas por esses novos intermediários que, ao substituírem os anteriores, também estabeleceram as suas próprias condições para fazê-lo. Não deixamos de lado os *gatekeepers* de antes: no limite, a Web 2.0 promoveu uma troca de guarda, pela qual plataformas passam a administrar a disponibilização e a veiculação de informações nesse novo ambiente, mitigando – para dizer o mínimo – os resultados da democratização inicialmente prometida. Elas são os novos intermediários, de um novo espaço, construído sobre uma nova internet:

*“Em outras palavras, para nos vermos livres de intermediários, nós aceitamos novos intermediários. As plataformas respondem à questão de distribuição diferentemente da internet, em seu início, ou da mídia tradicional. Mas eles oferecem o mesmo acordo básico: nós vamos administrar a distribuição para você – mas termos e condições serão aplicáveis. Esses termos podem ser menores e menos impositivos, mas você pode ser chamado a fazer mais do trabalho de postar, remover, manter, marcar [tagging], e assim em diante. Mas a plataforma continua a agir como um provedor”*<sup>237</sup>.

A partir dessa contextualização e da síntese proporcionada por Gillespie, é possível definir o que é uma plataforma. Plataformas da internet são serviços e páginas na internet que (i) “*hospedam, organizam e circulam, para os seus usuários, o conteúdo que eles compartilham e as interações que fazem*”; (ii) “*sem ter produzido ou comissionado (o*

---

<sup>236</sup> “[...] *despite the rhetoric of 'democratisation' Web 2.0 is a commercial and lightly regulated market*”. BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>>, n. p..

<sup>237</sup> “*In other words, to be free of intermediaries, we accepted new intermediaries. Platforms answer the question of distribution differently from the early web or traditional media. But they do offer the same basic deal: we'll handle distribution for you—but terms and conditions will apply. These terms may be fewer and less imposing, though you may be asked to do more of the labor of posting, removing, maintaining, tagging, and so on. But the platform still acts as a provider*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 24 (Kindle, n. p.).

*grosso) desse conteúdo*”; (iii) “[fazendo isso] *com base em uma infraestrutura que [...] realiza o processamento de dados para prestar serviços ao usuário, fazer anúncios e lucrar*”<sup>238</sup>.

Esse é um conceito preciso porque capta o que é essencial nessas estruturas, produzindo uma definição que distingue, para fins de um trabalho sobre a moderação de conteúdo, as plataformas de outros serviços que, por operar na internet – e, por vezes, até junto a essas plataformas – às vezes são com elas confundidos, como fez Balkin. É importante compreender, nesse sentido, a especificidade da definição de Gillespie: ela traduz o papel desempenhado pelas plataformas, primeiro, como viabilizadoras da circulação da informação de seus usuários na internet (item *(i)*); segundo, como partes de uma infraestrutura digital da liberdade de expressão, uma vez que mobilizam o espaço que criam em favor das manifestações de outras pessoas – e não da própria empresa, como na mídia tradicional (item *(ii)*); e, terceiro, como um serviço que não é disponibilizado gratuitamente, mas exercido por uma empresa que busca o lucro e, por isso, tem seu preço – mesmo que ele não seja tão evidente.

Mas o que exatamente está – e não está – incluído nessa definição? Estão incluídos, de uma maneira direta, os serviços que utilizei como exemplo na análise casuística que fiz no capítulo passado, isto é, o Facebook, o Twitter e o YouTube, que são, ao menos no Brasil e nos Estados Unidos<sup>239</sup>, as plataformas mais populares e familiares e que, por isso, continuarão a ser utilizadas como paradigmas a partir dos quais conduzirei

---

<sup>238</sup> “*For my purposes, platforms are: online sites and services that a) host, organize, and circulate users’ shared content or social interactions for them, b) without having produced or commissioned (the bulk of) that content, c) built on an infrastructure, beneath that circulation of information, for processing data for customer service, advertising, and profit*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 26/27 (Kindle, n. p.).

<sup>239</sup> “*There was not much choice about it, but over the last fifteen years, three American companies — YouTube, Facebook, and Twitter — have established themselves as dominant platforms in global content sharing and online speech.*” (“*Não há muita escolha quanto a isso, mas nos últimos quinze anos três companhias americanas – YouTube, Facebook e Twitter – estabeleceram-se como plataformas globais dominantes no compartilhamento de conteúdo e de manifestações online*”). KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.617. Esse é um recorte que se justifica, como afirma Klonick, pelo papel que essas plataformas tem tanto no compartilhamento quanto na moderação de conteúdo; outros, entretanto, são possíveis, a depender, por exemplo, do enfoque de pesquisa – uma análise da moderação de conteúdo em redes sociais voltadas ao desenvolvimento profissional certamente incluiria o LinkedIn, por exemplo.

este trabalho. Em todos esses casos, o que se tem é uma interface que hospeda, organiza e faz circular o conteúdo de usuários, viabilizando que eles se manifestem na Web 2.0. Gillespie traz mais alguns exemplos de serviços que são “*diretamente*” plataformas:

*“Aqui há uma lista representativa mas não exaustiva das plataformas de redes sociais que eu tenho em mente, e que serão centrais para as minha preocupação neste livro: sites de redes sociais como Facebook, LinkedIn, Google+, Hi5, Ning, NextDoor e Foursquare; provedores para blogging e microblogging como Twitter, Tumblr, Blogger, Wordpress e Livejournal; sites de compartilhamento de fotos e imagens como Instagram, Flickr, Pintarest, Photobucket, DevianArt e Snapchat; sites de compartilhamento de vídeos como YouTube, Vimeo e Dailymotion; ferramentas de discussão, opinião e fofoca como Reddit, Digg, Secret e Whisper; aplicativos de namoro e para saídas como OK Cupid, Tinder e Grindr; ferramentas de conhecimento colaborativo como Wikipedia, Ask e Quora; aplicativos de lojas como iTunes e Google Play; aplicativos de transmissão ao vivo como Facebook Live e Periscope”<sup>240</sup>.*

Neste trabalho, irei me concentrar nesse núcleo duro apresentado por Gillespie, acessando-o, pelos exemplos mais representativos em termos de número de usuários – notadamente o Facebook, que é a maior rede social do mundo, com 2,23 bilhões de utilizadores. É importante mencionar, entretanto, que a moderação de conteúdo não é uma atividade necessariamente restrita a essas plataformas, mas também se apresenta, mesmo que de modo auxiliar, em uma outra categoria de serviços, que têm funcionalidades equivalentes às das plataformas e que, por isso, também lidam com esse tipo de problema. Assim, por exemplo, plataformas de recomendação, como o TripAdvisor – uma página que concentra opiniões das pessoas sobre hotéis, atrações turísticas, etc. – tem como foco primordial a venda de serviços patrocinados de viagem; mas, como as pessoas podem ali se

---

<sup>240</sup> “*Here is a representative but not exhaustive list of the social media platforms I think about, and that will be central to my concern in this book: social network sites like Facebook, LinkedIn, Google+, Hi5, Ning, NextDoor, and Foursquare; blogging and microblogging providers like Twitter, Tumblr, Blogger, Wordpress, and Livejournal; photo- and image-sharing sites like Instagram, Flickr, Pinterest, Photobucket, DeviantArt, and Snapchat; videosharing sites like YouTube, Vimeo, and Dailymotion; discussion, opinion, and gossip tools like Reddit, Digg, Secret, and Whisper; dating and hookup apps like OK Cupid, Tinder, and Grindr; collaborative knowledge tools like Wikipedia, Ask, and Quora; app stores like iTunes and Google Play; live broadcasting apps like Facebook Live and Periscope*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 26 (Kindle, n. p.).



manifestar a respeito do serviço prestado, há um serviço de moderação de conteúdo desses comentários, sujeito, inclusive, a controvérsias<sup>241</sup>. Nas palavras de Gillespie, esses são serviços que, mesmo “*sem se adequar precisamente na definição de uma plataforma, enfrentam muitos dos mesmos desafios da moderação de conteúdo em jeitos semelhantes às plataformas*”:

“[...] sites de recomendação e avaliação como Yelp e TripAdvisor; plataformas de troca que auxiliam no compartilhamento de bens, serviços, valores ou trabalho, como Etsy, Kickstarter, Craigslist, Airbnb e Uber; mundos de videogame como League of Legends, Second Life e Minecraft; mecanismos de busca como Google, Bing e Yahoo.”<sup>242</sup>

Por sua vez, estão excluídos do conceito de plataforma serviços que, mesmo que prestados pelo oferecimento de uma interface entre usuário e internet, não enfrentam os problemas associados à moderação de conteúdo, isto é, o desafio de decidir aquilo que as pessoas podem dizer em um espaço aberto ao público. Um exemplo disso – citado, erroneamente, por Balkin, como uma plataforma – é o de serviços na internet que permitem às pessoas trocarem mensagens de forma privada, seja via e-mail – como o Gmail – seja de forma instantânea – como é o caso do Messenger, do Facebook. Nesses casos, assim como nas plataformas, empresas prestam um serviço a usuários da internet, criando uma ferramenta que lhes permita utilizar a rede sem ser programadores. Não há, aqui, entretanto, atividade de hospedar, organizar ou circular o conteúdo produzido por seus usuários (item *i*). Esse dado, que é constitutivo da definição de Gillespie, faz com que esses serviços não sejam incluídos na categoria “*plataforma*” porque, ao não exercerem essa atividade, eles deixam de se expor aos desafios da moderação de conteúdo, que estão relacionados justamente à publicidade que essas plataformas conferem às manifestações das pessoas: “*como eles [os serviços de mensagem] funcionam, geralmente, de pessoa*

---

<sup>241</sup> Cf., por exemplo, a controvérsia que surgiu quando o TripAdvisor se recusou a retirar da plataforma páginas de acomodações em que mulheres foram estupradas em WARREN, R. Hotels at centre of rape allegations promoted on TripAdvisor. **The Guardian**, 5.3.2019, disponível em <<https://www.theguardian.com/world/2019/mar/05/hotels-at-centre-of-allegations-promoted-on-tripadvisor>>, acesso em 23.7.2019.

<sup>242</sup> “[...] *recommendation and rating sites like Yelp and TripAdvisor; exchange platforms that help share goods, services, funds, or labor, like Etsy, Kickstarter, Craigslist, Airbnb, and Uber; video game worlds like League of Legends, Second Life, and Minecraft; search engines like Google, Bing, and Yahoo*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 26 (Kindle, n. p.).

*para pessoa ou de grupo para grupo, e na esmagadora maioria das vezes entre contatos conhecidos, eles evitam muitos dos problemas que criam dificuldades para plataformas que oferecem visibilidade pública e contato com estranhos*<sup>243</sup>. Isso não significa que cada um desses serviços não tenha também regulamentações internas<sup>244</sup>; elas não são, entretanto, o objeto deste trabalho.

Definido o que é uma plataforma, o desafio seguinte é oferecer um conceito que explique o que é a moderação de conteúdo nas plataformas da internet.

Assim como no caso das plataformas, esse não é um termo novo neste trabalho: ao longo do primeiro capítulo, referi à moderação de conteúdo nos casos que foram reconstruídos, identificando-a com ações realizadas por plataformas – no caso, o Facebook, o Twitter e o YouTube – em relação ao conteúdo que é disponibilizado por seus usuários. Para que essa relação entre plataforma e conteúdo, exemplificada casuisticamente no capítulo anterior, ganhe contornos conceituais a seguir, proponho, a partir do conceito que elaborei acima, um aprofundamento quanto ao papel desses atores como intermediários na internet.

Plataformas, como disse, são os intermediários da Web 2.0: elas mediam a comunicação produzida por usuários na internet, disponibilizando-a na rede por mecanismos tecnológicos que a hospedam, organizam e distribuem nos espaços virtuais. Mas como exatamente funciona essa intermediação? Que tipo de intermediário surgiu com a Web 2.0?

Um primeiro e natural caminho para responder a essa pergunta é tentar achar, nas categorias de intermediários anteriores, alguma na qual as plataformas possam se encaixar. Nesse sentido, a ação de intermediários sobre a produção e a transmissão de informações era pensada, antes da internet, a partir de dois extremos<sup>245</sup>. Em um, estão as

---

<sup>243</sup> “*Because they are generally person-to-person or group-to-group, and overwhelmingly between known contacts, they sidestep many of the problems that plague platforms that offer public visibility and contact with strangers*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 29 (Kindle, n. p.).

<sup>244</sup> Cf., por exemplo, os termos de serviço do Gmail em GOOGLE. **Termos de Serviço**. Disponível em <<https://policies.google.com/terms?hl=pt-BR>>, acesso em 23.7.2019.

<sup>245</sup> “*On the one hand, we have “trusted interpersonal information conduits”— the telephone companies, the post office [...]. [...] On the other hand, we have “media content producers”— radio, film, magazines, newspapers, television, video games—that make entertainment for us [...]*” (“*De um lado, nós temos ‘condutores interpessoais de informações confiáveis – companhias telefônicas, os correios [...]. [...] De outro, nós temos os ‘produtores de conteúdo’ – radio, filme, revistas, jornais, televisões, vídeo games – que*

empresas responsáveis apenas por deslocar as informações de um lugar a outro, garantindo que a mesma mensagem que lhes é confiada por quem a confecciona é a que chega a um destinatário determinado. Esse é o caso, por exemplo, de empresas como os correios, as companhias telefônicas e, na internet, os serviços de e-mail e de mensagens instantâneas. Esses “condutores” de informação têm como pressuposto de sua atividade o fato de que não atuam sobre aquilo que levam de um lugar a outro: o serviço que prestam – e pelo qual são, de alguma forma, remunerados – é o de transportar, com integridade, algo que alguém disse, fazendo chegar essa manifestação a um terceiro previamente definido. Paga-se, portanto, pela transmissão – e não pelo que é transmitido.

Em outro extremo, estão as empresas de mídia, a exemplo das editoras, das empresas de radiodifusão – aqui se incluindo as produtoras de rádio e televisão – e mesmo daquelas que elaboravam outros produtos de entretenimento, como jogos de videogame. Ao contrário dos condutores do parágrafo anterior, o que se espera desses agentes é um papel constitutivo: a eles incumbe mobilizar o que for necessário – fontes, repórteres, editores, técnicos, etc. – para produzir a informação, construindo-a de acordo com os seus próprios critérios. Paga-se pelo conteúdo transmitido por esses agentes, e, por isso, é natural esperar que eles ajam sobre aquilo que confeccionam, modelando a informação de acordo com essa responsabilidade:

*“[...] nós compreendemos a obrigação pública desses provedores como a de produzir informação e entretenimento para o consumo do público, e nós conferimos a eles o encargo de moderar os tipos de comunicação que nos incomodam: conteúdo sexual e sensível, violência e crueldade, tipos de informação ou conhecimento perigoso”<sup>246</sup>.*

---

*fazem entretenimento para nós [...]*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 48/49 (Kindle, n. p.).

<sup>246</sup> “We understand that the public obligation of these providers is to produce information and entertainment for public consumption, and we task them in that role with moderating away the kinds of communication harms we worry about most: sexual and graphic content, violence and cruelty, dangerous kinds of information and knowledge”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 49 (Kindle, n. p.).

O problema é que as plataformas da internet não se encaixam em nenhum desses dois extremos<sup>247</sup>. De um lado, elas não se restringem a transmitir a informação veiculada por seus usuários; de outro, elas também não chegam a produzi-la. Um exemplo pode ajudar a entender esse ponto: a partir de junho de 2019, o Twitter instituiu uma nova política de conteúdo quanto àquilo que os políticos dizem na plataforma<sup>248</sup>. Publicações contrárias às regras do Twitter que sejam feitas por candidatos a cargos públicos e membros do governo com mais de 100.000 seguidores não serão sempre retiradas, como eram antes. Caso a plataforma entenda que a informação, mesmo que violadora de suas regras, é de interesse público, a publicação será mantida no ar, mas encoberta – exigindo um clique adicional para ser acessada – e irá aparecer menos em pesquisas feitas pelos usuários, reduzindo, assim, o seu alcance. Essa nova política expõe os matizes da moderação de conteúdo: a informação que é disponibilizada por um usuário não simplesmente chega a quem endereçada, mas é sujeita às restrições que, impostas, pelo Twitter, modulam a sua disponibilização na internet, estabelecendo tanto critérios para a sua permanência, quanto balizas sobre o formato pelo qual ela será disponibilizada na rede. Plataformas, portanto, não simplesmente transportam a informação alheia. É um erro, entretanto, ir ao outro ponto do espectro, e dizer que o Twitter a “*produziu*”: a publicação que é objeto da restrição foi, indiscutivelmente, confeccionada por um usuário identificado, em função do qual, inclusive, foi adotada uma política diferenciada pela plataforma.

---

<sup>247</sup> “*This makes them distinctly neither conduit nor content, not only network or only media, but a hybrid that has not been anticipated by information law or public debate*” (“*Isso faz delas [as plataformas] distintamente nem condutoras nem produtoras de conteúdo, nem só rede, nem só mídia, mas um híbrido que não foi antecipado pelo direito da informação nem pelo debate público*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 49 (Kindle, n. p.).

<sup>248</sup> *Cf.*, a esse respeito, TWITTER SAFETY. Defining public interest on Twitter. **Twitter**, 27.6.2019, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2019/publicinterest.html](https://blog.twitter.com/en_us/topics/company/2019/publicinterest.html)>, acesso em 23.7.2019; DWOSKIN, E. ROMM, T. Twitter adds labels for tweets that break its rules — a move with potentially stark implications for Trump’s account. **The Washington Post**, 27.6.2019, disponível em <[https://www.washingtonpost.com/technology/2019/06/27/twitter-adds-labels-tweets-that-break-its-rules-putting-president-trump-companys-crosshairs/?utm\\_term=.284dbf6b9605](https://www.washingtonpost.com/technology/2019/06/27/twitter-adds-labels-tweets-that-break-its-rules-putting-president-trump-companys-crosshairs/?utm_term=.284dbf6b9605)>, acesso em 22.7.2019; KELLY, M. Twitter will now hide - but not remove - harmful tweets from public figures. **The Verge**, 27.6.2019, disponível em <<https://www.theverge.com/2019/6/27/18761132/twitter-donald-trump-rules-violation-tweet-hide-remove-political-figures>>, acesso em 22.7.2019.

Dessa forma, seja por um senso de responsabilidade corporativa, derivada da consciência que têm do papel que ocupam como um espaço de comunicação entre as pessoas – e no qual, portanto, se veem na obrigação de coibir determinados comportamentos, como o discurso de ódio<sup>249</sup>; seja porque as plataformas enxergam na moderação de conteúdo uma forma de criar um ambiente mais propício à interação, que é o que lhes gera lucro<sup>250</sup>, o fato é que quando alguém faz uma publicação no Facebook, envia um *Tweet* ou disponibiliza um vídeo no YouTube, essa informação não simplesmente chega aos seus destinatários. O trajeto da informação de um usuário a outro na Web 2.0 não é uma linha reta, mas um percurso de obstáculos que devem ser superados antes que ela venha à luz na internet. Como afirma Gillespie<sup>251</sup>, as plataformas se dispõem a prestar o serviço de mediar as comunicações disponibilizadas pelas pessoas na Web 2.0, mas “*termos e condições serão aplicáveis*” – isto é, condicionantes irão agir sobre essas manifestações, de modo que, sem perder o seu traço de autoria, elas se adequem àquilo que esse novo intermediário impõe, sob pena de, no limite, deixarem de existir. É nesse sentido que a intermediação feita pelas plataformas é específica, porque posicionada em algum lugar entre a transmissão e a produção, agindo de modo tanto a restringir aquilo que poderá ser veiculado na plataforma, quanto a definir a forma pela qual isso irá acontecer. É no espaço de hibridismo entre transmitir e produzir, particular da atuação das plataformas como intermediários que não criam, mas determinam onde, como e qual conteúdo irão veicular, que emerge a moderação de conteúdo.

Assim, sob uma *perspectiva geral*, o termo moderação de conteúdo é utilizado para designar o instrumento pelo qual as plataformas da internet exercem algum nível de controle sobre as manifestações de seus usuários, dispendo quanto à forma como elas virão à luz na rede sem nunca, entretanto, tomá-las como suas. Essa visão ampla do

---

<sup>249</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.625/1.627.

<sup>250</sup> A presença e a interação entre os usuários da plataforma gera lucro tanto por valores percebidos diretamente deles – via, por exemplo, um programa de acesso diferenciado ao conteúdo, como o YouTube Premium e o LinkedIn Premium –, quanto em relação a terceiros, como pela disponibilização de espaços a anunciantes e a cobrança de valores de desenvolvedores de programas inseridos e vendidos na plataforma. Cf., por exemplo, quanto ao Facebook, BOHN, D. How Facebook makes money: 85 percent from ads, 12 percent from Zynga. **The Verge**, 1.2.2012, disponível em <<https://www.theverge.com/2012/2/1/2764825/facebook-revenue-ads-12-percent-zynga>>, acesso em 23.7.2019.

<sup>251</sup> Cf. *supra*, nota n.º 237.

fenômeno adotada, exemplificadamente, por Grimmelmann<sup>252</sup> e, inicialmente, Gillespie<sup>253</sup>, abrange desde escolhas básicas quanto àquilo que a plataforma irá se dedicar e à forma como ela será estruturada – o YouTube, por exemplo, é um serviço de compartilhamento de vídeos; o Twitter, por sua vez, enfoca em pequenas manifestações feitas a partir de perfis individuais (*microblogging*); no Flickr, fotografias e imagens reunidas em álbuns elaborados pelos seus usuários são compartilhadas – até decisões sobre o tipo de conteúdo será permitido ou não nesse ambiente – o Facebook, por exemplo, veda a nudez; já o Twitter tem uma política mais flexível a esse respeito<sup>254</sup>.

---

<sup>252</sup> “‘By ‘moderation,’ I mean the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse” (“Por ‘moderação’ eu quero expressar os mecanismos de governança que estruturam a participação em uma comunidade para facilitar a cooperação e prevenir o abuso”) GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, p. 47. Grimmelmann adiciona a essa visão ampla duas funções específica – facilitar a cooperação e prevenir o abuso; no texto, deixo de adotar essa perspectiva porque minha intenção é também abranger, por essa perspectiva geral, a moderação de conteúdo como um instrumento de resguardo dos interesses das próprias plataformas, mesmo que isso não configure um óbice à cooperação ou um abuso – a exemplo da vedação, utilizada pelo Twitter, de que anúncios de terceiros sejam veiculados sem o consentimento da plataforma; cf. *infra*, Cap. I.2, 2.3.2.

<sup>253</sup> “How platforms are designed and governed not only makes possible social activity, it calls it into being, gives it shape, and affirms its basic legitimacy as a public contribution. [...] This includes how profiles and interactions are structured; how social exchanges are preserved; how access is priced or paid for; and how information is organized algorithmically, privileging some content over others, in opaque ways” (“Como as plataformas são construídas e governadas não apenas faz possível a atividade social, mas traz ela à luz, dá-lhe formato e afirma a sua legitimidade como uma contribuição pública [...] Isso inclui como perfis e interações são estruturadas, como trocas sociais são preservadas, como o acesso é precificado ou pago; e como a informação é organizada por meio de algoritmos, privilegiando alguns conteúdos sobre outros, de formas opacas”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 30 (Kindle, n. p.).

<sup>254</sup> A política de conteúdo do Facebook veda explicitamente a publicação de imagens com “*adultos despidos*” (FACEBOOK. **Padrões da Comunidade, Parte III, Item 13: Nudez Adulta e atividades sexuais**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019); a do Twitter, permite esse tipo de publicação, mas avisa ao usuário que “*Se você compartilhar esse tipo de conteúdo em seus Tweets, deverá classificar sua conta como sensível*”, o que cobre as imagens por um aviso prévio, permitindo às pessoas decidirem se querem acessar o conteúdo (TWITTER. **Regras do Twitter – Política de Mídia Sensível**. Disponível em <<https://help.twitter.com/pt/rules-and-policies/media-policy>>, acesso em 23.7.2019).

Entendidas as plataformas, assim, como intermediários capazes “*não apenas de mediar o discurso público, mas de constitui-lo*”<sup>255</sup>, a moderação de conteúdo, sob essa visão geral, é o instrumento pelo qual a modulação do discurso alheio é feita, seja por escolhas estruturais que conformam o modo como as pessoas podem se manifestar – limitando, por exemplo, a extensão do que pode ser dito<sup>256</sup> – seja agindo diretamente sobre esse conteúdo – banindo, por exemplo, um usuário porque usa a plataforma para veicular notícias falsas<sup>257</sup>.

Essa perspectiva geral não é o objeto direto deste trabalho – que se concentra, como descrevo abaixo, em um viés mais específico da moderação de conteúdo – , mas tê-la em mente é importante. Isso porque ao abranger desde decisões estruturais até a aplicação de padrões de comportamento a usuários determinados, ela evidencia as limitações de uma narrativa que, no início, vigeu quanto à moderação de conteúdo e que é denominada, por Gillespie, como a mitologia da plataforma neutra: a afirmação, ventilada por esses intermediários, de que eles seriam imparciais tanto no sentido de que não interviriam na forma como seus usuários se expressam, tanto que, ao fazê-lo, agiriam com neutralidade quanto ao conteúdo que eles disponibilizam<sup>258</sup>.

Nesse sentido, essa perspectiva geral é útil porque esclarece que moderar conteúdo – isto é, tomar decisões que se traduzem em algum nível de controle sobre aquilo

---

<sup>255</sup> “*Platforms don’t just mediate public discourse, they constitute it*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 30 (Kindle, n. p.).

<sup>256</sup> O exemplo mais evidente disso é o Twitter, nascido com a proposta de ser um serviço de *microblogging*, que, por isso, limita o número de caracteres que um usuário pode enviar a cada publicação – o limite inicial de 140 toques foi ampliado, em 2017, para 280. Cf. ROSEN, A., IHARA, I. Giving you more characters to express yourself. **Twitter**, 26.9.2017, disponível em <[https://blog.twitter.com/en\\_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html](https://blog.twitter.com/en_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html)>, acesso em 23.7.2019; KASTRENAKES, J. Twitter says people are tweeting more, but not longer, with 280-character limit. **The Verge**, 8.2.2018, disponível em <<https://www.theverge.com/2018/2/8/16990308/twitter-280-character-tweet-length>>, acesso em 22.7.2019.

<sup>257</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>258</sup> “*The promise that platforms are impartial is a powerful one, and it is easy to embrace that promise, at two levels: that social media platforms are impartial in that they do not intervene, and that social media platforms are impartial in the way that they intervene*” (“*A promessa que as plataformas são imparciais é poderosa, e é fácil se deixar envolver por essa promessa, em dois níveis: que as plataformas de mídia social são imparciais no sentido de que elas não intervêm, e que elas são imparciais nas formas como intervêm*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 50 (Kindle, n. p.).

que as pessoas se manifestam – não é algo opcional, mas inerente às plataformas<sup>259</sup>. Não existe plataforma que não modere conteúdo, porque fazê-lo é uma parte essencial do ser uma plataforma: quaisquer que sejam os seus âmbitos de atuação, esses intermediários precisam produzir, como uma parte de sua própria constituição, escolhas quanto a de que forma os seus usuários irão disponibilizar, por meio deles, conteúdos na internet. Assim, e como exemplifiquei acima, desde os seus momentos mais embrionários, as plataformas tomam decisões que impactam diretamente sobre como as pessoas podem se manifestar: se por meio de vídeos, imagens ou textos<sup>260</sup>; se de forma aberta a qualquer pessoa ou sujeita a cadastro prévio<sup>261</sup>; se instituindo a necessidade de que as pessoas façam perfis que reflitam suas identidades reais ou não<sup>262</sup>. Cada escolha dessas não só constitui a plataforma, mas cria distinções entre o que é permitido e o que é proibido fazer naquele ambiente, balizando o modo como as pessoas poderão se manifestar ali.

Talvez o mais presente – e por vezes imperceptível – exemplo disso sejam os limites impostos pela plataforma quanto a onde as pessoas podem dizer algo. Como detalho a seguir<sup>263</sup>, as páginas do Facebook, Twitter e YouTube não são quadros brancos em que as pessoas podem livremente se expressar, mas são organizadas por campos

---

<sup>259</sup> Cf., a esse respeito, que Gillespie adiciona um quarto elemento à sua definição – “*platforms do, and must, moderate the content and activity of users, using some logistics of detection, review, and enforcement*” (“*plataformas moderam, e tem que moderar o conteúdo e a atividade de usuários, usando alguma logística de detecção, revisão e aplicação*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 29/31 (Kindle, n. p.).

<sup>260</sup> Cf. *supra*, as distinções quanto a plataformas como o YouTube, Flickr e Blogger, notas n.º 228, 227 e 225.

<sup>261</sup> Plataformas frequentemente modulam a participação das pessoas, condicionando o acesso e a participação às suas comunidades a um cadastro prévio, que usualmente toma a forma da criação de um “*perfil*” que as identifica. O Facebook, por exemplo, confere uma visualização muito limitada a quem não tem um perfil na plataforma; o Instagram e o YouTube, de outra forma, permitem que imagens e vídeos publicadas por contas públicas sejam visíveis a usuários não cadastrados; há plataformas “*equiparadas*” – como caixas de comentário em páginas de jornais – que permitem que publicações sejam veiculadas, mesmo que anonimamente.

<sup>262</sup> Há plataformas que, como o Facebook, exigem que os seus usuários forneçam as suas identidades reais no momento em que se cadastram – o chamado “*real-name system*”; outras, como o Twitter, não fazem essa exigência. Para a análise de uma controvérsia a esse respeito, cf. *infra*, Cap. I.2, 2.4.2.1, em que relato a polêmica com os nomes utilizados em perfis de *drag queens*.

<sup>263</sup> Cf. *infra*, Cap. I.2, 2.4.2.1., em que explicada a interface utilizada para se publicar algo nessas plataformas.



específicos, estabelecidos pelas plataformas, dentro dos quais as pessoas podem dizer algo, desde que obedecidas limitações de extensão e formato. Qualquer um que tente, por exemplo, veicular um texto na página inicial do YouTube, um Tweet em uma conta bloqueada ou uma publicação em um perfil fechado do Facebook simplesmente não conseguirá fazê-lo, adstrito que é às escolhas estruturais realizadas por essas plataformas ao longo de sua constituição.

Para além dessa perspectiva geral, entretanto, existe um uso específico do termo, que se refere a uma das várias formas de controle exercido pelas plataformas sobre as manifestações de seus usuários. Sob essa perspectiva *específica*, moderação de conteúdo é a atividade das plataformas de determinar, dentro da modulação do discurso de seus usuários que é capaz de fazer, *qual conteúdo deve ou não deve ser permitido veicular nos espaços por elas administrados*. Assim, Klonick afirma que usa “os termos ‘moderar’, ‘curar’ e, às vezes ‘regular’ para descrever o comportamento dessas plataformas privadas tanto de manter quanto de retirar o conteúdo gerado por usuários”<sup>264</sup>; fala-se, ainda, do trabalho dos moderadores de conteúdo<sup>265</sup>, pessoas que são responsáveis por verificar a conformidade de publicações de usuários das plataformas com as suas políticas; e ainda, sob o viés comercial, de empresas que, na lógica da terceirização, oferecem serviços de moderação de conteúdo, que a definem como “a prática de monitorar e aplicar um conjunto de regras e padrões pré-determinados a submissões advindas de usuárias para determinar se uma comunicação (uma postagem, em particular), é permissível ou não”<sup>266</sup>.

Moderar conteúdo, portanto, e nesse sentido restrito, é a prática de estabelecer e aplicar um conjunto de regras que determinam aquilo que os usuários de uma plataforma podem ou não dizer. Essa visão abarca, essencialmente, duas atividades das plataformas, que concretizam essa prática e lhe dão visibilidade. A primeira é a de estabelecer padrões prévios de comportamento de seus usuários, definindo aquilo que é ou

---

<sup>264</sup> “[...] use the terms ‘moderate,’ ‘curate,’ and sometimes ‘regulate’ to describe the behavior of these private platforms in both keeping up and taking down user-generated content”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, nota n.º 11.

<sup>265</sup> Cf. *infra*, Cap. I.2, 2.4.2.2., em que detalho o trabalho desses moderadores.

<sup>266</sup> “[...] the practice of monitoring and applying a pre-determined set of rules and guidelines to user-generated submissions to determine best if the communication (a post, in particular) is permissible or not”. TASKUS. Outsourcing Glossary. Disponível em <<https://www.taskus.com/glossary/content-moderation/>>, acesso em 22.7.2019. A TaskUs é uma empresa que é contratada por páginas da internet para terceirizar o trabalho de moderação de conteúdo.

não admissível dizer nesses ambientes<sup>267</sup>. O Facebook, por exemplo, tem “*Padrões de Comunidade*” que proíbem seus usuários de publicarem conteúdo que “*exalte a violência ou celebre a humilhação ou o sofrimento de outras pessoas*”<sup>268</sup>. A segunda é a de aplicar esses padrões, fazendo-os valer em situações concretas nas quais um usuário veicula – ou, ao menos, tenta veicular – um determinado conteúdo na plataforma. Os três casos do capítulo anterior são desdobramentos dessa última atividade.

As divisões conceituais que estabeleci acima – tanto os usos geral e específico do conceito de moderação de conteúdo; quanto as concretizações desse último, por meio do estabelecimento de padrões prévios e pela sua aplicação, são mais instrumentos didáticos do que categorias estanques e incomunicáveis entre si. Nesse sentido, e, por um lado, esta é uma dissertação enfocada mais na perspectiva restrita da moderação de conteúdo do que a geral. Isso, entretanto, não significa que seja possível tomar a definição do que se pode ou não dizer em uma plataforma como algo completamente independente, por exemplo, da forma como ela foi, desde o início, estruturada. Assim, e como é a regra nos textos sobre moderação de conteúdo, esta dissertação também transitará entre as suas perspectivas geral e específica.

Por outro, e dentro da perspectiva específica, a transformação que tematizo na segunda parte da dissertação ocorre quanto à aplicação dos padrões de comportamento dos usuários e da plataforma, e não quanto à sua constituição<sup>269</sup>. Entretanto, como o caso da imagem “*O Horror da Guerra*” expôs no capítulo anterior, a constituição e a aplicação dessas regras não acontecem de forma dissociada: naquela oportunidade, a partir do que aconteceu após a fotografia ter sido removida, os Padrões da Comunidade do Facebook foram alterados, passando a admitir que a figura passasse a ser veiculada na plataforma, porque seria uma imagem com um caráter “*icônico*”. Por isso, apresentarei, a seguir, primeiro, o funcionamento da moderação de conteúdo enquanto um conjunto de padrões de comportamento instituídos pelas plataformas e endereçados aos seus usuários; para, apenas em seguida, chegar ao viés dessa atividade enquanto uma aplicação desses padrões – cerne da problematização que trabalho na segunda parte da dissertação.

---

<sup>267</sup> Utilizo o termo “*dizer*” de forma ampla, abarcando quaisquer tipos de manifestações viabilizadas pelas plataformas – isto é, textos, imagens, sons, vídeos, etc.

<sup>268</sup> FACEBOOK. **Padrões da Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019.

<sup>269</sup> Cf. *infra*, Cap. II.3, 2.

## 2.3 A moderação de conteúdo como normatização

### 2.3.1 As regras

Como mencionei na seção anterior, a moderação de conteúdo tem um sentido específico, que se manifesta em dois tipos de atividades: primeiro, definir regras sobre o que os usuários de uma plataforma podem dizer; e, segundo, aplicar essas regras em situações concretas nas quais alguém publica ou tenta publicar algo na internet por esses intermediários. O foco deste trabalho é uma transformação localizada na segunda dessas atividades; entendê-la, entretanto, exige que a primeira seja apresentada em termos gerais, mesmo porque ambas frequentemente dialogam entre si, a exemplo de um dos casos do capítulo anterior, em que uma situação tirada da aplicação de uma regra que vedava a nudez no Facebook levou à sua alteração<sup>270</sup>.

Para fazer isso, irei analisar, de modo abreviado, as principais regras das três plataformas que compõem a linha de frente da moderação de conteúdo: o Facebook, o YouTube e o Twitter. Para ter subsídios para fazê-lo, proponho, antes, uma referência ao modo como essas regras evoluíram, contextualizando as que são acessíveis – como os Padrões de Comunidade que o Facebook publica em sua página principal – em um quadro maior, do qual também fazer partes regulamentações que, tal como a moderação de conteúdo de uma forma geral, têm como marca a opacidade.

De início, é importante delimitar o que quero dizer por “*normatização*”. Em síntese, por normatizar, refiro-me à atividade das plataformas de (i) tomarem decisões a respeito do que seus usuários podem ou não veicular por meio delas; e (ii) fixarem abstratamente essas decisões como padrões que devem ser seguidos por toda a comunidade que administram. Sob essa perspectiva, Klonick colhe, a partir de entrevistas com os pioneiros das estruturas de moderação de conteúdo do Facebook e do YouTube<sup>271</sup>, um retrato de como essa atividade funcionava inicialmente. Em síntese, em ambas as plataformas a estrutura normativa que orientava a atuação de moderadores resumia-se a uma lista não divulgada que identificava publicações que deveriam ser removidas, a partir do que Klonick chama de “*padrões*” – diretivas de comportamento “*frequentemente vagas*

---

<sup>270</sup> Cf. *supra*, o caso “*O Horror da Guerra*”, Cap. I.1, 3.2.2.1.

<sup>271</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.631.

e abertas”<sup>272</sup> – que deixavam de definir aspectos essenciais para que esse tipo de decisão fosse tomada de forma uniforme.

Um relato de um moderador à época oferece uma imagem exemplificativa a esse respeito:

*“O time de Mora-Blanco – dez pessoas no total – era apelidado de O Esquadrão (Departamento de Segurança, Qualidade e Advocacia pelo Usuário). Eles trabalhavam em times de quatro a seis [pessoas], às vezes fazendo turnos durante o dia e às vezes durante a noite, avaliando vídeos todo o tempo. [...] Mora-Blanco sentava próxima a Misty Ewing-Davis, que, estando no trabalho há alguns meses, contava como uma veterana. Na mesa diante delas havia um pedaço único de papel, dobrado em dois para mostrar uma lista de instruções dividida em itens: remova vídeos de abuso animal. Remova vídeos mostrando sangue. Remova nudez visível. Remova pornografia”<sup>273</sup> (g. n.).*

Não é difícil vislumbrar as dificuldades desse modelo: além de ele não orientar os usuários a respeito daquilo que podiam fazer – as listas eram distribuídas apenas aos moderadores –, a diversidade de conteúdos torna impossível estabelecer decisões uniformes com critérios tão genéricos. O que é, por exemplo, um vídeo “mostrando sangue”? Se um grupo de médicos resolve compartilhar imagens de uma cirurgia bem-sucedida, por exemplo, isso deveria ser retirado da plataforma?

Assim, para compreender como essas regras estão dispostas hoje, é preciso traçar a forma como esse modelo inicial, de padrões genéricos e ocultos, transformou-se na estrutura complexa normativa que rege o conteúdo das plataformas atualmente. Klonick sustenta, nesse sentido, que um motor fundamental dessa transição foram as transformações pelas quais as plataformas passaram à medida que o uso da internet se popularizou, trazendo para esses ambientes um número cada vez maior e mais heterogêneo

---

<sup>272</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.632.

<sup>273</sup> “Mora-Blanco’s team — 10 people in total — was dubbed The SQUAD (Safety, Quality, and User Advocacy Department). They worked in teams of four to six, some doing day shifts and some night, reviewing videos around the clock. [...] Mora-Blanco sat next to Misty Ewing-Davis, who, having been on the job a few months, counted as an old hand. On the table before them was a single piece of paper, folded in half to show a bullet-point list of instructions: Remove videos of animal abuse. Remove videos showing blood. Remove visible nudity. Remove pornography”. BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

de usuários. No artigo, ela relata, por exemplo, a percepção de Dave Willner<sup>274</sup>, executivo do Facebook que, nesse contexto, viu-se diante do desafio de reformular completamente a estrutura da moderação de conteúdo da plataforma, na medida em que ela deixava de ser um espaço restrito a estudantes universitários dos Estados Unidos e incorporava usuários do mundo todo<sup>275</sup>. Lidando com volumes cada vez maiores e mais diversos de conteúdo, as plataformas se obrigadas a profissionalizar a atividade de moderação, no sentido tanto de mobilizar um corpo técnico adequado, quanto de fornecer orientações claras a ele e aos usuários a respeito do que poderia ou não ser veiculado nesses espaços. Isso não significou eliminar o modelo inicial – de padrões genéricos e ocultos –, mas sim utilizar diferentes variedades normativas – específicas ou genéricas, públicas ou ocultas – de acordo com o que era mais conveniente para se moderar conteúdo.

Por isso, em um movimento detectado por Klonick como “*dos padrões às regras*”<sup>276</sup>, as plataformas passaram a adotar, de forma paralela aos padrões genéricos de antes, regras cada vez mais particularizadas, indicando, de forma pormenorizada tanto aos seus usuários, quanto, principalmente, aos seus moderadores, aquilo que poderia permanecer ou não no ambiente que administravam. Em um cenário de crescimento do número de utilizadores da plataforma – e, por consequência, de aumento de informações –, essa passagem foi fundamental para garantir uma maior consistência nas decisões tomadas

---

<sup>274</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.633.

<sup>275</sup> “*But when Facebook expanded its target membership in September 2006 from college students to anyone age thirteen or over, the influx of adults to the platform brought with it a much wider range of activities, well beyond the specific interests of well-to-do American undergraduates*” (“*Mas quando o Facebook expandiu o alvo de associados em setembro de 2006 de estudantes universitários para qualquer um com idade igual ou superior a treze anos, o influxo de adultos para a plataforma trouxe consigo um espectro muito mais amplo de atividades, muito além dos específicos interesses dos prósperos estudantes de graduação americanos*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 155 (Kindle, n. p.).

<sup>276</sup> “*The development at YouTube and Facebook from standards to rules for content moderation reflects these trade-offs. A simple standard against something like gratuitous violence is able to reach a more tailored and precise measure of justice that reflects the norms of the community, but it is vague, capricious, fact dependent, and costly to enforce*”. (“*O desenvolvimento no YouTube e no Facebook dos padrões às normas reflete esses tradeoffs. Um simples padrão contra algo como a violência gratuita consegue abranger uma medida de justiça mais precisa e ajustada, que reflete as normas da comunidade, mas é vago, inconstante, dependente de fatos e caro para aplicar*”). KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.632.

pela moderação de conteúdo, porque forneceu balizas a um número cada vez maior de moderadores, orientando-os como trabalhar mesmo ante a multiplicidade de culturas envolvidas nessa atividade, algo que exigiu, inclusive, que essas pessoas fossem recrutados em países diferentes, para que fossem capazes de compreender as publicações de acordo com as especificidades linguísticas dos usuários<sup>277</sup>.

Assim, por um lado, é possível detectar, nas plataformas, a permanência de padrões normativos com maior grau de generalidade, presentes, por exemplo, na introdução de cada um dos itens dos Padrões da Comunidade do Facebook:

*“Nós removemos conteúdo que exalte a violência ou celebre a humilhação ou o sofrimento de outras pessoas, pois tal conteúdo pode criar um ambiente que desestimula a participação”<sup>278</sup>.*

Por outro, constam nesses códigos também disposições específicas, que esclarecem aos participantes da comunidade aquilo que podem ou não compartilhar, fazendo-o de uma forma imperativa. Nesse sentido, a continuidade do excerto que citei acima detalha o que o Facebook entende como conteúdo violento e não permitido:

*“Não poste:*

*[...]*

*Vídeos de pessoas agonizando, feridas ou mortas contendo:*

*- Desmembramento, salvo se em contexto médico.*

*- Órgãos internos visíveis.*

*- Pessoas queimadas ou carbonizadas, exceto no contexto de cremação ou autoimolação como forma de discurso político ou como algo digno de notícia.*

*- Vítimas de canibalismo”<sup>279</sup>.*

---

<sup>277</sup> Em 2017, por exemplo, o Facebook resolveu abrir mais centros de moderação de conteúdo nos Estados Unidos, para que esse controle fosse feito por pessoas mais familiarizadas com as gírias do inglês. Cf. NEWTON, C. Bodies in Seats. **The Verge**, 19.6.2019, disponível em <<https://www.theverge.com/2019/6/19/18681845/facebook-moderator-interviews-video-trauma-ptsd-cognizant-tampa>>, acesso em 23.7.2019.

<sup>278</sup> FACEBOOK. **Padrões da Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019.

<sup>279</sup> FACEBOOK. **Padrões da Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019.

A utilidade de se especificar o conteúdo normativo permitido aparece com mais clareza, entretanto, não no âmbito das normativas publicadas pelas plataformas, mas nas orientações que elas fornecem aos moderadores de conteúdo. Como pessoas podem ter interpretações diferentes sobre situações submetidas ao controle da plataforma, parte do trabalho de empresas como o Facebook, o Twitter e o Google é construir manuais objetivos que unifiquem entendimentos a respeito de certas publicações, orientando os responsáveis por, no fim da cadeia de moderação de conteúdo, manter ou retirar um conteúdo. Em um exemplo disso, documentos vazados do Facebook indicam que a plataforma fornece a seus empregados tabelas com símbolos de grupos de ódio que, uma vez detectados, devem ser removidos:

**Figura 11 – “A lista do ódio”**

COUNTRY	NAME	ALTERNATIVE NAME	SYMBOLS / SLOGANS	SYMBOLS / SLOGANS	SYMBOLS / SLOGANS
Indonesia	Front Jihad Islam	FJI			
Indonesia	Laskar Pembela Islam - LPI	Soldiers of the Defenders of Islam			
International	International Gaym Party	IGP			
Israel	Lehava	לסניגורת הבחולות אכאז ויהודים			
Italy	Comunità Militante Dei Dodici Raggi, Do.Ra.	Do.Ra.			
Italy	Lealtà - Azione				
Italy	MAB Manipolo D'avanguardia Bergamo	MAB			

Fonte: *The New York Times*<sup>280</sup>

Além do espectro de normativas gerais/específicas, há outra dualidade que surge com a profissionalização da moderação de conteúdo: a de normativas públicas e ocultas. Como expus acima, inicialmente, as normas que determinavam o que podia ser

<sup>280</sup> FISHER, M. Inside Facebook’s Secret Rulebook for Global Political Speech. **The New York Times**, 27.12.2018, disponível em <<https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>>, acesso em 22.7.2019.

veiculado nas plataformas eram ocultadas, parte de um código interno de como os funcionários responsáveis pela moderação de conteúdo deveriam trabalhar. Novamente, o movimento de profissionalização não significou a eliminação dessas regras, mas a adoção de uma maior variabilidade normativa, que aqui se reflete na adoção também de códigos públicos, escritos especificamente para os seus usuários e pelos quais as plataformas detalham o que é ou não permitido dizer ali.

Cada plataforma tem, em geral, dois códigos públicos<sup>281</sup>. O primeiro, denominado comumente de “*termos de serviço*”<sup>282</sup>, é normalmente escrito na forma de um contrato, que o usuário é obrigado a aceitar para ter acesso à plataforma. Ele estabelece “*os termos dentro dos quais usuário e plataforma interagem, as obrigações que os usuários devem aceitar como uma condição para a sua participação e os meios adequados para solucionar uma disputa, caso ela ocorra*”<sup>283</sup>. O segundo, que tem títulos variáveis a depender da plataforma, usualmente fazendo referência a ela como uma “*comunidade*” – no Facebook, são os “*Padrões da Comunidade*”; no YouTube, “*Diretrizes da Comunidade*”; no Twitter, as “*Regras do Twitter*” – é um documento com linguagem mais acessível e voltada propriamente aos usuários, enunciando “*os princípios da plataforma e listando proibições, com variáveis níveis de explicação e justificativas*”<sup>284</sup>.

Paralelamente à estratégia de estabelecer esses códigos públicos, as plataformas mantiveram e detalharam compilações de regras específicas e direcionadas aos seus moderadores, explicitando aquilo que deveria ser ou não removido da plataforma.

---

<sup>281</sup> “*Most platforms present users with the rules through two main documents*” (“A maioria das plataformas apresenta aos usuários suas regras por dois documentos principais”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 54 (Kindle, n. p.).

<sup>282</sup> Cf., respectivamente, quanto a Facebook, Twitter e YouTube, FACEBOOK. **Termos de Serviço**. Disponível em <<https://www.facebook.com/legal/terms>>, acesso em 23.7.2019; TWITTER. **Termos de Serviço**. Disponível em <<https://twitter.com/pt/tos>>, acesso em 23.7.2019; GOOGLE. **Termos de Serviço**. Disponível em <<https://policies.google.com/terms?hl=pt-BR>>, acesso em 23.7.2019.

<sup>283</sup> “*The ‘terms of service’ is the more legal of the two, a contract that spells out the terms under which user and platform interact, the obligations users must accept as a condition of their participation, and the proper means of resolving a dispute should one arise*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 54 (Kindle, n. p.).

<sup>284</sup> “*It also announces the platform’s principles, and lists prohibitions, with varying degrees of explanation and justification*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 54 (Kindle, n. p.).



Diferente do que acontece com os termos de serviço e os padrões de comunidade – que são propositadamente públicos, porque a sua intenção é induzir o comportamento dos usuários da plataforma – essas regras não são publicadas, mas mantidas, internamente, sob sigilo. O conhecimento que se tem disponível quanto a essas normativas advém, basicamente, de vazamentos internos, como o *Facebook Files*, que referi acima<sup>285</sup>, em que slides utilizados para orientar novos moderadores vazaram para o jornal *The Guardian*; e as exposições subsequentes de arquivos desse tipo a veículos como *Motherboard*<sup>286</sup> e *The New York Times*<sup>287</sup>. Além disso, outros meios informais também revelam essas normativas: como reconstruído no capítulo anterior, a controvérsia quanto à negação do Holocausto no Facebook foi reacendida após Mark Zuckerberg revelar, em uma entrevista, que esse tipo de publicação era permitida na plataforma – algo que não era claro, à época, nas políticas de conteúdo publicadas pelo Facebook, mas era parte das orientações mantidas sob sigilo e entregues aos moderadores de conteúdo<sup>288</sup>.

Em síntese: a profissionalização da moderação de conteúdo, impulsionada pelo aumento tanto do número de usuários, quanto do nível de heterogeneidade entre eles, ocasionou uma transformação na forma como regras são estruturadas para definir aquilo que as pessoas podem dizer nas plataformas. Isso significou, em resumo, uma maior variabilidade normativa nesses espaços, que identifiquei, exemplificadamente, acima, a partir das transições entre generalidade / especificidade e opacidade / publicidade. O quadro que segue resume esse diagnóstico:

---

<sup>285</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>286</sup> COX, J. Leaked Documents Show Facebook’s ‘Threshold’ for Deleting Pages and Groups. **Vice**, 18.7.2018, disponível em <[https://www.vice.com/en\\_us/article/ne5nxz/leaked-documents-facebook-threshold-delete-pages-groups](https://www.vice.com/en_us/article/ne5nxz/leaked-documents-facebook-threshold-delete-pages-groups)>, acesso em 23.7.2019.

<sup>287</sup> FISHER, M. Inside Facebook’s Secret Rulebook for Global Political Speech. **The New York Times**, 27.12.2018, disponível em <<https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>>, acesso em 22.7.2019.

<sup>288</sup> Isso fica evidente mesmo pela diferença de extensão entre as normas publicadas e as utilizadas para orientar os moderadores: como relato abaixo (cf. *infra*, Cap. I.2, 2.3.2), os Padrões da Comunidade do Facebook são relativamente curtos, quando comparados com as quase 200 páginas que, de acordo com os jornais que tiveram acesso aos *Facebook Files*, compõem as orientações para identificar apenas o discurso de ódio. Cf. FISHER, M. Inside Facebook’s Secret Rulebook for Global Political Speech. **The New York Times**, 27.12.2018, disponível em <<https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>>, acesso em 22.7.2019, em específico “*The guidelines for identifying hate speech [...] run to 200 jargon-filled, head-spinning pages*” (“*As orientações para identificar o discurso de ódio [...] chegam a 200 páginas de jargões de virar a cabeça*”).

**Quadro 6 – A variabilidade normativa nas plataformas da internet**

<b><u>Normas</u></b>	<b>Gerais</b>	<b>Específicas</b>
<b>Opacas</b>	- Modelo utilizado no início da moderação de conteúdo, em que diretivas opacas e gerais eram elaboradas para os moderadores.	- Modelo utilizado nos manuais elaborados para os moderadores de conteúdo atualmente.
	Ex.: “[...] <i>remova vídeos de abuso animal. Remova vídeos mostrando sangue. Remova nudez visível. Remova pornografia</i> ” <sup>289</sup> .	Ex.: tabela em que identificados símbolos de grupos extremistas que devem ser removidos; diretiva para que publicações que neguem o Holocausto não sejam excluídas.
<b>Públicas</b>	- Modelo utilizado nos termos de serviço e nos padrões de comunidade e atualmente.	- Modelo utilizado nos termos de serviço e nos padrões de comunidade atualmente.
	Ex.: “ <i>Nós removemos conteúdo que exalte a violência ou celebre a humilhação ou o sofrimento de outras pessoas, pois tal conteúdo pode criar um ambiente que desestimula a participação</i> ” <sup>290</sup> .	Ex.: “ <i>Não poste: [...] Vídeos de pessoas agonizando, feridas ou mortas contendo: - Desmembramento, salvo se em contexto médico</i> ” <sup>291</sup> .

### 2.3.2 Os exemplos

A seguir, concluo o acesso a essas normativas exemplificando-as a partir dos padrões de comunidade adotados pelo Facebook, YouTube e Twitter. Adotei essas três plataformas por razões que já expliquei antes<sup>292</sup>; no mais, como os padrões de comunidade são os documentos que consolidam, especificamente, as políticas de conteúdo para os seus usuários – tanto que, no capítulo anterior, eles eram referidos como os documentos

<sup>289</sup> Cf. *supra*, nota n.º 273.

<sup>290</sup> Cf. *supra*, nota n.º 278.

<sup>291</sup> Cf. *supra*, nota n.º 279.

<sup>292</sup> Cf. *supra*, nota n.º 239.

fundamentais a esse respeito<sup>293</sup> – terei como foco o seu conteúdo, e não o que consta nos termos e serviços, uma postura comum à da literatura a esse respeito<sup>294</sup>.

No caso do Facebook, os Padrões da Comunidade são divididos em uma introdução, seis partes (“*comportamento violento e criminoso*”, “*segurança*”, “*conteúdo questionável*”, “*integridade e autenticidade*”, “*com respeito à propriedade intelectual*”, “*solicitações relativas a conteúdo*”) e um anexo. Na introdução, a plataforma esclarece que desenvolveu “*um conjunto de Padrões de Comunidade que detalham o que é ou não permitido no Facebook*”, para “*incentivar a expressão e criar um ambiente seguro*”, adotando-se como “*princípios*”, a “*segurança*”, a “*voz*” e a “*igualdade*”<sup>295</sup>.

O documento segue, nas seis partes que descrevi acima, via de regra apresentando, no início de cada uma, uma descrição geral da política de conteúdo – a regra geral –, seguida de especificações que indicam o que não deve ser publicado – a regra específica. Na primeira parte, denominada “*comportamento violento e criminoso*”, estão reunidas as proibições à violência e à sua incitação; à presença de organizações e indivíduos perigosos na plataforma; à promoção e organização de condutas criminosas e a negociação de produtos controlados. Na segunda, “*segurança*”, estão reunidas as restrições a publicações que exponham ou promovam de alguma forma a “*automutilação e o suicídio*”; a “*nudez infantil e exploração sexual de crianças*”; a “*exploração sexual de adultos*”; o “*bullying e o assédio*”; e que impliquem em violações da privacidade e direitos de privacidade de imagem. Na terceira parte, “*conteúdo questionável*”, a plataforma expõe proibições ao discurso de ódio; à violência e conteúdo explícito; à nudez adulta e atividades sexuais – Facebook, ao contrário do Twitter, por exemplo, veda a pornografia –; a abordagens sexuais, mesmo que entre adultos; e, por fim, a conteúdo de teor cruel ou insensível.

---

<sup>293</sup> Esse papel fundamental aparece tanto no caso Egeland, em que os padrões foram utilizados para justificar a exclusão da imagem, por violar a política sobre nudez; quanto no caso Alex Jones, em que uma mudança nessas regras foi a justificativa para a exclusão do ativista do Facebook. Cf. *supra*, Cap. I.1, 3.2.2.1 e 3.2.2.2.

<sup>294</sup> É a postura adotada, por exemplo, por Gillespie; cf. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 55 e ss (Kindle, n. p.).

<sup>295</sup> As citações dos Padrões da Comunidade do Facebook foram retiradas de FACEBOOK. **Padrões da Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019

Na quarta parte, há a proibição ao spam, a exigência – que não é comum a todas as plataformas<sup>296</sup> – de que as pessoas utilizem o Facebook com o seu nome real, a proibição de notícias falsas – que baseou, ao menos em parte, a exclusão de Alex Jones – e a possibilidade de que, com o falecimento, o perfil de uma pessoa seja transformada em um memorial. A quinta parte é dedicada exclusivamente à proteção à propriedade intelectual, tanto responsabilizando os usuários pelo que veiculam – “*você é proprietário de todas as informações e conteúdos que publica no Facebook*” – quanto vedando a utilização da plataforma para infringi-la – “*Os Termos e Serviços do Facebook não permitem a publicação de conteúdo que viole os direitos de propriedade intelectual de outras pessoas, inclusive direitos autorais e de marca comercial*”. Na sexta parte, a plataforma esclarece que atende solicitações de retirada de conteúdo advindas de usuários, familiares e governos, a depender do caso – esses pedidos são respectivamente acolhidos, por exemplo, quando o usuário pede que a própria conta seja excluída, quando um familiar pede a retirada do perfil de alguém que faleceu ou quando o Estado pede a remoção de uma imagem de abuso infantil. Por fim, também há um anexo em que o Facebook esclarece como as pessoas podem participar na elaboração dos Padrões da Comunidade.

No caso do YouTube, as “*Diretrizes da Comunidade*”<sup>297</sup> estão em uma página de “*Políticas e Segurança*” e divididas em doze grupos temáticos, organizados, de forma semelhante à do Facebook, em uma descrição geral da política (por exemplo, “*não é permitido exigir no YouTube conteúdo que coloca em perigo o bem-estar emocional e físico de menores*”), esclarecida tanto por diretivas que traduzem isso em proibições (“*não poste [...] atos nocivos ou perigosos envolvendo menores*”) e exemplos (“*mostrar menores envolvidos em atividades perigosas, por exemplo, praticando acrobacias físicas, manejando armas ou explosivos ou usando substâncias controladas, como álcool e nicotina*”).

No primeiro grupo, há as proibições à nudez e ao conteúdo sexual – à semelhança do Facebook, o YouTube as veda –; no segundo, ao conteúdo prejudicial ou perigoso, como “*vídeos que incentivem outras pessoas, especialmente crianças, a realizar ações que possam causar lesões corporais graves*”; no terceiro, à incitação ao ódio; no

---

<sup>296</sup> Cf. *supra*, nota n.º 262, e *infra*, Cap. I.2, 2.4.2.1, ambos quanto à controvérsia do uso de nomes reais por perfis de *drag queens* no Facebook.

<sup>297</sup> As citações das Diretrizes da Comunidade do YouTube foram retiradas de YOUTUBE. **Diretrizes da Comunidade**. Disponível em <<https://www.youtube.com/intl/pt-BR/yt/about/policies/#community-guidelines/>>, acesso em 23.7.2019

quarto, ao conteúdo explícito ou violento, em específico “*com a intenção principal de chocar, impressionar ou desrespeitar o público*”; no quinto, veda-se o assédio e o bullying; e, no sexto, a utilização da plataforma para enviar spam ou realizar golpes.

No sétimo grupo, há a vedação à realização de ameaças; no oitavo – tal como o Facebook – há um item dedicado exclusivamente à proteção dos direitos autorais, algo sensível em uma plataforma de vídeos como o YouTube, que, por isso, veda o envio de “*vídeos que não tenham sido feitos por você ou [que utilizem] [...] conteúdo de propriedade de terceiros, como músicas, trechos de programas protegidos por direitos autorais ou vídeos feitos por outros usuários sem as autorizações necessárias*”. No nono grupo, há a proteção à privacidade – proibindo-se a divulgação não consentida de informações pessoais de terceiros –; no décimo, a proibição à falsidade; no décimo primeiro, um conjunto de regras voltadas à proteção infantil, que vedam, por exemplo, a “*sexualização de menores*”, “*atos nocivos ou perigosos envolvendo menores*”, a “*imposição de sofrimento emocional em menores*”, “*conteúdo familiar enganoso*” e “*bullying virtual e assédio envolvendo menores*”. Por fim, no décimo segundo item, são reunidas políticas adicionais sobre linguagem de baixo calão, contas inativas, publicações que “*encorajam a violação dos Termos de Serviço*” e restrições etárias.

No Twitter, esses padrões estão resumidos na área “*As Regras do Twitter*”<sup>298</sup>. Como no caso do Facebook, há uma introdução, na qual a plataforma esclarece que “*nossas regras foram criadas para garantir que todas as pessoas possam participar do diálogo público com liberdade e segurança*”. Em seguida, as regras são divididas em cinco temáticas diferentes, disponíveis em uma versão abreviada e outra distendida, que, como nas plataformas anteriores, é composta por aspectos genéricos (“*temos uma política contra ameaças de violência no Twitter [...] definimos ameaças violentas como declarações com a intenção de matar ou ferir gravemente uma pessoa ou um grupo de pessoas*”) e específicos (“*Violações dessa política incluem, mas não se limita a: [...] ameaçar de matar alguém*”). Adicionalmente, em cada temática há uma seção à parte, que explicita a sanção em caso de descumprimento (“*Suspenderemos imediatamente e permanentemente as contas que publicarem ameaças violentas*”).

A primeira temática é a de segurança, e inclui vedações à violência, ao extremismo, ao abuso e assédio, ao suicídio e à automutilação, à veiculação de mídias

---

<sup>298</sup> As citações das Regras do Twitter foram retiradas de TWITTER. **Regras do Twitter – Política de Mídia Sensível**. Disponível em <<https://help.twitter.com/pt/rules-and-policies/media-policy>>, acesso em 23.7.2019.

sensíveis – dentre as quais conteúdo adulto violento; o Twitter, entretanto, permite a nudez, ao contrário do Facebook e do YouTube –, e à “*venda, compra ou facilitação de transações em produtos ou serviços ilegais*”. A segunda temática, relativa à privacidade, veda a publicação de informações privadas, e a exposição de nudez não consensual; a terceira, quanto à “*autenticidade*”, proíbe o spam, a manipulação da plataforma, o uso do Twitter para interferir na integridade nas eleições, inclusive a “*publicação ou o compartilhamento de conteúdo que possa diminuir o comparecimento dos eleitores*”, a assunção da identidade de terceiros e a violação de direitos autorais e de marca registrada. Na quarta temática, o Twitter explica qual é a sua abordagem na “*elaboração de políticas*” e qual é a sua “*filosofia de medidas corretivas*”; e, por fim, na quinta, há a proibição de se veicular publicidade de terceiros sem o consentimento prévio da plataforma.

Os exemplos acima expõem, para além do diagnóstico do início desta seção, algo importante quanto à moderação de conteúdo – seja sob a perspectiva ampla, seja sob a perspectiva restrita: a despeito da existência de diferenças quanto àquilo que é permitido ou proibido – o Facebook e o YouTube, por exemplo, vedam a nudez; o Twitter, não –, existe uma tendência à uniformização entre as diferentes manifestações desse mecanismo. De uma forma geral, nas três plataformas: (i) as normativas são organizadas de forma semelhante, dispostas a partir de temáticas compartilhadas, como segurança, violência, conteúdo adulto, proteção à criança e violação a direitos autorais; (ii) é comum que essas normas não tenham apenas comandos aos usuários, mas sejam robustecidas por justificativas que apresentem as razões por trás de uma determinada postura da plataforma – o Twitter, por exemplo, baseia sua proibição discurso de ódio a partir de seu compromisso de “*combater o assédio motivado por ódio, preconceito ou intolerância, particularmente aquele que tem o objetivo de silenciar as vozes de quem é historicamente marginalizado*”; (iii) as normas da comunidade, inicialmente voltadas apenas ao conteúdo disponibilizado pelos usuários – isto é, a moderação de conteúdo em sentido estrito – também são aproveitados como uma oportunidade de esclarecimento, pelas plataformas, de aspectos relativos à moderação de conteúdo em sentido amplo, como, por exemplo, quanto às pessoas que são admitidas ou não no Facebook – assim, o tópico, do Facebook, “*comportamento violento e criminoso*”, elenca tanto a proibição à “*promoção ou divulgação de crimes violentos, roubos e/ou fraudes*”, quanto a vedação a que “*organizações ou indivíduos envolvidos nas atividades abaixo estejam no Facebook: atividade terrorista; ódio organizado*”.

Apresentada a moderação de conteúdo, em sentido estrito, sob a sua atividade de estabelecer padrões de comportamento aos seus usuários – algo marcado, durante o processo de profissionalização das plataformas, por uma maior variabilidade normativa, adotando-se regras específicas e genéricas, públicas e sigilosas, a depender da circunstância –, analiso, na seção seguinte, a sua segunda forma de se manifestar: a aplicação das regras expostas acima em casos concretos, nos quais um usuário publica ou tenta publicar algo em uma plataforma e ela resolve agir, via moderação de conteúdo.

## 2.4 A moderação de conteúdo como adjudicação

### 2.4.1 Introdução

Introduzi este trabalho a partir de uma situação concreta: o caso Lipman<sup>299</sup>. Como relatei acima, um grupo nacionalista israelense publicou, em 2016, uma imagem no Facebook em que indicava um território em disputa na Oriente-Médio e o identificava como sendo Israel, e não a Palestina. Sem maiores explicações, a publicação foi deletada pela plataforma e os administradores da página foram proibidos de fazer postagens no Facebook por três dias, bem como ordenados a retirar outras publicações consideradas “*problemáticas*”. Para testar a coerência do Facebook, Dov Lipman, um parlamentar israelense, publicou a mesma imagem com os nomes “*Israel*” e “*Palestina*” invertidos e reportou, de outro perfil, a figura à mesma moderação de conteúdo do Facebook, que decidiu, dessa vez, que publicação deveria permanecer ativa porque “*não violava os nossos Padrões de Comunidade*”. A partir desse contexto, Lipman escreveu uma carta aberta ao Facebook, relatando o experimento e exigindo uma explicação do porquê uma imagem foi mantida e a outra retirada. Sem dar uma resposta específica a essas perguntas, o Facebook reestabeleceu a publicação originária e enviou uma mensagem padronizada de desculpas à página.

Esta seção se dedica explicar o que aconteceu em casos como o de Dov Lipman: seu objetivo é esclarecer como funciona o mecanismo utilizado pela moderação de conteúdo para aplicar os padrões estabelecidos na seção anterior a situações concretas. Daqui em diante, irei me referir a ele com a ideia de “*adjudicação*”, que tomo aqui como a de aplicação de padrões normativos abstratos a casos concretos.

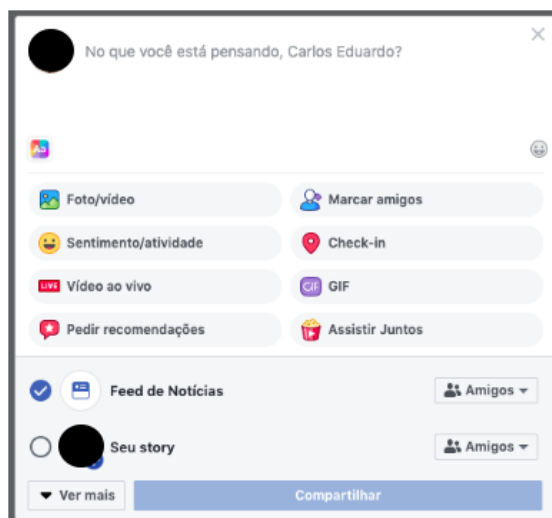
---

<sup>299</sup> Cf. *supra*, Introdução.

De início, é preciso compreender como os usuários das plataformas veiculam algo na internet por meio delas. Para isso, utilizarei, novamente, como paradigma, o que acontece no Facebook, no Twitter e no YouTube<sup>300</sup>. Em síntese, em todas essas plataformas há campos específicos que delimitam – em uma manifestação da moderação de conteúdo em sentido amplo – um local no qual as pessoas podem se manifestar, fazendo-o de acordo com um formato – tipo de letra, tamanho de arquivo, extensão, etc. – previamente determinado. Uma vez composta a publicação – seja ela uma mensagem, uma foto ou um vídeo –, o usuário aciona um comando específico – “*compartilhar*”, “*publicar*”, “*tweetar*”, etc. – que envia a publicação para a plataforma que, em seguida, veicula-a no espaço apropriado. Ter em mente essa sequência de atos é importante porque, como detalho a seguir, a moderação de conteúdo pode agir em diferentes momentos entre a inserção do conteúdo pelo usuário até depois de a publicação estar na plataforma.

Dessa forma, no Facebook, publica-se algo na plataforma por meio de campos tais como o que reproduzi abaixo, no qual é possível escrever uma mensagem e anexar, a ela, fotos e vídeos, e introduzir informações adicionais a respeito da localização ou de pessoas que estejam a ela vinculados. Pelo comando “*compartilhar*”, a informação é enviada à plataforma e, depois, publicada. Mecanismo semelhantes são utilizados em outros âmbitos do ambiente virtual criado pela plataforma para, por exemplo, comentar uma postagem de outra pessoa ou se expressar em um grupo fechado.

**Figura 12 – Publicação no Facebook**



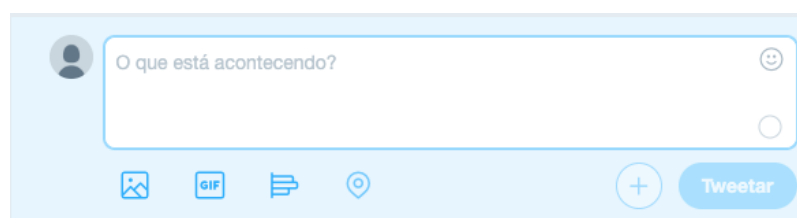
<sup>300</sup> Para a justificativa desse recorte, cf. *supra*, nota n.º 239.



Fonte: Facebook<sup>301</sup>

No Twitter, o mecanismo é semelhante: há um campo específico, em que o usuário digita sua mensagem e anexa a ela mídias, como fotos e vídeos, agregando, se quiser, a sua localização, uma sequência de mensagens, etc. O Twitter, por ser um serviço de *microblogging*, que se propõe a viabilizar a troca de mensagens rápidas, limita o número de caracteres a 280 – antes, eram 140<sup>302</sup> – algo para o que chamei a atenção ao expor a moderação de conteúdo em sentido amplo: mesmo as decisões mais iniciais das plataformas, tais como as que definem o seu nicho de atuação, já representam alguma forma de limitação quanto à forma como as pessoas poderão se expressar.

**Figura 13 – Publicação no Twitter**



Fonte: Twitter<sup>303</sup>

No YouTube, a formatação da área de publicação é diferente das anteriores, porque a plataforma se dedica ao compartilhamento de vídeos – novamente uma manifestação da moderação de conteúdo em sentido amplo. Além disso, é interessante perceber que, diferente dos casos anteriores, a plataforma reafirma ao usuário que veicula algo ali que ele “*concorda com os Termos de Serviço e com as Diretrizes da comunidade do YouTube*” – que serão aplicadas justamente pelo mecanismo de moderação de conteúdo que exponho a seguir.

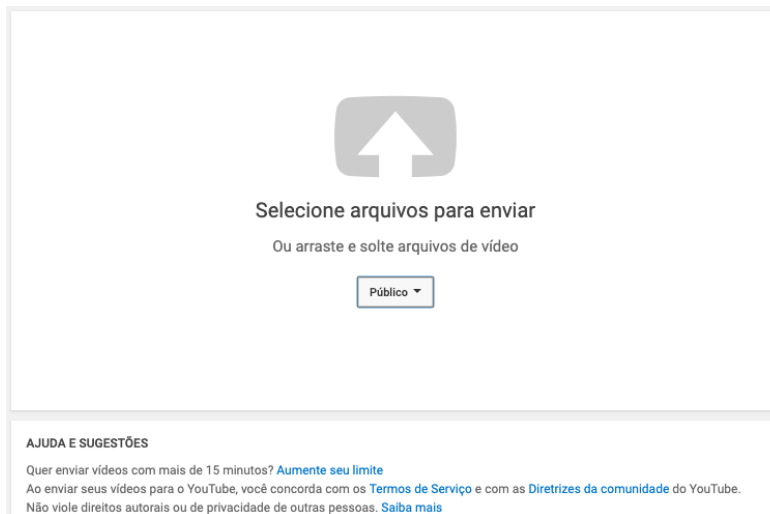
---

<sup>301</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“*screenshot*”); para tanto, acessei, como usuário, uma conta do Facebook. A figura é um retrato do que aparece na parte superior central da tela, após um clique no campo “*No que você está pensando*”.

<sup>302</sup> Cf. *supra*, nota n.º 256.

<sup>303</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“*screenshot*”); para tanto, acessei, como usuário, uma conta do Twitter. A figura é um retrato do que aparece na parte superior central da tela.

**Figura 14 – Publicação no YouTube**



Fonte: YouTube<sup>304</sup>

Exposto como as pessoas se manifestam no Facebook, Twitter e YouTube, o passo seguinte é analisar como funciona a moderação de conteúdo, uma vez que um usuário decida publicar algo – em particular, a partir do momento em que ele respectivamente aciona o comando de “*compartilhar*”, “*tweetar*” ou “*enviar*” – ou até antes<sup>305</sup>. Em síntese, a pergunta que precisa ser respondida é a seguinte: como algo que um usuário decide publicar é controlado pela moderação de conteúdo da plataforma?

O primeiro passo para compreender como funciona esse mecanismo é entender que a moderação de conteúdo funciona em etapas, isto é, momentos distinguíveis pela forma como a plataforma interage com o conteúdo publicado pelo usuário (2.4.2). Há

<sup>304</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“*screenshot*”); para tanto, acessei, como usuário, uma conta do YouTube. A figura é um retrato do que aparece na parte central da tela, no menu “*Meus vídeos*”, opção “*Enviar vídeo*”.

<sup>305</sup> Como uma consequência da opacidade da moderação de conteúdo, não há como se saber ao certo se a ferramenta também é utilizada de alguma forma quanto ao conteúdo que o usuário digita em campos da plataforma, mesmo antes de enviá-lo. É certo, entretanto, que essas informações são enviadas à plataforma: em 2013, um estudo de dois cientistas de dados causou polêmica ao analisar a censura imposta aos usuários a si mesmos, a partir do conteúdo que 3.9 milhões de pessoas digitaram mas não publicaram na plataforma. Cf. DAS, S., KRAMER, A. Self-Censorship on Facebook. **Proceedings of the Seventh International Conference on Weblogs and Social Media**, 2013, disponível em <<https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6093/6350>>, acesso em 23.7.2019; GOLBECK, J. On Second Thought... **Slate**, 13.12.2013, disponível em <<https://slate.com/technology/2013/12/facebook-self-censorship-what-happens-to-the-posts-you-dont-publish.html>>, acesso em 23.7.2019.

duas etapas essenciais que, a depender da situação, ocasionam uma terceira. A primeira é a de *detectar* o conteúdo (2.4.2.1); a segunda, que lhe é subsequente, é a de *tomar uma decisão* quanto a ele (2.4.2.2); e a terceira, que ocorre apenas se verificada uma violação, é a de *aplicar uma sanção* pela sua veiculação (2.4.2.3). Abordo-as a seguir, dedicando uma seção a cada uma delas e, em seguida, propondo uma conclusão que as articula em uma espécie de taxonomia, que organiza possibilidades em uma sistematização que ajuda a entender o que aconteceu em casos como os que mencionei no capítulo anterior.

## 2.4.2 As Fases da Moderação de Conteúdo

### 2.4.2.1 Fase n.º 1: Detectar

Detectar conteúdo significa selecionar aquilo que será submetido à análise dos mecanismos de moderação de conteúdo – que irão decidir, na fase seguinte, se algo permanece ou não na plataforma. Em um cenário ideal, a plataforma seria capaz de analisar todo o conteúdo que os seus usuários veiculam, submetendo-o a toda a extensão de suas políticas de conteúdo, garantindo, assim, que as publicações presentes no ambiente que administra observam as suas regras. O problema é que, como expus antes, o principal motor da profissionalização da moderação de conteúdo foi a popularização da internet, que trouxe um número imenso de usuários – e conteúdo – para esses ambientes, tornando-os cada vez mais heterogêneos. Como o exemplo do Facebook expõe, aquilo que antes era restrito aos universitários dos Estados Unidos agora é utilizado por pessoas do mundo inteiro. Alguns dados podem tornar isso mais concreto: o Facebook tinha 2,4 bilhões de usuários ativos mensalmente em 2019<sup>306</sup> – estatísticas de 2013, fornecidas pela própria plataforma em um momento em que o número de usuários era menor, indicavam que 350 milhões de fotos eram veiculadas diariamente em seu ambiente, em média, totalizando, à época, 250 bilhões de imagens enviadas<sup>307</sup>; o Twitter tem 321 milhões de usuários

---

<sup>306</sup> FACEBOOK. Stats. **Facebook Newsroom**, 2018, disponível em <<https://newsroom.fb.com/company-info/#statistics>>, acesso em 23.7.2019; um usuário mensalmente ativo é aquele que entrou pelo menos uma vez na plataforma no último mês.

<sup>307</sup> “More than 250 billion photos have been uploaded to Facebook, and more than 350 million photos are uploaded every day on average” (“Mais de 250 bilhões de fotos foram enviadas para o Facebook, e mais de 350 milhões de fotos são enviadas todos os dias, em média”). INTERNET.ORG. **A focus on efficiency: a white paper from Facebook, Ericsson and Qualcomm**. 16.9.2013, disponível em

mensalmente ativos<sup>308</sup> – que enviam, aproximadamente, 500 milhões de tweets por dia<sup>309</sup>; o YouTube tem 1,9 bilhões de usuários ativos<sup>310</sup> – que, segundo o dado de 2015, levam à plataforma 400 horas de vídeos a cada minuto<sup>311</sup>.

O estabelecimento de um mecanismo de detecção, nesse sentido, traduz uma situação de escassez: se, por um lado, a quantidade de conteúdo veiculado nas plataformas é quase incomensurável, por outro, os recursos disponibilizados para analisá-los – sejam eles automatizados ou humanos – são escassos. Essa escassez é explicada, por alguns autores, como Gillespie<sup>312</sup>, como “*o problema da escala*”: a mais óbvia diferença entre a mídia tradicional e as plataformas – “*o imenso número de usuários, a imensa quantidade de conteúdo e o ritmo incansável no qual eles circulam*”<sup>313</sup> – traduz-se na mais óbvia diferença entre a moderação feita, por exemplo, pelo conselho editorial de um jornal, e a que se exige de uma plataforma como o Facebook – a escala que a segunda precisa atingir para ser eficiente. É dessa necessidade que decorre a exigência de que duas seleções sejam realizadas: por um lado, é preciso escolher *o que* será analisado; por outro, é preciso

---

<<https://pt.scribd.com/document/168671112/A-Focus-On-Efficiency-Internet-org-Whitepaper>>, acesso em 23.7.2019.

<sup>308</sup> SHABAN, H. Twitter reveals its daily active user numbers for the first time. **The Washington Post**, 7.2.2019, disponível em <[https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-first-time/?utm\\_term=.84f6d67871a1](https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-first-time/?utm_term=.84f6d67871a1)>, acesso em 23.7.2019.

<sup>309</sup> INTERNET LIVE STATS. **Twitter Usage Statistics**. Disponível em <<https://www.internetlivestats.com/twitter-statistics/>>, acesso em 23.7.2019.

<sup>310</sup> YOUTUBE. **YouTube in Numbers**. Disponível em <<https://www.youtube.com/intl/en-GB/yt/about/press/>>, acesso em 23.7.2019.

<sup>311</sup> STATISTA. **Hours of video uploaded to YouTube every minute 2007-2019**. 15.7.2019, disponível em <<https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>>, acesso em 23.7.2019.

<sup>312</sup> “*What is different for today’s social media platforms is that they host and oversee an unprecedented amount of content and an unprecedented number of people, and must moderate on a qualitatively different scale*” (“*O que é diferente para as plataformas de mídia social hoje é o fato de que elas hospedam e fiscalizam uma quantidade sem precedentes de conteúdo e um número sem precedentes de pessoas, e precisam moderar em uma escala qualitativamente diferente*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 83/84 (Kindle, n. p.).

<sup>313</sup> “*The most obvious difference is the sheer number of users, the sheer amount of content, and the relentless pace at which they circulate*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 83 (Kindle, n. p.).

circunscrever *diante de que parte* das normas da comunidade essa análise será feita. É isso que os mecanismos de detecção fazem: ao mesmo tempo em que eles pinçam determinados conteúdos do universo das plataformas, eles também predisõem a análise decisória, posicionando-os de acordo com a regra que a publicação teria violado.

No caso Lipman, por exemplo, o parlamentar reportou a imagem que havia postado, levando-o à consideração da plataforma por um formulário digital – um dos meios de detecção disponível –, em que é possível assinalar qual seria a violação pela qual a publicação é denunciada. No caso Egeland, o primeiro do capítulo anterior, é provável que a imagem tenha sido enviada à moderação de conteúdo do Facebook por mecanismos automáticos de detecção de nudez; por sua vez, tanto as publicações de Alex Jones, quanto as páginas que negaram o Holocausto chegaram à administração da plataforma por outros usuários, que se viram incomodados com o conteúdo das publicações – respectivamente a acionando por formulários internos e por meios de pressão externos, como a carta enviada pela comunidade judia<sup>314</sup>.

Na linha que fez Balkin com a regulação da liberdade de expressão, classificando os mecanismos pelos quais os Estados intervêm na capacidade de as pessoas se manifestarem na era digital, proponho a seguir uma taxonomia dos mecanismos de detecção das plataformas, fazendo-o a partir de um diálogo com a literatura especializada e investigações jornalísticas a esse respeito e adotando, nesse sentido, a visão pluralista desse tipo de regulação, como argumentei no final do capítulo anterior.

Meu objetivo, nesse sentido, não é esgotar as classificações possíveis a respeito da moderação de conteúdo, mas sim estruturar as suas principais modalidades para, em seguida, utilizá-las para entender melhor como esse mecanismo funciona, contextualizando-o com casos concretos e organizando melhor a forma como ele é apresentado na literatura. Em específico, irei dispor do modo como autores de atuação especificamente voltada às plataformas da internet – notadamente Tarleton Gillespie<sup>315</sup> e Kate Klonick<sup>316</sup> – receberam abordagens mais gerais – como a feita por James

---

<sup>314</sup> Cf. *supra*, nota n.º 159.

<sup>315</sup> Gillespie faz isso, em específico, no primeiro capítulo de seu trabalho, expressamente citando Grimmelman, por exemplo, em GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, Cap. I.1, nota n.º 74 (Kindle, n. p.).

<sup>316</sup> Klonick evidencia essa particularização de uma abordagem mais geral, por exemplo, ao citar – no texto que tenho utilizado como base para citar o seu trabalho – a taxonomia de Grimmelman e caracterizá-la como

Grimmelmann<sup>317</sup> –, propondo, ao final, quatro formas de classificar a moderação de conteúdo a partir do modo como aquilo que os usuários da plataforma publicam – ou tentam publicar – é detectado e levado à atividade propriamente decisória.

A primeira classificação utiliza como critério a origem da qual provém a provocação, feita à plataforma, de que algo deve ser por ela analisado – e que pode advir tanto de fora, quanto de dentro do ambiente que ela administra.

Nesse sentido, é preciso se ter em conta que, a despeito de a plataforma controlar aquilo que as pessoas dizem apenas em seu próprio ambiente, posicionamentos – e, principalmente, controvérsias – quanto aquilo que é dito nesse espaço não ficam restritos a ele, mas, como os casos do capítulo anterior expõem, frequentemente se desenvolvem em contextos extraplataforma, desencadeando exigências quanto à moderação de conteúdo, que detecta essas situações a partir de estímulos externos – a exemplo de cartas abertas, editoriais, matérias de jornais e notificações de governos e empresas. No capítulo anterior, um exemplo disso é o modo como o Facebook foi instado a se posicionar a respeito de páginas que negavam o Holocausto, e que lhe foram apresentadas, primeiro, por uma carta aberta de vítimas judias<sup>318</sup> e, segundo, por uma série de reportagens posteriores a uma entrevista, de Zuckerberg, na qual disse que não deletaria publicações com esse teor<sup>319</sup>.

No mesmo sentido, a atuação da moderação de conteúdo do Facebook nos Atentados de Christchurch, que ocorreram na Nova Zelândia. O caso se desenvolveu em 15 de março de 2019, após Brenton Tarrant, um australiano de 28 anos, entrar em uma mesquita e abrir fogo contra fiéis muçulmanos, matando 51 e ferindo 49 pessoas<sup>320</sup>, em um ataque vinculado ao supremacismo branco<sup>321</sup>. A ação foi transmitida ao vivo, pelo atirador,

---

*“fundacional, mas que fala de forma mais geral quanto à toda moderação da internet, ao invés de se referir, especificamente, às plataformas que publicam conteúdo”.* KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, nota n.º 261.

<sup>317</sup> Cf. GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, pp. 42/109, em que o autor expressamente propõe uma taxonomia da moderação de conteúdo.

<sup>318</sup> Cf. *supra*, nota n.º 159.

<sup>319</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>320</sup> Cf., por exemplo, TARABAY, J., GRAHAM-MCLAY, C. Could the Christchurch Attacks Have Been Prevented? **The New York Times**, 18.6.2019, disponível em <<https://www.nytimes.com/2019/06/18/world/australia/new-zealand-terrorism-christchurch.html>>, acesso em 23.7.2019.

<sup>321</sup> Cf., por exemplo, KINGSLEY, P. New Zealand Massacre Highlights Global Reach of White Extremism. **The New York Times**, 15.3.2019, disponível em <<https://www.nytimes.com/2019/06/18/world/australia/new-zealand-terrorism-christchurch.html>>, acesso em 23.7.2019.

utilizando-se de uma ferramenta do Facebook chamada “*livestream*”<sup>322</sup> e permaneceu sendo exibida e compartilhada na plataforma até que ela foi alertada pela polícia a respeito, acionando a moderação de conteúdo que, em 24 horas, removeu 1,5 milhão de vídeos do Facebook e do Instagram – dos quais 1,2 milhão foram retirados antes que fossem efetivamente publicados<sup>323</sup>. Em um sentido semelhante, o Twitter deletou uma publicação de Donald Trump que utilizava a trilha sonora do filme *The Dark Knight* após ter sido formalmente notificado pela Warner Bros., por mecanismos externos à plataforma, de que o uso não havia sido autorizado e violava os seus direitos autorais<sup>324</sup>.

Para além do obtido extraplataforma, empresas como o Facebook, o Twitter e o Google também detectam conteúdo por meio de mecanismos intraplataformas, isto é, posicionados dentro desses ambientes e que levam o conteúdo que é neles compartilhado à moderação de conteúdo. Assim, como descrevi no caso Lipman, a forma que o parlamentar encontrou para submeter a imagem por ele modificada à plataforma foi reportando-a, isto é, utilizando-se de uma funcionalidade interna que submete uma publicação à moderação de conteúdo. Outra possibilidade, que discuto com mais detalhes adiante, é a existência de ferramentas automatizadas que, independentemente da ação de usuário, detectam conteúdo potencialmente violador dos padrões das plataformas e os levam à moderação de conteúdo – em 2017, por exemplo, o Twitter informou que removeu 270 mil contas porque veicularam publicações que promovem o terrorismo, sendo que, segundo o porta-voz da plataforma, “*a esmagadora maioria dessas contas foram detectadas por nossa própria tecnologia, sendo que apenas 0,2% das contas suspensas em 2017 o foram em razão de*

---

<sup>322</sup> A ferramenta permite que o usuário ative a câmera do celular e faça uma transmissão ao vivo na plataforma. Cf. FACEBOOK. **Facebook Live**. Disponível em <<https://live.fb.com>>, acesso em 23.7.2019.

<sup>323</sup> A informação foi veiculada na conta do Twitter do Facebook; cf. FACEBOOK NEWSROOM (Facebook). Sem título. Sem localização, 17 de março de 2019. **Twitter**: fbnewsroom. Disponível em <[https://twitter.com/fbnewsroom/status/1107117981358682112?ref\\_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref\\_url=https%3A%2F%2F](https://twitter.com/fbnewsroom/status/1107117981358682112?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref_url=https%3A%2F%2F)>, acesso em 23.7.2019; no mesmo sentido, LIPTAK, A. Facebook says that it removed 1.5 million videos of the New Zealand mass shooting. **The Verge**, 17.3.2019, disponível em <<https://www.theverge.com/2019/3/17/18269453/facebook-new-zealand-attack-removed-1-5-million-videos-content-moderation>>, acesso em 23.7.2019.

<sup>324</sup> STATT, N. Twitter removes Trump’s video featuring music from *The Dark Knight* for copyright infringement. **The Verge**, 10.4.2019, disponível em <<https://www.theverge.com/2019/4/10/18303841/trump-twitter-the-dark-knight-rises-copyright-infringement-video-removed-hans-zimmer>>, acesso em 23.7.2019.

*sinalizações enviadas pela polícia*”<sup>325</sup>. Essa é, provavelmente, a via utilizada no primeiro caso do capítulo anterior, em que a nudez da imagem publicada por Egeland foi levada à moderação de conteúdo por mecanismos que a detectaram automaticamente.

A segunda classificação, apresentada nesses termos por Klonick<sup>326</sup>, e por ela apreendida – e especificada – a partir de Grimmelmann<sup>327</sup>, toma como critério o momento em que o conteúdo objeto da moderação é detectado – se antes ou depois de ser publicado. Há, assim, a moderação de conteúdo *ex ante* – isto é, preventiva –, realizada antes de o conteúdo ter sido publicado na plataforma; e a moderação de conteúdo *ex post* – repressiva – que ocorre quando o conteúdo já foi publicado.

Quanto à moderação *ex ante*, ela materializa o que chamei acima de modelo ideal da moderação de conteúdo, porque, por um lado, propõe uma análise de tudo que é veiculado nesse espaço; e, por outro, faz isso antes que publicação venha à luz – e possa, nesse sentido, desencadear o seu potencial de causar problemas. Trata-se de algo próximo ao que Gillespie chama de “*revisão editorial*”<sup>328</sup>, em uma analogia ao que fazem as publicações tradicionais – como jornais e revistas – antes de enviar a versão final à publicação.

Dois exemplos podem ajudar a compreender como isso funciona: na classificação anterior, mencionei o caso dos Atentados de Christchurch, transmitidos ao vivo no Facebook por uma ferramenta da plataforma. Respondendo a comunicações das autoridades, a plataforma começou a retirar os vídeos – tanto a transmissão original quanto suas replicações – e publicou duas mensagens no Twitter, informando essas providências:

---

<sup>325</sup> “*The overwhelming majority of these accounts were detected by our own technology, with just 0.2% of the accounts we suspended in 2017 being flagged by the police*”. PRESS ASSOCIATION. Twitter bans 270,000 accounts for ‘promoting terrorism’. **The Guardian**, 5.4.2018, disponível em <<https://www.theguardian.com/technology/2018/apr/05/twitter-bans-270000-accounts-to-counter-terrorism-advocacy>>, acesso em 22.7.2019.

<sup>326</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.636 e ss.

<sup>327</sup> Cf. GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, pp. 67/69.

<sup>328</sup> “*Editorial Review*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 86 e ss (Kindle, n. p.).



*“Nas primeiras 24 horas nós removemos 1.5 milhões de vídeos do ataque globalmente, dos quais mais de 1.2 milhões foram bloqueados no momento do upload...”*

*“Por respeito às pessoas afetadas por essa tragédia e preocupações das autoridades locais, nós estamos também removendo todas as versões editadas do vídeo que não mostram conteúdo sensível”- Mia Garlich, Facebook da Nova Zelândia.”<sup>329</sup>*

A divisão informada entre os vídeos que foram retirados quando já estavam na plataforma e os que “foram bloqueados no momento do upload” refere-se exatamente à moderação de conteúdo *ex post* e *ex ante*: após o primeiro vídeo do ataque ter sido detectado – *ex post* –, suas características digitais foram inseridas na base da plataforma, de modo que sempre que um conteúdo idêntico tentasse ser enviado, ele era automaticamente bloqueado assim que o usuário acionasse o comando “compartilhar”. Como informa Klonick, é por isso que, “quando um usuário faz o upload de um vídeo no Facebook, uma mensagem aparece: ‘Processando vídeos: O vídeo na sua publicação está sendo processado. Nós vamos mandar uma notificação quando isso estiver pronto e a sua publicação estiver pronta para ser visualizada’<sup>330</sup> – trata-se do momento em que a plataforma compatibiliza o vídeo com um banco de dados de proibições pré-determinadas, detectando, *ex ante*, se o conteúdo veiculado deve ser enviado à moderação de conteúdo.

Um outro exemplo, citado tanto por Klonick<sup>331</sup>, quanto Gillespie<sup>332</sup> é a ferramenta do PhotoDNA, criada pela Microsoft. Ela torna possível que qualquer figura

---

<sup>329</sup> *“In the first 24 hours we removed 1.5 million videos of the attack globally, of which over 1.2 million were blocked at upload...Out of respect for the people affected by this tragedy and the concerns of local authorities, we’re also removing all edited versions of the video that do not show graphic content.”*. FACEBOOK NEWSROOM (Facebook). Sem título. Sem localização, 17 de março de 2019. **Twitter**: fbnewsroom. Disponível em <[https://twitter.com/fbnewsroom/status/1107117981358682112?ref\\_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref\\_url=https%3A%2F%2F](https://twitter.com/fbnewsroom/status/1107117981358682112?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref_url=https%3A%2F%2F)>, acesso em 23.7.2019.

<sup>330</sup> *“When a user uploads a video to Facebook, a message appears: “Processing Videos: The video in your post is being processed. We’ll send you a notification when it’s done and your post is ready to view.”* KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.636.

<sup>331</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.636/1.637 e nota n.º 264.

<sup>332</sup> *“The automatic detection of child pornography images is largely accomplished through a tool called PhotoDNA, introduced by Microsoft in 2009 and since offered for free to major social media platforms and online content providers”* (“A detecção automática de imagens de pornografia infantil é largamente realizada por uma ferramenta chamada PhotoDNA, introduzida pela Microsoft em 2009 e desde então

seja transformada em um código único por uma técnica denominada *hashing*, que traduz pixels em números e cria uma espécie de espécie de “*impressão digital*” – isto é, uma sequência de dados única – da imagem. Dessa forma, é possível criar um banco de dados numérico de imagens proibidas e compará-lo com a que o usuário enviou para publicação na plataforma, detectando-a como uma potencial violação das políticas de uso antes que ela seja veiculada<sup>333</sup>. A principal utilidade dessa ferramenta, atualmente, é o combate à pornografia infantil: o Centro Nacional de Crianças Desaparecidas e Exploradas, uma instituição dos Estados Unidos sem fins lucrativos, mantém, por meio do Projeto VIC<sup>334</sup> um banco de dados de códigos que correspondem a imagens já apreendidas de pornografia infantil. Quando um usuário de plataformas como o Facebook, Twitter e YouTube veicula uma imagem ou um vídeo nessas plataformas, eles são transformados em um código numérico, que é comparado com esse banco de dados. Caso haja alguma coincidência, a publicação não é veiculada e mecanismos internos de comunicação às autoridades são acionados. Um método muito semelhante é utilizado para a proteção aos direitos autorais: o YouTube, por exemplo, desenvolveu uma ferramenta – o ContentID – que cria uma “*impressão digital*” no conteúdo criado por alguém, de modo a identificá-lo sempre que ele é levado à plataforma, permitindo que, se o caso, ele seja bloqueado<sup>335</sup>.

---

*oferecida gratuitamente na maioria das plataformas de mídias sociais e provedores de conteúdo online”*).

GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 108 (Kindle, n. p.).

<sup>333</sup> “*PhotoDNA creates a unique digital signature (known as a “hash”) of an image which is then compared against signatures (hashes) of other photos to find copies of the same image. When matched with a database containing hashes of previously identified illegal images, PhotoDNA is an incredible tool to help detect, disrupt and report the distribution of child exploitation material*” (“*PhotoDNA cria uma assinatura digital única (conhecida como ‘hash’) de uma imagem que é então comparada com assinaturas (hashes) de outras fotos para encontrar cópias da mesma imagem. Quando combinado com uma base de dados com hashes de imagens ilegais anteriormente identificadas. PhotoDNA é uma ferramenta impressionante para ajudar a detectar, interromper e reportar a distribuição de material advindo da exploração de crianças*”). MICROSOFT. **How does PhotoDNA technology work?** Disponível em <<https://www.microsoft.com/en-us/photodna>>, acesso em 23.7.2019.

<sup>334</sup> Cf. PROJECT VIC. **About us**. Disponível em <<https://www.projectvic.org/about-2/>>, acesso em 23.7.2019.

<sup>335</sup> Cf. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.637; GOOGLE. Como funciona o ContentID. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2797370?hl=pt-BR>>, acesso em 22.7.2019; e *infra*, Cap. I.2, 2.4.2.2.

Se a moderação de conteúdo *ex ante* é um meio possível de inviabilizar que as consequências deletérias de se veicular um conteúdo proibido em uma plataforma ocorram, por que, afinal de contas, ela não é usada sempre? Por que existem outros tipos de moderação de conteúdo que não a *ex ante*?

Para responder a essa pergunta – que levará, naturalmente, ao outro oposto dessa segunda classificação, a moderação de conteúdo *ex post* – é preciso retomar o “*problema da escala*”, que reconstruí acima: em síntese, com a popularização da internet e o aumento do número de usuários e de conteúdo que circula nas plataformas, um problema fundamental enfrentado pela moderação de conteúdo é o de estabelecer métodos capazes de analisar volumes crescentes de informações. Nesse sentido, o diagnóstico é de que “*há simplesmente muito conteúdo e atividade para conduzir uma avaliação proativa na qual um moderador examine cada contribuição antes que ela seja divulgada*”<sup>336</sup>. É impossível analisar, previamente, 350 milhões de fotos diariamente, ou 500 milhões de Tweets por dia, ou, ainda, 400 horas de vídeo por minuto<sup>337</sup>, seja porque é simplesmente muito material para conferir – uma questão quantitativa – seja porque analisa-lo raramente é uma atividade tão simples como a de comparar um vídeo ou uma imagem com um registro de um banco de dados, como nos exemplos acima – uma questão qualitativa<sup>338</sup>. Nesse sentido, as situações que exemplifiquei acima – o vídeo do atentado e o PhotoDNA – têm particularidades que não se repetem em todo o conteúdo que é moderado: a existência de uma decisão prévia quanto a um material determinado, que é simplesmente replicada e a

---

<sup>336</sup> “*For instance, there is simply too much content and activity to conduct proactive review, in which a moderator would examine each contribution before it appeared*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 83 (Kindle, n. p.).

<sup>337</sup> Cf. *supra*, notas n.º 307 a 311.

<sup>338</sup> “*But the question of scale is more than just sheer number of users. Social media platforms are not just big; at this scale they become fundamentally different than they once were. They are qualitatively more complex; while these platforms may speak to their online ‘community,’ singular, at two billion active users there simply can be no such thing*” (“*Mas a questão da escala é mais do que apenas o número puro de usuários. Plataformas de mídias sociais não são apenas grandes; nessa escala elas se tornam fundamentalmente diferentes daquilo que um dia foram. Elas são qualitativamente mais complexas; enquanto essas plataformas às vezes falem de sua ‘comunidade’ online no singular, ao nível de dois bilhões de usuários ativos é simplesmente impossível haver algo assim*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 84 (Kindle, n. p.).

possibilidade de se identificar com facilidade e rapidez, por ferramentas tecnológicas, a inserção do conteúdo na plataforma. Se, entretanto, alguém decidir enviar um vídeo de um novo ataque ou uma imagem que não está no banco de dados, ambos não serão detectados por esses mecanismos e novas ferramentas terão que entrar em ação – algo que tanto expõe as limitações da moderação *ex ante* quanto evidencia a necessidade de uma moderação *ex post*. Além disso, é importante perceber que plataformas funcionam a partir de uma dinâmica de interações instantâneas entre as pessoas, algo que é diretamente associado à sua capacidade de gerar lucro<sup>339</sup>: impor barreiras que retirem essa dinamicidade, por exemplo, sujeitando cada manifestação a um análise caso a caso – que não depende, portanto, de um banco de dados prévio como o do Photo ID – esvaziaria um aspecto que é essencial do tipo de serviço que prestam.

A alternativa, nesse sentido, é a moderação *ex post*, isto é, aquela que ocorre quando a publicação já foi veiculada na plataforma. Por esse mecanismo, o usuário faz uma publicação na plataforma que, após ser disponibilizada para os outros usuários, é detectado como algo sujeito à moderação de conteúdo e, por isso, é levada a esses mecanismos decisórios. Os três casos do capítulo anterior são hipóteses de moderação *ex post*: no primeiro, a imagem de Egeland já estava no Facebook quando foi removida; no segundo, a página de Jones – e seus vídeos, a exemplo dos que espalharam a notícia falsa sobre Sandy Hook – já existiam há bastante tempo quando foram excluídas; no terceiro, as publicações que negavam o Holocausto estavam e permaneceram no Facebook nas oportunidades em que as controvérsias emergiram.

Dois outros exemplos podem ajudar a compreender como a moderação *ex post* funciona: em fevereiro de 2019, um grupo de pais reportou, por uma ferramenta interna do YouTube – que detalho abaixo – vídeos a respeito de um jogo de videogame, que escondiam instruções sobre suicídio a crianças<sup>340</sup>; nas eleições de 2016 nos Estados

---

<sup>339</sup> Cf. *supra*, nota n.º 250; uma diminuição em 5 milhões no número de usuários do Twitter no início de 2019 foi associada a uma queda significativo no valor de mercado da companhia; cf. SWENEY, M. Twitter shares fall amid concern over falling user numbers. **The Guardian**, 7.2.2019, disponível em <<https://www.theguardian.com/technology/2019/feb/07/twitter-shares-fall-amid-concern-over-falling-user-numbers>>, acesso em 23.7.2019.

<sup>340</sup> BEVER, L. A pediatrician exposes suicide tips for children hidden in videos on YouTube and YouTube Kids. **The Washington Post**, 24.2.2019, disponível em <[https://www.washingtonpost.com/technology/2019/02/24/pediatrician-exposes-suicide-tips-children-hidden-videos-youtube-youtube-kids/?noredirect=on&utm\\_term=.b9389fe41524](https://www.washingtonpost.com/technology/2019/02/24/pediatrician-exposes-suicide-tips-children-hidden-videos-youtube-youtube-kids/?noredirect=on&utm_term=.b9389fe41524)>.

Unidos, o Twitter detectou e baniou, da plataforma, mais de 200 páginas vinculadas à Rússia por terem disseminado informações falsas e spam em seu ambiente<sup>341</sup>.

A terceira classificação divide a moderação de conteúdo de acordo com quem detecta o conteúdo sujeito à avaliação: se a própria plataforma, que vasculha, *proativamente*; ou, então, se os seus usuários ou outros atores, que o leva à análise decisória, que age *reativamente*. Klonick afirma, quanto à primeira forma, que “*esse é um método em grande parte confinado à moderação de manifestações extremistas e terroristas*”<sup>342</sup>; um exemplo disso são as ações tomadas pelo Facebook em 2016 contra a utilização da plataforma como um modo de disseminação do terrorismo. Pressionado por governos – isto é, por um formato de detecção extraplataforma –, o Facebook estabeleceu uma espécie de força-tarefa contraterrorista, vasculhando ativamente a plataforma contra publicações, por exemplo, feitas pelo grupo radical islâmico Boko Haram<sup>343</sup>. O Twitter, em uma ação parecida, suspendeu, proativamente, em seis meses, 125 mil contas vinculadas ao ISIS em 2016<sup>344</sup>. Algo semelhante tem ocorrido nas vésperas de eleições: após a polêmica envolvendo a influência russa nos Estados Unidos em 2016<sup>345</sup>, o Facebook organiza ações proativas prévias a sufrágios, vasculhando a plataforma para excluir páginas falsas e evitar interferências de outros países nas escolhas dos eleitores. Nesse sentido, uma declaração feita por Zuckerberg no final de 2018:

---

<sup>341</sup> TWITTER PUBLIC POLICY. **Update: Russian interference in the 2016 US presidential election.** 28.9.2017, disponível em <[https://blog.twitter.com/official/en\\_us/topics/company/2017/Update-Russian-Interference-in-2016-Election-Bots-and-Misinformation.html](https://blog.twitter.com/official/en_us/topics/company/2017/Update-Russian-Interference-in-2016-Election-Bots-and-Misinformation.html)>, acesso em 23.7.2019. Cf., também, LECHER, C. Twitter says it banned more than 200 Russia-linked propaganda accounts. **The Verge**, 28.8.2017, disponível em <<https://www.theverge.com/2017/9/28/16381134/twitter-russia-election-ban>>, acesso em 23.7.2019.

<sup>342</sup> “*Currently, this method is largely confined to the moderation of extremist and terrorist speech*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.638.

<sup>343</sup> ANDREWS, N., SEETHARAMAN, D. Facebook Steps Up Efforts Against Terrorism. **The Wall Street Journal**, 11.2.2016, disponível em <<https://www.wsj.com/articles/facebook-steps-up-efforts-against-terrorism-1455237595>>, acesso em 23.7.2019.

<sup>344</sup> BROOMFIELD, M. Twitter shuts down 125,000 Isis-linked accounts. **The Independent**, 6.2.2016, disponível em <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/125000-isis-linked-accounts-suspended-by-twitter-a6857371.html>>, acesso em 23.7.2019.

<sup>345</sup> Cf. *supra*, nota n.º 341.

*“Hoje, o Facebook está mais preparado para esse tipo de ataque. Nós identificamos e removemos contas falsas antes das eleições na França, Alemanha, Alabama, México e Brasil. Nós encontramos e retiramos campanhas de influência estrangeira provenientes da Rússia e do Irã, tentando interferir nos Estados Unidos, Reino Unido, Oriente Médio e em outros locais – assim como grupos no México e no Brasil que têm sido ativos em seus próprios países”<sup>346</sup>.*

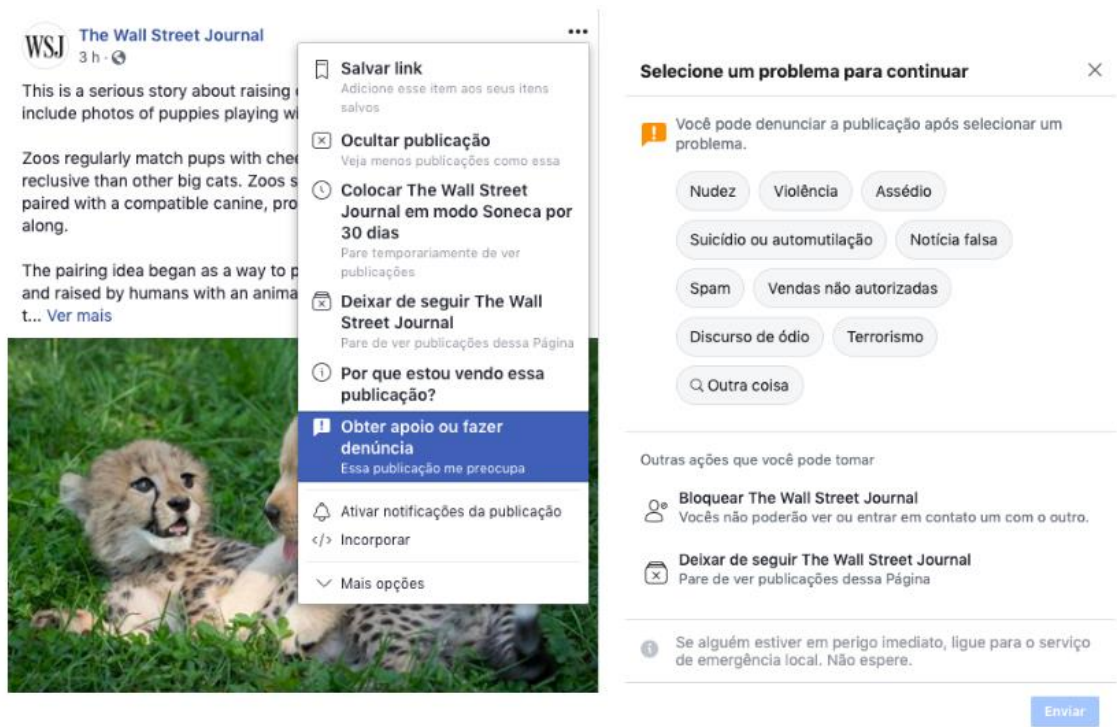
Além de *proativamente*, o conteúdo pode ser moderado *reativamente*, quando aquilo que deve ser submetido à decisão da plataforma é detectado não por ela, mas por seus usuários. Isso acontece, via de regra, por meio de um mecanismo de sinalização [*“flagging”*], pelo qual o usuário da plataforma submete um determinado conteúdo à moderação de conteúdo a partir de formulários específicos, que reportam a ela que, por exemplo, uma determinada publicação viola os padrões daquela comunidade. Essa via tem as suas desvantagens – não existe sinalização *ex ante*, mas apenas *ex post*, o que significa que o conteúdo ofensivo já desencadeou problemas no ambiente da plataforma, no mínimo para o usuário que o reporta –, mas ela lida melhor justamente com o problema da escala, porque estabelece uma forma de cooperação entre usuário e plataforma, aproveitando-se justamente do imenso número de pessoas que passaram a se manifestar nesses ambientes para melhor controlá-los.

Via de regra, esse mecanismo funciona inserindo, em cada manifestação feita na plataforma, uma ferramenta pela qual o usuário pode selecioná-la e a encaminhar para a moderação de conteúdo. No caso do Facebook, trata-se de uma caixa de seleção localizada na parte superior direita de cada publicação, que remete o usuário para um formulário pelo qual ele pode reportar a publicação, detalhando a violação:

---

<sup>346</sup> “Today, Facebook is better prepared for these kinds of attacks. We've identified and removed fake accounts ahead of elections in France, Germany, Alabama, Mexico, and Brazil. We've found and taken down foreign influence campaigns from Russia and Iran attempting to interfere in the US, UK, Middle East, and elsewhere -- as well as groups in Mexico and Brazil that have been active in their own country”. ZUCKERBERG, M. **Preparing for Elections**. 13.9.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/preparing-for-elections/10156300047606634/>>, acesso em 23.7.2019.

**Figura 15 – Caixa de seleção e formulário para se reportar uma publicação no Facebook**

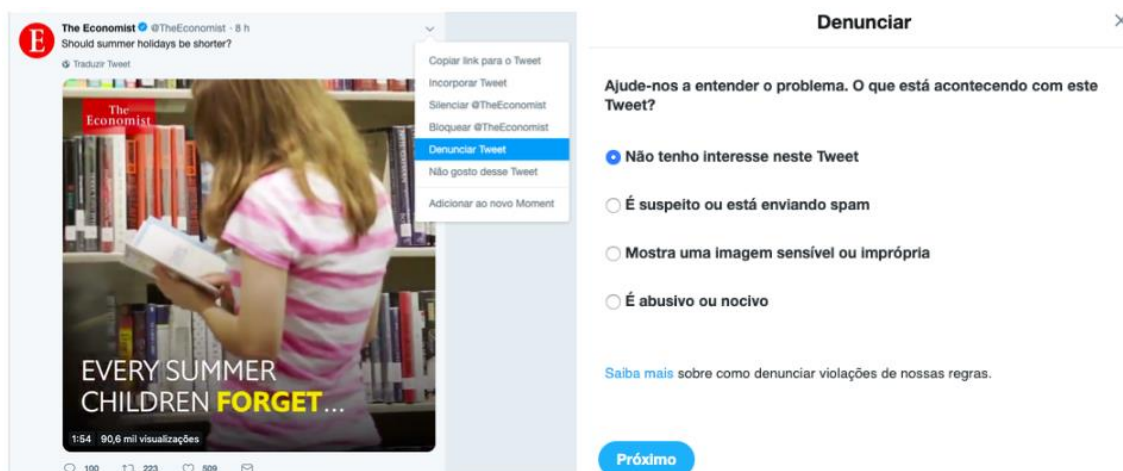


Fonte: Facebook<sup>347</sup>

Esse mesmo mecanismo está disponível tanto no Twitter quanto no YouTube:

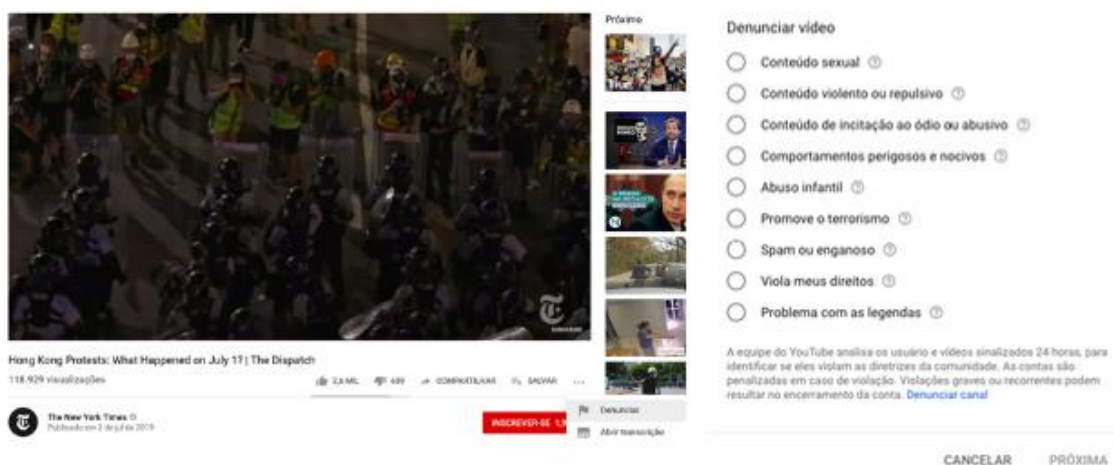
<sup>347</sup>A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“screenshot”); para tanto, acessei, como usuário, uma conta do Facebook. A figura é de uma publicação aleatória escolhida da página do *The Wall Street Journal* <<https://www.facebook.com/wsj/>> na plataforma, em 7.7.2019.

**Figura 16 – Caixa de seleção e formulário para se reportar uma publicação no Twitter**



Fonte: Twitter<sup>348</sup>

**Figura 17 – Caixa de seleção e formulário para se reportar uma publicação no YouTube**



Fonte: YouTube<sup>349</sup>

<sup>348</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“screenshot”); para tanto, acessei, como usuário, uma conta do Twitter. A figura é de uma publicação aleatória escolhida da página do *The Economist* <<https://twitter.com/TheEconomist/status/1147890839630352388/>> na plataforma, em 7.7.2019.

<sup>349</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“screenshot”); para tanto, acessei, como usuário, uma conta do YouTube. A figura é de uma publicação aleatória escolhida da página do *The New York Times* <<https://www.youtube.com/watch?v=koTdqWz77AU>> na plataforma, em 7.7.2019.



Como os exemplos acima expõem, a moderação reativa tem ainda outra vantagem: as ferramentas que viabilizam aos usuários reportarem conteúdo podem ser configuradas para que direcionem a publicação para providências específicas, a depender do problema detectado – predispondo, assim, a moderação de conteúdo. Gillespie relata, por exemplo, que no YouTube “*algumas categorias (por exemplo, ‘conteúdo sexual: conteúdo envolvendo menores’) são revisados imediatamente, porque as plataformas têm obrigações estritas de reportar pornografia infantil*”, sendo que “*o resto é colocado em uma fila, em ordem de importância: vídeos sinalizados como ‘conteúdo sexual: nudez sexual explícita’ podem ser priorizados sobre os sinalizados como ‘atos perigosos e danosos: abuso de fármacos ou de drogas’*”<sup>350</sup>.

Evidentemente, a moderação reativa não soluciona todas as dificuldades das plataformas: o problema de escala, por exemplo, toma uma proporção menor que a colocada, por exemplo, na moderação *ex ante*, mas mesmo assim continua presente. Em 2016, os usuários do Facebook reportaram um milhão de publicações para que a plataforma os analisasse – por dia<sup>351</sup>. Isso significa que o problema de escala que se apresentava na detecção acaba sendo levado à decisão, o que, como exponho na seção seguinte, estimulou a construção de soluções que automatizassem a análise do conteúdo denunciado, em uma tentativa de minimizar a ação humana – naturalmente mais lenta e cara.

Ainda, ao dividir o trabalho da moderação com os usuários, as plataformas abrem caminho para que a detecção de conteúdo seja manipulada, como foi em 2014: na época, o Facebook começou a receber sinalizações de que perfis de *drag queens* estariam violando as suas políticas de conteúdo, porque os nomes utilizados por eles para se apresentarem no espaço não seriam os “*reais*”, mas sim aqueles utilizados em performances artísticas. Consequentemente, centenas de perfis foram suspensos pela

---

<sup>350</sup> “*For instance, some categories (for example, ‘sexual content: content involving minors’) are reviewed immediately, as platforms have strict legal obligations to report child pornography. The rest can be queued up in order of importance: videos flagged as ‘sexual content: graphic sexual nudity’ might be prioritized over those flagged as ‘harmful dangerous acts: pharmaceutical or drug abuse, as YouTube sees fit’*”. GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 98 (Kindle, n. p.).

<sup>351</sup> BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

plataforma, desencadeando reclamações da comunidade LGBTQ+, que organizou uma marcha até a sede do Facebook para dizer que “a política [de nomes reais] é discriminatória contra pessoas que usam nomes alternativos”, a exemplo de “*drag queens e trabalhadores do mercado do sexo*” ou mesmo “*sobreviventes de abusos sexuais que podem fazê-lo por razões de segurança*”<sup>352</sup>. Algum tempo depois, o Facebook reestabeleceu os perfis e fez uma publicação se desculpando e informando que a política da plataforma nunca foi a de proibir a criação de perfis com esses nomes alternativos<sup>353</sup>. De acordo com a executiva que publicou a mensagem, a plataforma foi induzida a erro por um único usuário, que resolveu sinalizar centenas de perfis de *drag queens*, notadamente da região de São Francisco. Uma investigação independente descobriu, algum tempo depois, que o responsável havia se baseado em motivações religiosas para fazê-lo de forma intencional<sup>354</sup>.

A quarta e última classificação diz respeito ao modo como o material veiculado na plataforma é detectado e levado à moderação de conteúdo: se por meio da ação humana, ou se por mecanismos automatizados. Primeiro, é preciso compreender o que se quer dizer com essa distinção. É evidente que, em última análise, a detecção de qualquer conteúdo no sentido de leva-lo à análise da moderação de conteúdo é um ato humano. Mesmo técnicas automatizadas que detêm mecanismos de autoaprendizagem – isto é, capazes de se adequar a novas situações mesmo sem a intervenção humana – foram, um dia, programadas para agir daquela forma pela ação humana. Mesmo assim, há diferenças entre por um lado, o conteúdo que é detectado diretamente por um ser humano, que vasculha a plataforma buscando-o; e uma máquina, previamente programada por alguém – e esclarecê-las é importante, porque elas indicam as vantagens e as limitações de cada um desses mecanismos.

---

<sup>352</sup> “[...] *the policy discriminates against people who use alternative names drag queens and sex workers do it because that's how they self-identify*”. KELLY, H. Drag queens lead "real names" protest at Facebook. **CNN Business**, 1.6.2015, disponível em <<https://money.cnn.com/2015/06/01/technology/facebook-protest-names/index.html>>, acesso em 23.7.2019.

<sup>353</sup> COX, C. (Chris Cox). Sem título. Menlo Park, 1 de outubro de 2014. **Facebook**: chris.cox. Disponível em <<https://www.facebook.com/chris.cox/posts/10101301777354543>>, acesso em 23.7.2019.

<sup>354</sup> HATMAKER, T. RealNamePolicy and the real story behind Facebook's name policy tumble. **The Daily Dot**, 3.10.2014, disponível em <<https://www.dailydot.com/debug/realnamepolice-facebook-real-names-policy/>>, acesso em 23.7.2019.

A detecção humana se traduz na ação de um moderador que, por exemplo, após receber uma notificação externa, informando que um grupo terrorista está agindo em uma determinada página da plataforma, acessa-a e busca identificar se há alguma publicação suspeita. Nesse sentido – e por guardar certas semelhanças com a atividade do censor estatal – há pouco para se destacar a esse respeito, além das evidentes limitações dessa metodologia: dedicar seres humanos a fazer isso é um processo demorado e custoso, em específico tendo-se em conta a dimensão da quantidade de conteúdo veiculado nessas plataformas na última década – no YouTube, por exemplo, a última estatística disponível indica que, diariamente, os usuários assistem mais de um bilhão de horas em vídeos<sup>355</sup>. Isso ainda é útil, entretanto, em específico nos casos em que se sabe exatamente o que se está procurando: no caso de reclamações extraplataforma, em que a informação sobre o que deve ser moderado é fornecida com alguma precisão – por exemplo, quando o Facebook recebeu reclamações de vítimas sobre a publicação de Alex Jones a respeito do atentado à escola Sandy Hook<sup>356</sup> –, a detecção manual funciona melhor, porque ensinar uma máquina a buscar a informação pode levar mais tempo do que acessá-la.

Por sua vez, a detecção automática funciona a partir de *softwares* que, agindo *ex ante* ou *ex post*, são programados para identificar determinados conteúdos, como nudez e palavras agressivas, e submetê-los à moderação de conteúdo. A técnica tem vantagens evidentes em relação à detecção humana, uma vez que ela oferece instrumentos adequados para a solução do problema de escala, como explorei com mais detalhes acima, na detecção *ex ante* de pornografia infantil pela ferramenta do PhotoDNA. Em síntese, moderar conteúdo com mecanismos automáticos de detecção viabiliza atuar em uma quantidade maior de conteúdo, fazendo-o de forma mais rápida e, em casos como o Facebook, Twitter e YouTube, que lidam com proporções imensas de publicações de seus usuários, barata: “*Dá mais trabalho dizer a um computador o que ele deve fazer do que dizê-lo a um humano, mas uma vez que o trabalho de programação está feito, é barato utilizá-lo em milhares ou milhões de casos*”<sup>357</sup>. Além disso, como aponta Gillespie, a

---

<sup>355</sup> YOUTUBE. **YouTube in Numbers**. Disponível em <<https://www.youtube.com/intl/en-GB/yt/about/press/>>, acesso em 23.7.2019.

<sup>356</sup> Cf. *supra*, nota n.º 129.

<sup>357</sup> “*It takes more work to tell a computer what to do than to tell a human, but once the programming work is done, it is cheap to use it in thousands or millions of cases*”. GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, p. 65.

automatização traz consigo a promessa do fim da subjetividade: diferente da atuação humana, a decisão automatizada seria sempre objetiva e, nesse sentido, coerente<sup>358</sup>.

A detecção automatizada do que as pessoas veiculam nas plataformas, entretanto, não é a solução final dos problemas da moderação de conteúdo, mas, assim como todas as modalidades que expus acima, uma ferramenta que se mostra mais ou menos adequada a depender da situação. Ela exige, mesmo que em um nível inicial, uma inserção básica de dados que lhe informem o que é inapropriado, o que a torna inefetiva em casos novos – a exemplo do primeiro vídeo veiculado no Facebook que veiculava os Atentados de Christchurch foram transmitidos. Além disso – e como discuto com mais detalhes na seção seguinte –, as particularidades das formas de expressão humana frequentemente não se deixam traduzir em um conjunto fechado de instruções que possa ser dado a essas ferramentas e, por isso, situações que dificilmente gerariam controvérsias na moderação manual geram falsos positivos na automatizada, ocasionando erros evidentes na detecção de conteúdo.

Um exemplo pode ajudar a compreender melhor isso: como expliquei no terceiro caso que reconstruí no capítulo anterior, o YouTube modificou a sua política de conteúdo quanto a vídeos que negam o Holocausto ou o nazismo, passando a classifica-los como uma forma de discurso de ódio vedado na plataforma<sup>359</sup>. Assim, a plataforma anunciou em uma publicação oficial que iria remover “*vídeos que promovam ou glorifiquem a ideologia nazista*” ou que “*negam que eventos violentos documentados, como o Holocausto e o tiroteio na Escola Básica Sandy Hook, ocorreram*”, de modo que começaria a “*aplicar essa política atualizada hoje [5.6.2019]; entretanto, ainda levará tempo até os nossos sistemas se atualizarem e nós estaremos gradualmente expandir a sua cobertura nos próximos meses*”<sup>360</sup>. No dia seguinte, entretanto, os canais de professores de

---

<sup>358</sup> “*In addition, these tools promise to solve the problem of subjectivity: what if automated mechanisms could identify and remove content fairly, without human bias, in any language?*” (“*Adicionalmente, essas ferramentas prometem resolver o problema da subjetividade: e se mecanismos automáticos pudessem identificar e remover conteúdo de forma justa, sem o enviesamento humano, em nenhuma linguagem?*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 105 (Kindle, n. p.).

<sup>359</sup> Cf. *supra*, nota n.º 160.

<sup>360</sup> “[...] *videos that promote or glorify Nazi ideology, we will remove content denying that well-documented violent events, like the Holocaust or the shooting at Sandy Hook Elementary, took place*”. [...] “*We will begin enforcing this updated policy today; however, it will take time for our systems to fully ramp up and we’ll be gradually expanding coverage over the next several months*”. YOUTUBE. Our ongoing work to

história que veiculavam na plataforma vídeos educativos sobre o nazismo – como documentários antigos da BBC – foram retirados do ar, sob o argumento de que eles promoviam o discurso de ódio. Após os proprietários das contas reclamarem, os canais foram reestabelecidos e um porta-voz do YouTube veio a público dizer que “*a companhia usa uma combinação de tecnologia e pessoas para aplicar os seus padrões*”<sup>361</sup> – um indicativo de que, possivelmente, as contas foram retiradas por mecanismos automatizados que detectaram ali referências ao nazismo, sem serem capazes, entretanto, de diferenciar o material a ser removido de documentários de caráter educativo, uma vez que ambos têm alguns conteúdos semelhantes, como imagens de suásticas.

O quadro abaixo articula as classificações que propus nesta seção quanto às formas pelas quais as plataformas detectam o conteúdo publicado – ou em vias de ser publicado – por seus usuários, sempre sob o pressuposto de que essa é uma taxonomia não-exaustiva, que se concentra em expor as metodologias apontadas de forma mais comum pela literatura.

**Quadro 7 – Fase n.º 1: A detecção de conteúdo no procedimento adjudicatório**

<u>Critério</u>	<u>Classificações</u>	<u>Exemplos</u>
<b><u>Origem</u></b> De onde vem a provocação para que a plataforma modere?	<u>Extraplataforma:</u> de mecanismos localizados fora da plataforma.	Caso Alex Jones (Cap. 1, 3.2.2.2); Negação do Holocausto (Cap. 1, 3.2.2.3); Atentados de Christchurch.
	<u>Intraplataforma:</u> de mecanismos localizados dentro da plataforma.	Caso Lipman; Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); Uso do PhotoDNA.
<b><u>Momento</u></b> Quando o	<u>Ex ante / preventiva:</u> antes de ele ser publicado.	Atentados de Christchurch; Uso do PhotoDNA.

tackle hate. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>>, acesso em 22.7.2019.

<sup>361</sup> “[...] *the company used a combination of technology and people to enforce the guidelines*” WATERSON, J. YouTube blocks history teachers uploading archive videos of Hitler. **The Guardian**, 6.6.2019, disponível em < <https://www.theguardian.com/technology/2019/jun/06/youtube-blocks-history-teachers-uploading-archive-videos-of-hitler>>, acesso em 23.7.2019.

conteúdo é detectado?	<i>Ex post / repressiva</i> : após ele ser publicado.	Caso Lipman; Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); Caso Alex Jones (Cap. 1, 3.2.2.2); Negação do Holocausto (Cap. 1, 3.2.2.3); Instruções de suicídio detectadas por pais em vídeos do YouTube; disseminação de notícias falsas por contas no Twitter provenientes da Rússia.
<b><u>Tipo de agente</u></b>  Quem detecta o conteúdo?	<i>Proativa</i> : a plataforma vasculha o ambiente que administra.	Manifestações extremistas e terroristas, como do Boko Haram e do ISIS; ações prévias a sufrágios; Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); Caso Alex Jones (Cap. 1, 3.2.2.2); Negação do Holocausto (Cap. 1, 3.2.2.3).
	<i>Reativa</i> : outros participantes da comunidade reportam algo que lhes causa desconforto.	Caso Lipman; Casos de <i>flagging</i> em geral; caso em que se discutiu o nome utilizado por <i>drag queens</i> .
<b><u>Modo</u></b>  Como o conteúdo é levado à moderação?	<i>Ação humana</i> : um moderador humano busca o conteúdo.	Caso Lipman’ Caso Alex Jones (Cap. 1, 3.2.2.2).
	<i>Ação automatizada</i> : um <i>software</i> busca o conteúdo.	Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); PhotoDNA; Negação do Holocausto (Cap. 1, 3.2.2.3); buscas feitas pelo YouTube após mudar sua política sobre a negação ao Holocausto.

#### 2.4.2.2 Fase n.º 2: Decidir

Uma vez detectado que um determinado conteúdo na plataforma deve ser submetido à moderação de conteúdo porque há a suspeita de que ele viole alguma regra da plataforma, o passo seguinte é tomar uma decisão sobre isso. Aqui, a pergunta é a seguinte: como as plataformas definem que o conteúdo *x*, veiculado pelo usuário *y*, viola a regra *z* da plataforma? Há duas formas principais de classificar essa fase: de acordo com o modo como a decisão é organizada e conforme quem toma a decisão.

A primeira classificação adota como critério o modo como essa tomada de decisão é organizada: se de forma centralizada ou distribuída ao longo da plataforma<sup>362</sup>. Em síntese, os mecanismos aos quais fiz referência nos exemplos ao longo deste trabalho traduzem uma moderação de conteúdo centralizada, porque eles refletem decisões que se aplicam a todo o conteúdo e a todo o espaço das plataformas da internet. Assim, a nudez é proibida em todo o Facebook, da mesma forma como o discurso de ódio é vedado e páginas que negam o Holocausto são permitidas – algo que reflete orientações centralizadas pela plataforma e que são por ela aplicadas, também de forma centralizada, via moderação de conteúdo. Centralização, entretanto, não significa impossibilidade de se delegar poderes: a profissionalização da moderação de conteúdo, a que fiz referência acima, por exemplo, para explicar uma maior variabilidade normativa, também significou o escalonamento dos polos decisórios internos das plataformas. Um exemplo disso é o Facebook: na moderação *ex post*, manual e proativa, o conteúdo sinalizado pelos usuários da plataforma é encaminhado para uma equipe de moderadores, que é distribuído em três níveis, cada qual com funções específicas, como exponho com mais detalhes a seguir<sup>363</sup>.

No lado oposto à moderação de conteúdo centralizada, há a que ocorre de forma distribuída ao longo da plataforma: nesse caso, a companhia por trás desse ambiente confere, dentro de certos limites, poderes aos seus usuários para que eles mesmos tomem as decisões sobre o conteúdo uns dos outros, descentralizando-as. Esse é um modelo adotado com limitações em plataformas extensas, como o Facebook, o Twitter e o YouTube, mas que se faz presente, por exemplo, em comunidades que emergem dentro desses ambientes, nas quais se atribui, a algumas pessoas, poderes semelhantes aos da moderação de conteúdo<sup>364</sup>. Isso é feito conferindo prerrogativas a determinados usuários-

---

<sup>362</sup> “Moderation decisions can be made either centrally by a single moderator whose decision affects the entire community, or by multiple distributed moderators whose individual decisions affect only part of the community” (“Decisões de moderação podem ser feitas ou de forma central, por um moderador sozinho, cuja decisão afeta a comunidade inteira, ou por múltiplos moderadores distribuídos cuja decisão individual afeta apenas parte da comunidade”). GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, p. 69.

<sup>363</sup> Cf. *infra*, Cap. I.2, 2.4.2.3.

<sup>364</sup> “Platforms that are designed to allow smaller subgroups to exist within the broader population of users can empower moderators to be responsible for the activity of a subgroup” (“Plataformas que são construídas para permitir que subgrupos menores existam dentro de uma população mais ampla de usuários podem empoderar moderadores para serem responsáveis pela atividade do subgrupo”). Gillespie, T. **Custodians of**

membros dessas comunidades – frequentemente denominados moderadores –, que têm acesso a funcionalidades específicas, como, por exemplo, a de excluir alguém ou de impedir que um determinado membro faça publicações ali por um certo tempo. O Facebook, por exemplo, recomenda essas providências contra pessoas que “*publicam spam continuamente em sua página*”<sup>365</sup>. Outro exemplo é a possibilidade que os usuários dessas plataformas têm de bloquear determinados usuários de interagirem com eles: no caso do Twitter, é possível que um usuário impeça outro de visualizar as informações da sua página por um recurso desse tipo – o qual, inclusive, foi objeto de judicialização, após o Presidente dos Estados Unidos utilizá-lo contra cidadãos americanos<sup>366</sup>.

O quadro abaixo, adaptado a partir da Central de Ajuda do Facebook, mostra os poderes que a plataforma confere a membros de páginas da plataforma – no caso, o seu administrador e o seu moderador –, descentralizando decisões, por exemplo, que digam respeito a “*remover ou banir pessoas da página*” ou “*responder e excluir comentários e publicações na página*”:

**Quadro 8 – Descentralização de poderes no Facebook**

	<b>Administrador</b>	<b>Moderador</b>
<b>Gerenciar as funções e configurações da Página</b>	✓	
<b>Editar a Página e adicionar aplicativos</b>	✓	✓
<b>Criar e excluir publicações em nome da Página</b>	✓	✓

---

**the Internet: Platforms, content moderation and the hidden decisions that shape social media.** New Haven: Yale University Press, 2018, pp. 132 e ss (Kindle, n. p.).

<sup>365</sup> FACEBOOK. Como faço para banir ou cancelar o banimento de alguém da minha Página? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/289207354498410?helpref=about\\_content](https://www.facebook.com/help/289207354498410?helpref=about_content)>, acesso em 23.7.2019.

<sup>366</sup> Cf. Knight First Amendment v. Trump, n.º 1:17-cv-05205; um resumo do andamento processual e dos documentos pode ser encontrado em KNIGHT FIRST AMENDMENT INSTITUTE. **Knight Institute v. Trump - Lawsuit Challenging President Trump's Blocking of Critics on Twitter**. Disponível em <<https://knightcolumbia.org/content/knight-institute-v-trump-lawsuit-challenging-president-trumps-blocking-critics-twitter>>, acesso em 23.7.2019.



<b>Enviar mensagens como a Página</b>	✓	✓
<b>Responder e excluir comentários e publicações na Página</b>	✓	✓
<b>Remover ou banir pessoas da Página</b>	✓	✓

Fonte: Facebook<sup>367</sup>

A segunda classificação toma como critério quem toma essas decisões. Em síntese, elas podem ser tomadas manualmente, isto é, diretamente por um moderador humano; ou, então, de modo automatizado, por um programa estruturado para fazê-lo. Como afirma Grimmelmann, na ordem inversa, “*uma política contra linguagem pouco apropriada pode ser implementada ou por um software-filtro, que bloqueia os sete palavrões, ou por um censor que lê tudo e decide o que ultrapassa e o que não ultrapassa a linha*”<sup>368</sup>.

De início, é importante não confundir o modo como o conteúdo é detectado da forma pela qual decisões a respeito dele são tomadas. Mecanismos manuais e automatizados podem ser utilizados para detectar aquilo que será submetido à moderação de conteúdo da plataforma, mas isso não implica que uma ferramenta do mesmo tipo será mobilizada para tomar uma decisão a esse respeito. Assim, é possível que uma publicação seja detectada automaticamente e decidida manualmente – e a maioria dos casos ocorre dessa forma, como explico abaixo; detectada manualmente e decidida automaticamente – como no caso em que o usuário reporta uma violação de direitos autorais, que é verificada por um software específico; detectada e decidida manualmente – algo que provavelmente ocorreu no primeiro caso que reconstruí no capítulo anterior, da imagem de Tom Egeland;

<sup>367</sup> FACEBOOK. Quais são as diversas funções das Páginas e o que elas podem fazer? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/289207354498410?helpref=about\\_content](https://www.facebook.com/help/289207354498410?helpref=about_content)>, acesso em 23.7.2019.

<sup>368</sup> “[...] a policy against foul language could be implemented either through a software filter that blocks the seven dirty words or by a censor who reads everything and decides what does and does not cross the line”. GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, 2015, pp. 63/64. O termo “seven words” é uma referência a um rol de sete palavras popularmente consideradas chulas no idioma inglês.

e, por fim, detectada e decidida automaticamente – como no exemplo do PhotoDNA, da seção anterior.

Quanto às decisões tomadas manualmente, elas são o resultado de análises realizadas por moderadores de conteúdos que trabalham tanto em centrais distribuídas pelo mundo, quanto na sede das plataformas, analisando o que lhes é reportado e julgando se há ou não uma violação às políticas de conteúdo. Ocorre que, como afirma Klonick, “*depois que o conteúdo é sinalizado para a plataforma avalia-lo, a mecânica precisa do processo de formação da decisão se torna obscura*”<sup>369</sup> – algo que não é propriamente novidade em se tratando da moderação de conteúdo, um procedimento que, como tenho insistido, tem a opacidade como característica – e problema – marcante. Entretanto, seja a partir de entrevistas com pessoas envolvidas no processo, seja por reportagens e vazamentos de informações, é possível depreender as linhas gerais desse procedimento, que passo a descrever a seguir.

De início, é importante compreender que o trabalho da moderação manual de conteúdo ocorre de forma heterogênea: há níveis que escalonam atores de diferentes capacitações e com diferentes poderes, em uma estrutura profissionalizada vertical. Como relata Klonick, no Facebook, por exemplo, há pelo menos três níveis de moderadores:

*“No Facebook, há três níveis básicos de moderadores de conteúdo: moderadores ‘Nível 1’, que fazem a maioria da avaliação de conteúdo do dia-a-dia; moderadores ‘Nível 2’, que supervisionam os de Nível 3 e avaliam conteúdo prioritário ou a eles escalado; e moderadores de Nível 1, que são tipicamente advogados e aqueles que fazem as políticas, baseados na sede da empresa”*<sup>370</sup>.

Esse retrato da organização do Facebook expõe que, dentro da moderação de conteúdo manual, há uma clivagem entre o pessoal mais qualificado, identificado na organização do Facebook sob nível 1 e que, a exemplo dos advogados de Ammori, que

---

<sup>369</sup> “*After content has been flagged to a platform for review, the precise mechanics of the decision-making process become murky*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.639.

<sup>370</sup> “*At Facebook, there are three basic tiers of content moderators: ‘Tier 3’ moderators, who do the majority of the day-to-day reviewing of content; ‘Tier 2’ moderators, who supervise Tier 3 moderators and review prioritized or escalated content; and ‘Tier 1’ moderators, who are typically lawyers or policymakers based at company headquarters.*” KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.639/1.640.

mencionei na introdução<sup>371</sup>, é responsável por atuar na constituição das políticas e em casos muito delicados; e os responsáveis pela moderação de conteúdo do dia-a-dia, que, identificados sob os níveis 2 e 3 no Facebook, lidam com a maioria do conteúdo a partir de diretrizes provenientes dos estratos superiores.

Quanto ao estrato superior, ele é constituído, via de regra, por pessoas que têm vinculação direta com as companhias por trás das plataformas. São indivíduos com um grau sofisticado de formação acadêmica, provenientes das melhores faculdades dos Estados Unidos e que atuam na intersecção entre direito e tecnologia, tanto constituindo quanto decidindo, de forma vinculante, aquilo que os usuários das plataformas poderão ou não dizer naqueles ambientes. Como a análise de Ammori revela<sup>372</sup>, são advogados que perceberam que o futuro da advocacia voltada à liberdade de expressão não estava mais na atuação judicial em favor da mídia tradicional, mas sim na constituição de políticas de conteúdo em empresas de tecnologia, como o Facebook, Google e Twitter. Um exemplo citado com frequência é o de Nicole Wong<sup>373</sup>: Wong é formada em direito pela Universidade da Califórnia, Berkeley, com um mestrado em jornalismo, o que lhe garantiu posições de destaque em escritórios voltados à defesa, em juízo, de jornais e estações de rádio – área conhecida como “*First Amendment law*”, nos Estados Unidos<sup>374</sup>. Com o

---

<sup>371</sup> Cf. *supra*, Introdução.

<sup>372</sup> AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, p. 2.262.

<sup>373</sup> Cf., para referências a respeito da executiva, KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.618; AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.260/2.261; ROSEN, J. The Deciders: The Future of Privacy and Free Speech in the Age of Facebook and Google. **Fordham Law Review**, v. 80 n. 4, 2012, pp. 1.536 e ss; JOHNSON, E. Full Q&A: Former Google lawyer and deputy U.S. CTO Nicole Wong on Recode Decode. **Vox**, 12.9.2018, disponível em <<https://www.vox.com/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript>>, acesso em 24.7.2019.

<sup>374</sup> “*But I ended up doing First Amendment law when I came out of school. So, the first part of my career was representing newspapers, TV stations, radio stations in the Bay Area on First Amendment issues, which was fabulous*” (“*Mas eu acabei praticando direito da “Primeira Emenda” quando eu sai da faculdade. Então, a primeira parte da minha carreira foi representando jornais, estações de TV, estações de radio na área da Baía [de São Francisco] em questões da Primeira Emenda, o que foi fabuloso*”). JOHNSON, E. Full Q&A: Former Google lawyer and deputy U.S. CTO Nicole Wong on Recode Decode. **Vox**, 12.9.2018, disponível em <<https://www.vox.com/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript>>, acesso em 24.7.2019 (n. p.).

surgimento da internet e o crescimento das empresas de tecnologia na área de São Francisco, Wong aceitou uma posição no Google; pouco tempo depois, a empresa comprou o YouTube “e Wong foi colocada como responsável por criar e implementar políticas de moderação de conteúdo”<sup>375</sup>. Ela executou um projeto de profissionalização da área – em específico pela implantação dos padrões de comunidade que referi na seção anterior –, mas casos mais delicados – como, em 2006, quando vídeos da morte de Saddam Hussein foram publicados na plataforma – eram escalados até ela, que decidia se o conteúdo permanecia ou não na plataforma, algo que lhe rendeu o apelido de “*The Decider*” – aquela que decide.

Esse processo decisório é, portanto, localizado na cúpula da moderação de conteúdo, e concentrado em pessoas tal como Wong. Há pouca informação sobre como isso exatamente ocorre, mas um relato da época em que ela atuava como chefe da moderação de conteúdo do YouTube esclarece como existe uma tentativa, por parte dessas pessoas, de balancear a liberdade de expressão com outros valores importantes, como o respeito a culturas locais. Não há notícia de que exista uma padronização a esse respeito; como o excerto abaixo indica, as análises são casuísticas. O relato foi fornecido por Wong a Kate Klonick, em uma entrevista telefônica:

*[...] em novembro de 2006, o governo da Tailândia anunciou que bloquearia o YouTube para qualquer pessoa que utilizasse um IP tailandês a não ser que o Google removesse vinte vídeos ofensivos do site. Inobstante alguns dos vídeos ‘claramente violarem os termos de serviço do YouTube’, outros simplesmente apresentavam imagens com photoshop do Rei da Tailândia com o pé na cabeça. Na Tailândia, insultar o Rei era ilegal e punível com até 15 anos na prisão. Nicole Wong foi fortemente pressionada para considerar o conteúdo ofensivo. ‘Meu primeiro instinto foi o de que era um cartum. É um photoshop estúpido’, ela disse, ‘mas então, de repente, esse se tornou um momento de aprendizado para mim sobre padrões internacionais de expressão versus padrões de fala da Primeira Emenda e havia muito mais exceções à Primeira Emenda dos Estados Unidos [naquele espaço] que antes’. Wong viajou à Tailândia para resolver a disputa e ficou impressionada pelo amor popular que ela observou na população tailandesa por seu Rei. ‘Você não pode nem imaginar [o amor deles pelo Rei]’, ela relembra da viagem:*

*Toda segunda-feira, literalmente oitenta e cinco por cento das pessoas iam trabalhar com uma camisa dourada ou amarela e vestido e há uma razão histórica para isso: a única*

---

<sup>375</sup> “Wong was put in charge of creating and implementing content-moderation policies”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.619.

*fonte de estabilidade nesse país é o Rei... Eles absolutamente reverenciam o Rei... Alguém na Embaixada dos Estados Unidos o descreveu como uma 'mistura de George Washington, Jesus e Elvis'. Algumas pessoas... lágrimas vieram aos seus olhos quando elas falavam sobre insultos ao Rei e como isso as havia ofendido. Essa foi a parte que me fez recuar. Quem sou eu, uma advogada nos Estados Unidos sentada na Califórnia, para dizer a eles: 'Não, nós não vamos retirar isso. Você vai ter que viver com isso'.*

*Depois da viagem, Wong e seus colegas concordaram em remover os vídeos nos limites geográficos da Tailândia, com a exceção de críticas feitas aos militares”<sup>376</sup>.*

No estrato inferior – composto, no caso do Facebook, pelos Níveis 2 e 3 –, no qual a maioria do conteúdo analisado manualmente é moderado, a realidade é muito diferente. Ao contrário do trabalho artesanal e realizado por pessoas qualificadas no Nível 1, aqui o problema da escala da moderação de conteúdo, que referi acima, traduz-se na necessidade de que as plataformas construam estruturas capazes de analisar, diariamente, milhões de publicações feitas nos seus ambientes e reportadas por sistemas automatizados ou por seus usuários<sup>377</sup>. No que segue, descrevo como essa estrutura funciona, fazendo-o a partir de informações obtidas por vazamentos internos e reportagens que a investigaram a fundo.

---

<sup>376</sup> “[...] in November 2006, the Thai government announced that it would block YouTube to anyone using a Thai IP address unless Google removed twenty offensive videos from the site. While some of the videos “clearly violated the YouTube terms of service,” others simply featured Photoshopped images of the King of Thailand with feet on his head. In Thailand, insulting the King was illegal and punishable by as much as fifteen years in prison. I Nicole Wong was hard pressed to find the content offensive. “My first instinct was it’s a cartoon. It’s a stupid Photoshop,” she stated, “but then it suddenly became a kind of learning moment for me about international speech standards versus First Amendment speech standards and there was a lot more American First Amendment exceptionalism [in that space] than previously.” Wong traveled to Thailand to resolve the dispute and was overwhelmed by the popular love she observed in the Thai people for their King. “You can’t even imagine [their love for their King],” she recounted of the trip: Every Monday literally eighty-five percent of the people show up to work in a gold or yellow shirt and dress and there’s a historical reason for it: the only source of stability in this country is this King... They absolutely revere their King... Someone at the U.S. Embassy described him as a “blend of George Washington, Jesus, and Elvis.” Some people... tears came to their eyes as they talked about the insults to the King and how much it offended them. That’s the part that set me back. Who am I, a U.S. attorney sitting in California to tell them: “No, we’re not taking that down. You’re going to have to live with that.” After the trip, Wong and her colleagues agreed to remove the videos within the geographical boundaries of Thailand, with the exception of critiques of the military.” KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.623.

<sup>377</sup> Cf. *supra*, nota n.º 307.

Em síntese, a moderação de conteúdo dos níveis 2 e 3 – e seus correspondentes nas outras plataformas – é realizada, via de regra, em postos de trabalhos distribuídos pelo mundo e que se assemelham a *callcenters* de telefonia<sup>378</sup>. São estruturas constituídas por empresas terceirizadas que são contratadas pelas plataformas para moderar conteúdo, fazendo-o a partir da mobilização de mão-de-obra barata – nos Estados Unidos, por exemplo, paga-se 15 dólares a hora pelo trabalho<sup>379</sup>. Dia após dia, essas pessoas veem o que há de pior na internet – vídeos de pessoas sendo decapitadas, imagens de animais sendo submetidos a maus tratos, fotografias de estupro, pornografia infantil, etc. – e decidem, a partir de um treinamento prévio, se aquilo deve ou não permanecer na internet. Não é surpreendente, portanto, que a realidade desse tipo de atividade tem sido denunciada por inúmeras reportagens que enfocam os traumas psicológicos permanentes que o trabalho da moderação de conteúdo gera a quem o faz<sup>380</sup>, uma abordagem que não é propriamente o tema da dissertação, mas que se torna útil na medida em que expõe como esse trabalho é feito.

Assim, detectado um conteúdo potencialmente violador das políticas da plataforma, ele é direcionado para uma espécie de fila virtual em que aguarda para ser avaliado por um moderador. Essa fila é dividida de acordo com graus de prioridade, que atribuem os conteúdos, segundo o seu assunto e urgência, a moderadores de diferentes posições – no caso do Facebook, dos níveis 2 ou 3. Essas pessoas, após passarem por treinamentos nos quais têm contato tanto com as políticas de conteúdo da plataforma – é dessa fase que provêm, por exemplo, os slides vazados nos *Facebook Files*<sup>381</sup> – quanto

---

<sup>378</sup> “Today, Tier 3 moderators typically work in “call centers”<sup>294</sup> in the Philippines, Ireland, Mexico, Turkey, India, or Eastern Europe” (“Hoje, moderadores do Nível 3 tipicamente trabalham em ‘call centers’ nas Filipinas, Irlanda, México, Turquia, Índia e Leste Europeu”). KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.640.

<sup>379</sup> NEWTON, C. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

<sup>380</sup> “Several moderators told me they experienced symptoms of secondary traumatic stress — a disorder that can result from observing firsthand trauma experienced by others” (“Muitos moderadores me disseram que tiveram sintomas de stress traumático secundário – uma doença que pode resultar de observar o trauma tido diretamente por outras pessoas”). NEWTON, C. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

<sup>381</sup> Cf. *supra*, notas n.º 21 e 174.

com simulações a respeito de que tipo de decisão devem tomar<sup>382</sup>, têm acesso a uma espécie de fluxo interno das plataformas, no qual os itens a serem moderados são posicionados um em seguida do outro. O moderador, então, acessa o conteúdo – lê a mensagem, vê a imagem ou assiste ao vídeo – e decide o que fazer com ele, movendo a fila em diante<sup>383</sup>. As decisões dos moderadores entram imediatamente em vigor – seja excluir a publicação ou banir um usuário, por exemplo – e são avaliadas tanto quantitativamente quanto qualitativamente. Sob a perspectiva quantitativa, um empregado terceirizado de uma empresa contratada pelo Facebook, por exemplo, contabiliza em um relato que lhe são esperadas cerca de 400 decisões de conteúdo por dia de trabalho, cada uma das quais em menos de 30 segundos<sup>384</sup>. Qualitativamente, suas decisões são medidas por uma métrica de precisão, que leva em conta amostras do seu trabalho que são auditadas por moderadores de grau superior e, em alguns casos, por pessoas que trabalham diretamente para o Facebook: quanto maior o número de decisões coincidentes, maior a precisão – algo que impacta nas metas dispostas contratualmente entre plataforma e terceirizado<sup>385</sup>.

Evidentemente, essa forma de trabalho gera controvérsias, três das quais me proponho a exemplificar a seguir, por esclarecerem as dificuldades que esse modelo decisório enfrenta. Primeiro: por mais bem instruídos e treinados que estejam essas pessoas, a dinâmica da produção de conteúdo nas plataformas sempre vai gerar novas situações, em que moderadores de conteúdo terão que exercer juízos sobre o que veem a

---

<sup>382</sup> “For this portion of her education, Chloe will have to moderate a Facebook post in front of her fellow trainees” (“Como parte de seu treinamento, Chloe terá que moderar uma publicação do Facebook na frente de seus colegas trainees”). NEWTON, C. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

<sup>383</sup> Cf. *infra*, nota n.º 401.

<sup>384</sup> “He will spend less than 30 seconds on each item, and he will do this up to 400 times a day” (“Ele irá gastar menos de 30 segundos em cada item, e vai fazer isso até 400 vezes por dia”). NEWTON, C. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

<sup>385</sup> “Accuracy, in this case, means that when Facebook audits a subset of contractors’ decisions, its full-time employees agree with the contractors” (“Precisão, nesse caso, significa que quando o Facebook auditor uma parte das decisões dos contratados, seus empregados em tempo integral irão concordar com os contratados”). NEWTON, C. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

partir das normativas da comunidade. Nem sempre o seu trabalho se resumirá a lembrar se a imagem publicada por um usuário está do lado permitido ou proibido de uma tabela repassada no treinamento<sup>386</sup>: frequentemente, essas pessoas se veem diante da necessidade de avaliar se uma situação que lhes é apresentada é ou não uma ameaça, uma forma de discurso de ódio, uma influência negativa para crianças ou uma demonstração de violência, sem nunca terem sido apresentados a conteúdo semelhante. Pessoas treinadas para repetir padrões pré-determinados são chamadas, nesse contexto, a oferecer – em trinta segundos e de forma absolutamente descontextualizada<sup>387</sup> – respostas às controvérsias que têm ocupado há anos a literatura especializada e a jurisprudência dedicada à liberdade de expressão: quando uma manifestação passa a ser uma ameaça?<sup>388</sup> Até onde os direitos autorais podem restringir a liberdade de expressão?<sup>389</sup> O que é incitar alguém a fazer algo?<sup>390</sup>

---

<sup>386</sup> Cf. *supra*, Cap. I.2, 2.3.1.

<sup>387</sup> “*To speed things along, and also for reasons of privacy, the images and posts are usually detached from their original material. Moderators are generally unaware of the identity of the user, and are provided little of the surrounding context — the back and forth of a conversation, that user’s previous posts, and so on*” (“*Para tornar as coisas mais rápidas e também por razões de privacidade, as imagens e publicações são usualmente deslocadas de seu material original. Moderadores geralmente não conhecem a identidade do usuário e têm pouca informação sobre o contexto – as partes anterior e seguinte da conversação*”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 129 (Kindle, n. p.).

<sup>388</sup> Cf., por exemplo, o debate em *Virginia v. Black*, 538 U.S. 343 (2003), em que uma lei da Virgínia foi declarada inconstitucional na extensão em que ela proibia, sem restrições, o ato de queimar uma cruz – conduta historicamente associada à Ku Klux Klan nos Estados Unidos. O cerne do precedente é o debate entre qual é o limite entre um ato que expressa uma ideologia – e que é tutelado pelo direito à liberdade de expressão – e uma ameaça intimidatória real – que pode ser objeto de controle do Estado. Esse tipo de problema – definir o que é uma ameaça real – aparece, por exemplo, no relato de um moderador de conteúdo do Pintarest, que se viu diante de comentários de um homem que publicou comentários dizendo que tinha a intenção de cozinhar mulheres (cf. BUNI, C., CHEMALY, S. *The Secret Rules of the Internet*. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019).

<sup>389</sup> Cf., por exemplo, o debate em *Harper & Row v. Nation Enterprises*, 471 U.S. 539 (1985), em que a revista *The Nation* foi processada por publicar trechos das memórias de Gerald Ford, um material protegido por direitos autorais, argumentando que havia interesse público em fazê-lo. A Corte entendeu que a publicação não estava salvaguardada pela doutrina do “*fair use*” e, por isso, manteve a condenação. Esse tipo de controvérsia – quanto a se definir o que é *fair use* – é exatamente o tema de uma discussão que opôs produtores de vídeos e os moderadores de conteúdo do YouTube, por discordâncias quanto à política da



Segundo: ao mesmo tempo em que essa metodologia de trabalho exige que juízos sejam feitos por essas pessoas, ela também faz com que elas tenham que estar localizadas em diferentes países, na medida em que a língua<sup>391</sup> – expressamente a marcada por informalidades, típicas das redes sociais – pode ser obstáculo significativo à moderação de conteúdo. Isso expõe, entretanto, a dificuldade que se tem em unificar entendimentos a respeito dos mais básicos conceitos entre pessoas de passado cultural distinto: uma moderadora de conteúdo do YouTube, por exemplo, relata a divergência entre equipes dos Estados Unidos e da Índia quanto a um vídeo em que um grupo de estudantes circunda duas pessoas brigando, porque, enquanto os moderadores americanos viam ali crianças – algo vedado pela política da plataforma –, os indianos os consideravam adultos<sup>392</sup>.

---

plataforma; *cf.*, por exemplo, COOPER, R. Copyright laws are breaking YouTube. Here's how to fix the problem. **The Week**, 7.3.2016, disponível em <<https://theweek.com/articles/608700/copyright-laws-are-breaking-youtube-heres-how-fix-problem>>, acesso em 24.7.2019.

<sup>390</sup> *Cf.*, por exemplo, *Brandenburg v. Ohio*, 395 U.S. 444 (1969), em que discutida a possibilidade de se condenar falas realizadas por uma líder da Ku Klux Klan durante uma manifestação; o caso é importante por firmar o entendimento de que o Estado não pode punir discursos inflamados (“*inflammatory speech*”), exceto se “*directed to inciting or producing imminent lawless action and is likely to incite or produce such action.*” (“*dirigido a incitar ou produzir ações ilícitas iminentes, sendo provável que essa ação seria incitada e produzida*”). Nas plataformas, o tema apareceu, por exemplo, na controvérsia quanto a se uma página com o nome “*Kill George Zimmerman*” – referência a um vigilante comunitário que matou um adolescente na Flórida e foi absolvido – era uma forma de incitação à violência que deveria ser retirada da plataforma. *Cf.* KLEINMAN, A. Why ‘Kill George Zimmerman’ Facebook Pages Still Exist. **The Huffington Post**, 17.7.2013, disponível em <[https://www.huffpostbrasil.com/2013/07/17/kill-george-zimmerman-facebook\\_n\\_3611633.html](https://www.huffpostbrasil.com/2013/07/17/kill-george-zimmerman-facebook_n_3611633.html)>, acesso em 24.7.2019.

<sup>391</sup> *Cf.*, a esse respeito, as dificuldades que o Facebook tem para implantar métodos de moderação de conteúdo em Mianmar, porque há poucos moderadores que falem birmanês; essas barreiras foram relacionadas com a facilidade com que a plataforma foi utilizada no país para estimular a violência contra um grupo minoritário islâmico, como relatado em KOEBLER, J. e COX, J. The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People. **Motherboard**, 23.8.2018, disponível em <[https://motherboard.vice.com/en\\_us/article/xwk9zd/how-facebook-content-moderation-works](https://motherboard.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works)>, acesso em 24.7.2019; e STECKLOW, S. Why Facebook is losing the war on hate speech in Myanmar. **Reuters**, 15.8.2018, disponível em <<https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>>.

<sup>392</sup> BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

Terceiro: esse é um trabalho que tem destruído as pessoas. Pelo próprio modelo de trabalho adotado – contratos terceirizados permeados por cláusulas de sigilo, que vedam tanto as partes quanto seus empregados de discutirem a respeito do que fazem<sup>393</sup> – é ainda difícil estabelecer uma base empírica que permita conclusões confiáveis, mas é possível depreender, dos relatos disponíveis, que a atividade dos moderadores de conteúdo, tal como posta pela sistemática vigente, têm causado danos psicológicos em diferentes graus a essas pessoas. Há relatos de moderadores que tiveram alterações na vida sexual pela exposição contínua à pornografia<sup>394</sup>, surtos de paranoia mesmo quanto a pessoas próximas<sup>395</sup> e diagnóstico de stress pós-traumático<sup>396</sup> após exercerem a função, algo que coloca em dúvida o quão sustentável esse modelo é no futuro.

Quanto à segunda possibilidade dessa classificação – a decisão automatizada –, a despeito de ela constituir uma categoria à parte em relação à detecção – tanto que, como afirmei acima, é possível que haja decisões manuais após o conteúdo ser detectado de forma automatizada, bem como decisões automáticas após a detecção manual –, ela segue muito do que afirmei na seção anterior. Os mesmos mecanismos utilizados para detectar conteúdo de forma automática são utilizados para tomar decisões a esse

---

<sup>393</sup> “*Very little is known about how platforms set their policies — current and former employees like Mora-Blanco and others we spoke to are constrained by nondisclosure agreements*” (“*Muito pouco se sabe sobre como as plataformas determinam suas políticas – atuais e ex-empregados como Mora-Blanco e outros que nós falamos são vinculados a acordos de confidencialidade*”). BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

<sup>394</sup> “*Workers quit because they feel desensitized by the hours of pornography they watch each day and no longer want to be intimate with their spouses*” (“*Trabalhadores desistem porque eles se sentem insensibilizados pelas horas de pornografia que assistem a cada dia e não querem mais ser íntimos de seus companheiros*”). CHEN, A. The laborers who keep dick pics and beheadings out of your Facebook feed. **Wired**, 23.10.2014, disponível em <<https://www.wired.com/2014/10/content-moderation/>>, acesso em 24.7.2019.

<sup>395</sup> “*Constant exposure to videos like this has turned some of Maria's coworkers intensely paranoid*” (“*A constante exposição a vídeos como esse deixou alguns dos colegas de Maria intensamente paranóicos*”). CHEN, A. The laborers who keep dick pics and beheadings out of your Facebook feed. **Wired**, 23.10.2014, disponível em <<https://www.wired.com/2014/10/content-moderation/>>, acesso em 24.7.2019.

<sup>396</sup> Cf., por exemplo, HERN, A. Facebook moderators tell of strict scrutiny and PTSD symptoms. **The Guardian**, 26.2.2019, disponível em <<https://www.theguardian.com/technology/2019/feb/26/facebook-moderators-tell-of-strict-scrutiny-and-ptsd-symptoms>>, acesso em 24.7.2019.

respeito. Um exemplo que mencionei acima é o do ContentID: de forma semelhante ao PhotoDNA e ao convênio existente entre as plataformas e as instituições que protegem a infância, o YouTube firma parcerias com detentores de grandes quantidades de materiais protegidos por direitos autorais<sup>397</sup>. Por meio delas, a plataforma ganhou acesso a um banco de dados de conteúdos que não podem ser disponibilizados pelos seus usuários sem autorização e criou um mecanismo decisório que automaticamente compara e tira do ar vídeos que constem como vedados. Essa forma decisória pode ser precedida tanto pela detecção automática – usualmente realizada quando alguém envia o vídeo para publicação para a plataforma – quanto manual, via, por exemplo, *flagging* de um usuário que reporta um vídeo como violador dessas regras.

No mais, os métodos decisórios automáticos têm as mesmas vantagens e enfrentam os mesmos problemas das formas de detecção automática. Eles permitem uma forma barata, rápida e, por isso, eficiente de moderar conteúdo, lidando bem com o problema de escala que é inerente às plataformas atuais; mas têm limitações, porque dependem da inserção prévia de informações – como, no caso, o material que é protegido por direitos autorais e o que fazer uma vez que algo seja detectado – e são incapazes de compreender certos contextos em que conteúdos são veiculados – como, por exemplo, no caso da mudança de política quanto à negação do Holocausto, em que um mecanismo de detecção e decisão automática retirou do ar documentários que, como parte de sua função informativa, expunham símbolos nazistas<sup>398</sup>.

Como na seção anterior, o quadro abaixo articula as classificações que propus quanto às formas pelas quais decisões são tomadas pelas plataformas a respeito do conteúdo veiculado por seus usuários – elas serão, no final do capítulo, sistematizadas em um só quadro que as junte com as demais fases da moderação de conteúdo:

#### **Quadro 9 – Fase n.º 2: A decisão quanto ao conteúdo no procedimento adjudicatório**

<u>Critério</u>	<u>Classificações</u>	<u>Exemplos</u>
-----------------	-----------------------	-----------------

<sup>397</sup> GOOGLE. Como funciona o ContentID. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2797370?hl=pt-BR>>, acesso em 24.7.2019.

<sup>398</sup> Cf. *supra*, nota n.º 160.

<b>Modo</b> Como a decisão é tomada?	<u>Centralizado</u> : a decisão é tomada por um núcleo de poder único, que impacta toda a plataforma.	Caso Lipman, Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); Caso Alex Jones (Cap. 1, 3.2.2.2); Negação do Holocausto (Cap. 1, 3.2.2.3).
	<u>Descentralizado</u> : a decisão é tomada por núcleos de poder dispersos a, impactando apenas determinadas comunidades da plataforma.	Poder conferido pelo Facebook aos proprietários de comunidades na plataforma para excluírem membros que publicam spam.
<b>Agente decisório</b> Quem toma a decisão?	<u>Manual</u> : a decisão é tomada por um moderador de conteúdo humano.	Caso Lipman, Caso da imagem “ <i>O Horror da Guerra</i> ” (Cap. 1, 3.2.2.1); Caso Alex Jones (Cap. 1, 3.2.2.2); Negação do Holocausto (Cap. 1, 3.2.2.3).
	<u>Automatizada</u> : a decisão é tomada por um <i>software</i> .	Uso do PhotoDNA; uso do ContentID, no YouTube; decisões tomadas pelo YouTube após mudar sua política sobre a negação ao Holocausto.

#### 2.4.2.3 Fase n.º 3: Sancionar

Detectado o conteúdo sujeito à moderação de conteúdo e realizada uma decisão a seu respeito, em alguns casos disso decorrem consequências – que aqui estou albergando sob um uso geral da palavra “*sanção*”. É importante ressaltar, logo de início, que essa não é uma fase essencial da moderação de conteúdo em sentido estrito: frequentemente, conteúdos são reportados à plataforma, que os analisa e decide que não há nada de errado com eles. Também é importante perceber que, diferente do que acontece entre a detecção e a decisão – que são momentos, via de regra, independentes entre si, algo que justificou a abordagem separada de ambos acima –, a sanção geralmente é aplicada pela mesma via que a decisão é tomada. Assim, de decisões tomadas de forma centralizada decorrem sanções centralizadas; e de decisões tomadas de forma descentralizada decorrem sanções também descentralizadas. Da mesma forma, quando decisões são tomadas por mecanismos automatizados, a sanção também é automatizada; ao passo que, na moderação

de conteúdo manual, o moderador deve decidir, logo após enquadrar um determinado conteúdo como inapropriado, qual será a consequência disso.

Mesmo assim, é importante abrir um tópico à parte quanto às consequências das decisões tomadas pela moderação de conteúdo das plataformas, porque, à medida que essa instância decisória se tornou cada vez mais complexa, surgiram diferentes formas de responder à publicação de um conteúdo que viola as normas dessas plataformas. Essas diferentes formas se manifestaram nos casos do capítulo anterior – no primeiro, Egeland não chegou a ser banido do Facebook; Jones, de outra forma, o foi – e, por isso, entendê-las faz parte de um projeto que busca reduzir a opacidade da atividade de moderar conteúdo. A esquematização que segue, nesse sentido, divide essas sanções em duas classificações diferentes.

Tomada a decisão, pela moderação de conteúdo, de que a publicação de um usuário não atende às suas normas, dois tipos de consequências podem surgir: as que dizem respeito ao âmbito externo da plataforma; e as que atinem ao seu âmbito interno. As primeiras são específicas de situações em que se fazem necessárias medidas além da plataforma, tomadas seja por um motivo de urgência – alguém publica um vídeo dizendo que irá se suicidar, por exemplo – seja por uma obrigação legal – caso da pornografia infantil; as segundas são necessárias – mesmo que em um grau simbólico, de uma advertência particular, por exemplo –, porque não faria sentido, para a plataforma, estabelecer uma regra e não sancionar o seu descumprimento de alguma forma.

Quanto às consequências externas, elas decorrem da necessidade de que providências a respeito do conteúdo veiculado sejam tomadas em um ambiente que não seja o da plataforma. A mais evidente é a de uma situação de urgência, como casos em que alguém publica na plataforma um vídeo dizendo que irá se suicidar<sup>399</sup> ou que irá causar algum mal para outras pessoas. Há também casos em que a plataforma tem a obrigação legal de fazer algo para além dos seus limites a respeito de um conteúdo veiculado por meio delas. Assim, caso algum moderador detecte publicações em que crianças estão expostas a algum risco, ele deve não só a excluir, mas encaminhar os dados de quem a veiculou para um supervisor, que os enviará ao Centro Nacional de Crianças Desaparecidas e Exploradas – o mesmo que viabiliza o PhotoDNA – que comunicará as

---

<sup>399</sup> Cf., por exemplo, a situação de um homem na Turquia que cometeu suicídio ao vivo no Facebook, em ANDREWS, T. Turkish man, 22, fatally shoots himself on Facebook Live. **The Washington Post**, 13.10.2016, disponível em <[https://www.washingtonpost.com/news/morning-mix/wp/2016/10/13/turkish-man-22-fatally-shoots-himself-on-facebook-live/?utm\\_term=.4a95c3b1bd37](https://www.washingtonpost.com/news/morning-mix/wp/2016/10/13/turkish-man-22-fatally-shoots-himself-on-facebook-live/?utm_term=.4a95c3b1bd37)>, acesso em 22.7.2019.

autoridades para providências – uma obrigação legal nos Estados Unidos imposta às plataformas<sup>400</sup>.

Quanto às sanções internas, elas se voltam tanto ao conteúdo veiculado, quanto ao usuário que o publica. Elas aparecem refletidas no relato de um moderador de conteúdo quanto às opções que lhe são disponibilizadas uma vez que constata – em um caso de moderação manual *ex post* – a inadequação da publicação:

*“Após assistir um par de segundos de cada conteúdo, membros do SQUAD clicaram em um dos quatro botões que apareceram na parte superior direita de suas telas: ‘aprovar’-deixar o vídeo ficar; ‘inapropriado’ – marcar o vídeo como para maiores de 18 anos; ‘rejeitar’ – remover o vídeo sem penalidade; ‘penalizar’ – remover o vídeo com uma penalidade para a conta.”*<sup>401</sup>

No que diz respeito às sanções que afetam o conteúdo, uma perspectiva inicial identificaria apenas duas possibilidades ao moderador de conteúdo – autorizar a permanência da publicação ou retirá-la. A complexidade das manifestações veiculadas pelas plataformas e de suas políticas, entretanto, têm criado pontos intermediários entre esses extremos, de modo a viabilizar a permanência do conteúdo sem que isso prejudique a experiência de pessoas mais sensíveis na plataforma. Dessa forma, conteúdos capazes de impactar pessoas sensíveis – uma cena de violência, ou que contenha elementos de sexualidade, por exemplo –, mas que não chegam a violar as políticas de conteúdo das plataformas não são deletadas, mas marcadas como tal pelos moderadores. A ideia é

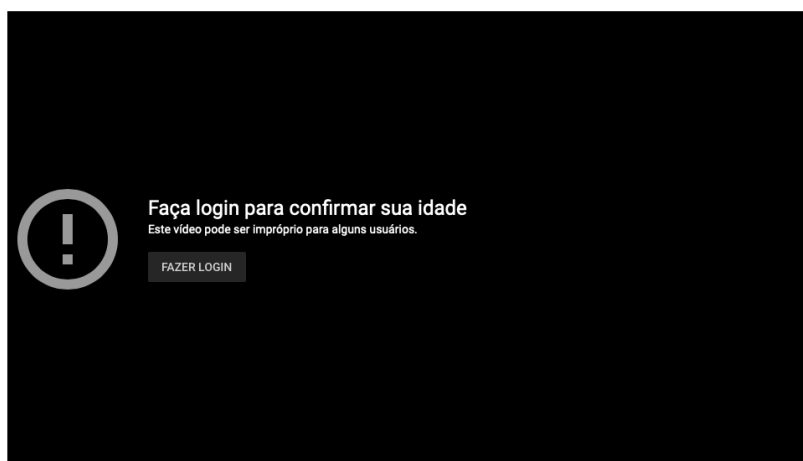
---

<sup>400</sup> “U.S. law imposes clear and special obligations on content and social media platforms, under which suspected content and users must be immediately reported to the National Center for Missing and Exploited Children (NCMEC)” (“O direito dos Estados Unidos impõe claras e específicas obrigações em plataformas de mídia social e de conteúdo, sob as quais conteúdo e usuários suspeitos devem ser imediatamente reportados para o Centro Nacional de Crianças Desaparecidas e Exploradas (NCMEC)”). GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, p. 62 (Kindle, n. p.).

<sup>401</sup> “After watching a couple seconds apiece, SQUAD members clicked one of four buttons that appeared in the upper right hand corner of their screens: “Approve” — let the video stand; “Racy” — mark video as 18-plus; “Reject” — remove video without penalty; “Strike” — remove video with a penalty to the account. Click, click, click. But that day Mora-Blanco came across something that stopped her in her tracks”. BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

estabelecer, por recursos da infraestrutura da plataforma, barreiras que permitam aos usuários optar por terem acesso ou não àquele conteúdo, privilegiando-se a permanência do que se veicula nas plataformas no lugar de sua simples retirada. O YouTube, por exemplo só permite que usuários cadastrados – e que tenham desativado o filtro de conteúdo em suas contas – tenham acesso a vídeos marcados como “18+” ou que contenham cenas de violência<sup>402</sup>. O Facebook, por sua vez, adicionou avisos a vídeos com conteúdo sensível, exigindo um clique adicional do usuário para que ele seja exibido<sup>403</sup>. De uma forma semelhante, o Twitter, como relatei acima<sup>404</sup>, estabeleceu uma política diferenciada para publicações de políticos e candidatos a cargos públicos que veiculem conteúdo contrário às suas regras, mas de interesse público: a publicação permanece na plataforma, mas também exige um clique adicional para ser acessada, bem como tem o alcance diminuído, sendo menos acessível às buscas realizadas na plataforma.

**Figura 18 – Restrição de acesso por idade no YouTube**



Fonte: YouTube<sup>405</sup>

<sup>402</sup> “Podemos aplicar a restrição de idade em vez de remover o conteúdo se o vídeo incluir uma das opções acima [de conteúdo violento ou explícito], mas fornecer contexto suficiente para compreendê-lo”. GOOGLE. Políticas de Conteúdo Violento ou Explícito. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2802008?hl=pt-BR/>>, acesso em 23.7.2019

<sup>403</sup> FACEBOOK. Por que vejo um aviso antes de poder visualizar uma foto ou um vídeo? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/814083248683500?helpref=uf\\_perma-link](https://www.facebook.com/help/814083248683500?helpref=uf_perma-link)>, acesso em 23.7.2019.

<sup>404</sup> Cf. *supra*, Cap. I.2, 2.2.

<sup>405</sup> A imagem foi obtida utilizando-se de recurso que permite extrair uma cópia da tela do computador (“*screenshot*”). A figura é um retrato do que aparece no local em que, usualmente, é visualizado o vídeo; utilizei como exemplo um vídeo de conteúdo violento aleatoriamente encontrado na plataforma.

Por sua vez, quanto ao usuário, as sanções ganham diferentes graus de modulação, que vão desde consequências que restringem a sua atuação na plataforma – proibindo-o de publicar por um certo tempo, como no caso Egeland<sup>406</sup> – até o banimento do perfil – como no caso Alex Jones<sup>407</sup>. Há, entretanto, diversos níveis intermediários entre um e outro: no caso Lipman, por exemplo, a exclusão do conteúdo foi acompanhado do impedimento de se publicar algo em um determinado período de tempo<sup>408</sup>; algo parecido ocorreu com Jones no Twitter, antes de ser banido<sup>409</sup>. Essa variabilidade de punibilidades pode ser acompanhada de uma gradação, em que sanções cada vez mais restritivas são aplicadas à medida que o usuário comete falhas mais graves ou reincide nas mais leves. As Políticas do YouTube fornecem um raro retrato de transparência quanto a como isso funciona, dividindo as sanções em “*alertas*”, que vão do primeiro – que gera uma advertência – até o terceiro – que gera o banimento da plataforma:

***“Alerta***

*Entendemos que erros acontecem e que as pessoas não querem violar nossas políticas. É por isso que a primeira violação é apenas um alerta. Note que você será alertado uma vez. Na próxima vez que seu conteúdo violar as diretrizes da comunidade, você receberá um aviso. Se achar que cometemos um erro, conteste o alerta.*

***Primeiro aviso***

*Se descobirmos que seu conteúdo não segue nossas políticas pela segunda vez, você receberá um aviso. Isso significa que você não poderá enviar vídeos, transmissões ao vivo, histórias nem criar miniaturas personalizadas ou postagens da comunidade por uma semana. Todos os privilégios serão restabelecidos automaticamente após esse período, mas o aviso continuará no canal por 90 dias.*

***Segundo aviso***

*Se você receber um segundo aviso dentro do mesmo período de 90 dias da primeira ocorrência, não poderá postar conteúdo por duas semanas. Se não houver outros problemas, todos os privilégios serão restaurados automaticamente depois desse período, mas cada aviso expira 90 dias depois de ser enviado.*

---

<sup>406</sup> Cf. *supra*, Cap. I.1, 3.2.2.1.

<sup>407</sup> Cf. *supra*, Cap. I.1, 3.2.2.2.

<sup>408</sup> Cf. *supra*, Introdução.

<sup>409</sup> Cf. *supra*, nota n.º 138.



### **Terceiro aviso**

*Se você receber três avisos em 90 dias, seu canal será removido permanentemente do YouTube. Além disso, observe que cada aviso expira 90 dias depois do envio e que a exclusão do conteúdo não removerá o aviso.*<sup>410</sup>

Restrições quanto àquilo que se pode publicar, entretanto, não são as únicas que podem ser aplicadas pelas plataformas. Recentemente, por exemplo, o YouTube adicionou uma forma de sanção distinta das que relatei acima: a desmonetização<sup>411</sup>. Em síntese, o modelo de negócios da plataforma remunera os usuários que veiculam vídeos de acordo com a sua capacidade de gerar visibilidade a anúncios, que são inseridos, por exemplo, antes do início da execução do conteúdo. Assim, quanto mais visualizações alguém é capaz de gerar, mais “monetizado” será o vídeo e mais recursos ele irá gerar para o seu criador. Essa, entretanto, não será mais uma opção para canais da plataforma que façam publicações contrárias às políticas de conteúdo: como anunciado em 5 de junho de 2019, pelo YouTube, “*canais que repetidamente atingem o limiar de nossas políticas de discurso de ódio serão suspensos do Programa de Parceria do YouTube, o que significa que eles não poderão executar anúncios em seus canais*”<sup>412</sup> – ou seja, serão desmonetizados como uma consequência sancionatória pela violação dessas políticas.

Como fiz nas seções anteriores, o quadro abaixo consolida a classificação acima:

### **Quadro 10 – Fase n.º 3: A aplicação de sanções no procedimento adjudicatório**

---

<sup>410</sup> YOUTUBE. Conceitos básicos sobre Avisos das diretrizes da comunidade. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2802032>>, acesso em 22.7.2019.

<sup>411</sup> Cf., a esse respeito, o anúncio oficial da medida em YOUTUBE. Our ongoing work to tackle hate. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>>, acesso em 22.7.2019; já há, entretanto, críticas pela sua inefetividade, porque as políticas do YouTube já têm feito com que os criadores dos canais procurem outra forma de monetização cf. ALEXANDER, J. YouTube looks to demonetization as punishment for major creators, but it doesn't work. **The Verge**, 25.6.2019, disponível em <<https://www.theverge.com/2019/6/25/18744246/youtube-demonetization-steven-crowder-patreonadvertising-merch>>, acesso em 22.7.2019.

<sup>412</sup> “Channels that repeatedly brush up against our hate speech policies will be suspended from the YouTube Partner program, meaning they can't run ads on their channel”. YOUTUBE. Our ongoing work to tackle hate. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>>, acesso em 22.7.2019.

<u>Critério</u>	<u>Classificação 1</u>	<u>Classificação 2</u>	<u>Exemplos</u>
<p><b><u>Abrangência</u></b></p> <p>Onde ocorrem as consequências diretas da sanção</p>	<p><u>Consequências externas</u>: a decisão repercute diretamente fora da plataforma.</p>	-	<p>Notificação das autoridades de saúde, em caso de emergências; notificação de autoridades policiais, na hipótese de pornografia infantil.</p>
	<p><u>Consequências internas</u>: a decisão repercute diretamente dentro da plataforma.</p>	<p><u>Voltadas ao conteúdo</u>: impactam na publicação feita pelo usuário, abrangendo desde restrições à sua circulação até a exclusão.</p>	<p>Política do YouTube de permitir apenas a usuários cadastrados o acesso a conteúdos inapropriados; ferramenta do Facebook que exige um clique adicional para visualizar vídeos com conteúdo sensível; política do Twitter quanto a manifestações de políticos.</p>
		<p><u>Voltadas ao usuário</u>: impactam na atuação do usuário na plataforma, abrangendo desde a desmonetização até o banimento.</p>	<p>Caso Lipman, Caso da imagem “<i>O Horror da Guerra</i>” (Cap. 1. 3.2.2.1); Caso Alex Jones (Cap. 1, 3.2.2.2);</p>

### 3. DA PRIMEIRA EMENDA À GOVERNANÇA PRIVADA: DUAS LEITURAS SOBRE A MODERAÇÃO DE CONTEÚDO

#### 3.1 Introdução

Como esclareci na introdução, a proposta deste capítulo é lançar luz sobre a opacidade que é característica – e problemática – da moderação de conteúdo. Para fazer isso, estabeleci dois passos: primeiro, expor, com mais detalhes – isto é, para além de uma abordagem casuística – o que é a moderação de conteúdo. Fiz isso, na seção anterior, a partir de uma sistematização em que a apresentei desde o seu conceito mais geral, até as formas específicas pelas quais as plataformas cominam sanções a usuários que violam a suas políticas de conteúdo. Segundo, expor as perspectivas a partir das quais a literatura especializada entende moderação de conteúdo, sistematizando, assim, as estruturas que

temos, hoje, para analisar essa forma de regular a liberdade de expressão – o que passo a fazer nesta seção. Esse é um passo importante, porque ao apresentar, no próximo capítulo, o problema de pesquisa que proponho com este trabalho, irei demonstrar, de forma concomitante, que as leituras atuais a respeito da moderação de conteúdo – que descrevo a seguir – não conseguem resolvê-lo, algo que servirá de justificativa para uma abordagem diferente das que até agora têm sido propostas.

As perspectivas que seguem são concretizações da clivagem que identifiquei no capítulo anterior, entre visões diádicas e pluralistas da regulação da liberdade de expressão. Às perspectivas diádicas, que tomam esse tipo de regulação exclusivamente como uma relação entre Estado e indivíduo, corresponde a primeira corrente, que tenta subsumir a moderação de conteúdo às categorias jurisprudenciais extraídas, pela Suprema Corte, de interpretações da Primeira Emenda, fazendo-o para identificar o regime jurídico aplicável às plataformas. Às perspectivas pluralistas, que tomam aquilo que as plataformas fazem em relação aos seus usuários como uma forma à parte de regulação, corresponde a segunda corrente, que mobiliza as ideias de governança privada e autorregulação para abordar a moderação de conteúdo não à luz de um regime jurídico particular, que lhes seria aplicável, mas sim quanto a como funcionam de forma relativamente independente, decidindo aquilo que as pessoas podem ou não fazer a partir de critérios e procedimentos por elas próprias criados.

Como justifiquei no capítulo anterior, esse é um trabalho que adota a perspectiva pluralista da regulação da liberdade de expressão, porque abordá-la de forma diádica significa deixar de compreender aquilo de verdadeiramente revolucionário que decorreu da transição entre as infraestruturas da liberdade de expressão pré-digital e digital – o surgimento não só de novos espaços de manifestação, mas também de espaços de decisão que, em paralelo aos Estados, determinam aquilo que as pessoas podem dizer. Tomei isso, inclusive, como uma justificativa para esse trabalho: fazê-lo enquanto um acesso à moderação de conteúdo, entendendo-a como uma forma de regulação da liberdade de expressão só faz sentido dentro de uma perspectiva pluralista. Seguindo a linha do capítulo anterior – em que reconstruí a posição do Balkin de 2014, como uma forma de compreender o Balkin de 2018 – acesso inicialmente a primeira corrente de entendimentos a respeito da moderação de conteúdo, vinculada à perspectiva diádica, para então abordar a visão pluralista, em que este trabalho se insere.

Como fiz com Balkin no capítulo anterior, não tenho a pretensão de reunir a seguir todas as posições de todos os autores que já escreveram a respeito da moderação de

conteúdo. Assim, irei me concentrar mais em definir as linhas gerais dessas perspectivas do que em descer às particularidades de autores determinados. Tomarei como guia em minha abordagem a reconstrução que Klonick – uma das partes desse debate – fez em um texto de 2018, o mais recente a respeito do tema, cotejando-a com os autores por ela citados.

A seção (3.2) está organizada em duas subseções, cada qual dedicada a uma dessas correntes. Na primeira (3.2.1), descrevo a corrente da “*Primeira Emenda*”, que enxerga a moderação de conteúdo a partir dos contornos jurisprudenciais da Suprema Corte dos Estados Unidos; na segunda (3.2.2), descrevo a corrente que surge a partir da crítica feita por Balkin e Klonick à primeira, e que sustenta, em síntese, que a forma mais adequada de se entender a moderação de conteúdo é a partir de modelos autorregulatórios de governança privada.

### 3.2 *Duas correntes*

#### 3.2.1 A moderação de conteúdo como uma categoria jurisprudencial

A moderação de conteúdo é um tema novo: seu surgimento, ao menos no formato atual – como um modo profissionalizado de tomar decisões sobre o que as pessoas podem dizer nas plataformas da internet – remonta à primeira década dos anos 2000. Nessa época – e ainda hoje<sup>413</sup> – uma reação comum entre aqueles que tentavam compreendê-la de uma forma mais sistematizada foi a de inseri-la nas formas de pensar a regulação da liberdade de expressão que se tinham disponíveis à época. Isso significou, em síntese, adotar como ponto de partida uma perspectiva essencialmente diádica da regulação da liberdade de expressão que, em paralelo à do Balkin de 2014, percebe-a exclusivamente como uma relação entre, de um lado, o Estado regulador e, de outro, o indivíduo que tem as suas manifestações reguladas.

De uma maneira geral, essa perspectiva diádica toma forma em uma metodologia comum a esses primeiros autores, baseada em uma característica da jurisprudência da Suprema Corte dos Estados Unidos: a de diferenciar as obrigações e os

---

<sup>413</sup> *Cf.*, a esse respeito, uma tentativa de adaptar as doutrinas da Primeira Emenda às plataformas em WU, T. Is the First Amendment Obsolete? **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2017, disponível em <<https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>>, acesso em 13.6.2019.

direitos que decorrem da Primeira Emenda – a cláusula que protege a liberdade de expressão na Constituição dos Estados Unidos – de acordo com certas categorias, como Estados, empresas, administradores de cidades operárias, donos de empresas de mídia e editores de jornais, ao redor das quais os precedentes estabelecem regimes jurídicos próprios. Assim, por exemplo, e de acordo com os precedentes que firmaram o “*State Action Requirement*”<sup>414</sup>, a Primeira Emenda estabelece um regime jurídico diferente à categoria “*Estado*” do que o imposto à categoria “*empresa*”, porque o primeiro tem o dever de respeitar a liberdade de expressão e a segunda – via de regra –, não.

Nesse sentido, o que marca os posicionamentos filiados a essa primeira corrente é a tentativa de se estabelecer uma analogia entre, de um lado, a moderação de conteúdo; e, de outro, as categorias da jurisprudência da Suprema Corte quanto à Primeira Emenda, compatibilizando-as de modo que os entendimentos desses precedentes sejam extensíveis à atuação das plataformas, explicando-as, por exemplo, como uma forma de atuação estatal ou um tipo específico de editoração jornalística. Por isso, o resultado dessa corrente são afirmações de que, na verdade, as plataformas agiriam tal como, por exemplo, Estados<sup>415</sup> ou editores de conteúdo<sup>416</sup> e que constatar isso esclareceria algo importante sobre elas – o regime jurídico que lhe seria aplicável.

A partir dessa metodologia comum, os autores dessa corrente divergem entre si na medida em que tomam categorias – e, por consequência, precedentes – diferentes para estabelecer essa relação de analogia. O que se discute aqui é qual categoria melhor se aproxima daquilo que as plataformas fazem quando moderam conteúdo e, por isso, lança mais luz sobre essa atividade, ao esclarecer melhor o regime jurídico que lhes seria aplicável. No que segue, exponho, exemplificadamente, três dessas abordagens, segundo a linha condutora apresentada por Klonick<sup>417</sup>.

---

<sup>414</sup> Cf. *supra*, Cap. I.1, 3.3.4.

<sup>415</sup> Cf., por exemplo, CHANDER, A. Facebookistan. **North Carolina Law Review**, v. 90 n. 5, pp. 1.807/1.844, 2012; e FARREL, H., LEVI, M., O’REILLY, T. Mark Zuckerberg runs a nation-state, and he’s the king. **Vox**, 10.4.2.2018, disponível em <<https://www.vox.com/the-big-idea/2018/4/9/17214752/zuckerberg-facebook-power-regulation-data-privacy-control-political-theory-data-breach-king>>, acesso em 24.7.2019.

<sup>416</sup> Tom Egeland, por exemplo, chamou o Facebook de “*o editor mais poderoso do mundo*”; cf. *supra*, Cap. I.1, 3.2.2.1.

<sup>417</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.658 e ss.

A primeira abordagem estabelece uma analogia entre as plataformas e os Estados. Ela é construída a partir do precedente da Suprema Corte dos Estados Unidos em *Marsh v. Alabama*<sup>418</sup>. O caso transcorreu em Chickasaw, uma cidade-operária no Alabama que foi estabelecida e mantida – inclusive em suas áreas de circulação geral de pessoas, como ruas e calçadas – pela Gulf Shipbuilding Corporation, uma empresa que construía embarcações no local. Grace Marsh, parte da comunidade local de Testemunhas de Jeová, tentou distribuir folhetos religiosos em uma calçada, mas foi avisada pelo xerife que não poderia veicular o material sem uma autorização da empresa, que era a proprietária da área. Marsh insistiu, foi presa e condenada por invadir a propriedade alheia (“*trespass*”). O caso chegou à Suprema Corte, que decidiu reverter a decisão de primeira instância, sustentando que, na medida em que a empresa abria a área da cidade para a livre circulação das pessoas, ela circunscrevia o seu direito à propriedade, de modo que, na situação concreta, ele não se sobrepunha ao direito à liberdade de expressão garantido pela Primeira Emenda detido por Marsh – mesmo que a área em que ela tentava fazer a distribuição fosse, a rigor, privada.

Encontrado o precedente, o passo seguinte do argumento é compatibilizá-lo à moderação de conteúdo, de modo que seja possível entendê-la sob a perspectiva do que decidiu a Suprema Corte no caso. Assim, como exemplifica Klonick, um caminho é extrair do caso *Marsh* uma regra geral, segundo a qual se uma “*entidade privada exerce poderes que são ‘tradicionalmente uma prerrogativa exclusiva do Estado’*”<sup>419</sup>, então as obrigações que decorrem da Primeira Emenda – notadamente a de respeitar o direito das pessoas se expressarem sobre a religião que professam – devem ser respeitadas. Estabelecer essa analogia se traduziria, portanto, em verificar em que medida as plataformas da internet, ao moderar conteúdo, exercem uma atividade “*tradicionalmente*” identificada com uma prerrogativa exclusiva do Estado, para, então, impor a elas o regime da Primeira Emenda.

Nesse sentido, o argumento é de que, assim como em *Marsh*, as plataformas exercem atividades muito próximas as de um Estado ao moderar conteúdo: elas estabelecem regras – a exemplo dos Padrões da Comunidade, do Facebook –, instituem um procedimento de aplicação dessas regras a casos concretos – como o que concluiu que

---

<sup>418</sup> *Marsh v. Alabama*, 326 U.S. 501 (1946).

<sup>419</sup> “*Subsequent courts have concluded that these facts bear on whether “the private entity has exercised powers that are ‘traditionally the exclusive prerogative of the State.’”* KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.658; a citação entre aspas é do caso *Blum v. Yaretsky* 457 U.S. 991 (1982).

aquilo que Alex Jones publicava era contrário às normas da plataforma – e cominam sanções, se o caso – como a que baniu Jones do Facebook. Elas, em síntese:

*“[...] desenvolveram uma infraestrutura similar [à dos Estados] para regular as manifestações dos usuários, fazendo-o com regras detalhadas, moderação ativa e passiva, confecção treinada de decisões humanas, raciocínios por analogia e inputs e outputs de fontes internas e externas”<sup>420</sup>.*

A principal consequência dessa analogia – de onde adviria o seu caráter esclarecedor a respeito da moderação de conteúdo – é a conclusão de que, se as plataformas agem como Estados, no sentido do caso *Marsh*, então, assim como a empresa que era proprietária da cidade-operária, também elas estão vinculadas à Primeira Emenda. Por isso, a atuação das plataformas da internet, quando tomam decisões a respeito daquilo que seus usuários veiculam por meio delas, deve ser analisada à luz do regime jurídico que a Primeira Emenda impõe aos Estados e que garante, às pessoas, o direito à liberdade de expressão. Nesse sentido, aquilo que ocorre nas plataformas não é uma forma diferente de regular a liberdade de expressão – como propõe, por exemplo, a visão pluralista do Balkin de 2018, com a qual concluí o capítulo anterior – mas sim a exata mesma forma pela qual o Estado o faz, utilizando-se a mesma regulamentação que, entretanto, é aplicada – e deve ser respeitada – por um ator privado.

Na prática, essa interpretação enfrenta, segundo Klonick, dois problemas<sup>421</sup>. O primeiro diz respeito à analogia entre a cidade-operária do caso *Marsh* e aquilo que as plataformas fazem. Diferente da empresa que é dona da cidade do precedente – e que, nesse sentido, invocou para si o direito de decidir, com exclusividade, o que poderia ser divulgado ali –, nenhuma plataforma da internet tem a capacidade de monopolizar o espaço da internet, para decidir, com exclusividade, o que pode ser ou não veiculado ali. Isso porque, por um lado, há outras plataformas: se alguém deseja veicular um conteúdo com algum grau de nudez que não é permitida no Facebook, por exemplo, é possível fazê-

---

<sup>420</sup> “*This Article has detailed how platforms have developed a similar infrastructure to regulate users’ speech through detailed rules, active and passive moderation, trained human decision-making, reasoning by analogy, and input from internal and external sources*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.658.

<sup>421</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.659 e ss.

lo no Twitter – que, como disse antes, tem uma política mais flexível a esse respeito<sup>422</sup>. Assim, a circunstância de o Facebook inadmitir um conteúdo não significa que ele não poderá ser veiculado na internet por meio de uma plataforma, mas apenas que aquela específica plataforma não o aceitará. Por outro, porque há outras formas de se veicular algo na internet, para além das plataformas. Assim, a despeito de elas exercerem o papel de novos intermediários da Web 2.0, na medida em que constroem interfaces que viabilizam a usuários comuns veicularem conteúdos na internet, essa não é a única forma de se produzir algo na rede. Alex Jones, por exemplo, não deixou de produzir e veicular conteúdo após ser banido do Facebook, Twitter e YouTube; ele continua fazendo-o por meio de uma página à parte desses espaços, mas também localizada na internet e, por isso, tão acessível quanto as plataformas, a *Infowars* (<https://www.infowars.com/>).

O segundo problema diz respeito ao tipo de comprometimento que essa analogia gera: se as plataformas atuam como Estados – no sentido do caso *Marsh* – e, por isso, estão vinculadas ao regime jurídico da Primeira Emenda, então uma parte significativa das decisões tomadas via moderação de conteúdo teria que ser revista. Um exemplo disso seria a pornografia: como a reconstrução das normas estabelecidas pelas acima expôs, uma parte significativa das plataformas – a exemplo do Facebook e do YouTube – vedam a disponibilização de nudez ou de conteúdo sexual por seus usuários, de modo que quaisquer fotos ou vídeos nesse sentido são retirados – algo rotineiro no trabalho dos moderadores de conteúdo. Mesmo plataformas como o Twitter, que tem uma política mais flexível, adotam alertas adicionais quanto a conteúdos desse tipo, impedindo que usuários os acessem de forma inadvertida. O problema é que, sob a Primeira Emenda, manifestações que tenham por objeto a nudez, e mesmo a pornografia, são protegidas como uma consequência do direito à liberdade de expressão<sup>423</sup>. Isso significa que, se as plataformas estão vinculadas à Primeira Emenda, elas não mais poderiam suprimir conteúdo desse tipo – algo que poderia criar, na afirmação de Klonick, uma “*internet que ninguém quer*”, em que “*praticamente todos os tipos de expressão seriam explicitamente permitidas e protegidas – fazendo dos atuais problemas online de discurso de ódio,*

---

<sup>422</sup> Cf. *supra*, nota n.º 254.

<sup>423</sup> Para o precedente mais atual a respeito, cf. o caso *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997), detalhado *supra*, nota n.º 72.



*bullying e terrorismo, em relação aos quais muitos ativistas estão preocupadas, inimaginavelmente piores”*.<sup>424</sup>

A segunda abordagem estabelece uma analogia entre plataformas – e sua atividade de moderar conteúdo – e aquilo que editores de jornais fazem quanto ao conteúdo que esse tipo de mídia propaga. Essa categoria também foi objeto de precedentes da Suprema Corte dos Estados Unidos, que fixou, por meio deles, um regime jurídico específico de aplicação da Primeira Emenda. O caso de referência é o julgamento da Corte em *Miami Herald Publishing Co. v. Tornillo*<sup>425</sup>. Pat Tornillo, uma candidata local ao legislativo da Flórida, foi criticada pelo jornal *Miami Herald*, que publicou dois editoriais contrárias à sua eleição para a posição. Tornillo apresentou ao jornal um pedido para que suas respostas a essas críticas fossem também publicadas, em cumprimento à legislação à época vigente na Flórida, que garantia esse direito de resposta a candidatos a cargos políticos. O jornal se recusou e, uma vez processado pela candidata, sustentou que a lei violava a Primeira Emenda – em específico, a cláusula que protege a liberdade de imprensa. A Corte entendeu que jornais, por selecionarem e decidirem tanto aquilo que será publicado, quanto o tipo de tratamento que será dado à informação, não são meros “*receptáculos passivos*”<sup>426</sup> de conteúdo, mas partes ativas que fazem escolhas capazes de refletir posições que lhe são internas a respeito de um determinado assunto. Na perspectiva da Corte, essas escolhas editoriais são uma forma de manifestação que traz os jornais que a exercem para uma posição semelhante à de alguém que externa um posicionamento sobre algo – e, nesse sentido, que fazem deles agentes protegidos pela Primeira Emenda. Por isso, a legislação da Flórida é inconstitucional: ela interfere nessas escolhas editoriais e obriga o órgão de imprensa a publicar algo que não reflete aquilo que acredita.

Nesse caso, a comparação passa por equivaler as plataformas às mídias tradicionais e a moderação de conteúdo à ação editorial. Nesse sentido, a moderação de conteúdo seria vista como uma atividade que, como nos jornais, molda aquilo que os

---

<sup>424</sup> “*All but the very basest speech would be explicitly allowed and protected — making current problems of online hate speech, bullying, and terrorism, with which many activists and scholars are concerned, unimaginably worse*”. KLONICK, K. *The New Governors: The People, Rules, And Processes Governing Online Speech*. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.659.

<sup>425</sup> *Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241 (1974).

<sup>426</sup> “[...] *passive receptacle*”. *Miami Herald Publishing Co. v. Tornillo* *apud* KLONICK, K. *The New Governors: The People, Rules, And Processes Governing Online Speech*. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.660.

usuários da plataforma conseguem ver: da mesma forma que o *The New York Times* exerce uma forma de curadoria sobre as informações, escolhendo-as e publicando-as de acordo com políticas próprias – definindo, por exemplo, o que estará na primeira página –, também as plataformas, ao decidir aquilo que permanece ou não nos espaços que administram, expressam uma posição que deve ser respeitada e que se manifesta, por exemplo, naquilo que é visível na página inicial de cada usuário.

A consequência dessa interpretação é que, inversamente à analogia com o Estado, que vinculava a plataforma a um dever de respeitar os direitos conferidos pela Primeira Emenda, aqui é ela que se resguarda com esse tipo de proteção: assim como jornais têm o direito de estabelecer uma linha editorial que reflita as suas posições, manifestando-a pelas escolhas que fazem a respeito do conteúdo que publicam, plataformas têm o direito de escolher aquilo que será veiculado por meio delas, definindo o que publicam de modo a externalizar suas posições a respeito do conteúdo veiculado por seus usuários. Se a posição da plataforma, portanto, é de vedar a nudez, então ela tem o direito não só de estabelecer políticas de conteúdo nesse sentido, como de “*editar*” aquilo que aparece para os seus usuários, retirando do ar publicações contrárias a esse posicionamento.

Essa forma de ver a moderação de conteúdo tem uma vantagem em relação à anterior: ela se encaixa melhor com a multiplicidade de políticas que existem a respeito de conteúdo, porque confere a cada uma das plataformas a possibilidade de constitui-las livremente e, mais do que isso, tutela essas decisões sob a perspectiva do direito à liberdade de expressão. As plataformas, nesse contexto, se deslocariam de partes das quais se exige o respeito à liberdade de expressão para instâncias que invocam, para si, esse direito.

A analogia, entretanto, tem um problema. Ela ignora aquilo que não apenas distingue as mídias tradicionais das plataformas, mas que constitui as últimas como uma forma diferente de intermediar informações: o fato de que, no caso das plataformas da internet, a informação não é produzida por elas, mas por seus usuários. Elas – e aqui vale a pena retomar o seu conceito – (i) “*hospedam, organizam e circulam, para os seus usuários, o conteúdo que eles compartilham e as interações que fazem*”; (ii) “*sem ter produzido ou comissionado (o grosso) desse conteúdo*”<sup>427</sup>. Essa, portanto, não é uma característica lateral, que pode ser ignorada em uma relação analógica, mas precisamente o

---

<sup>427</sup> Cf. *supra*, nota n.º 238.

discrímen que faz as plataformas serem plataformas. Como discuti antes<sup>428</sup>, essa característica as distancia, de um lado, do extremo das mídias tradicionais – que confeccionam as informações e, por isso, responsabilizam-se pelo seu conteúdo; e, de outro, do extremo dos serviços de envio de informação – que apenas as conduzem de um lugar ao outro, sem nenhum tipo de responsabilidade pelo seu conteúdo. É por isso que é difícil conceber uma plataforma como um jornal: assim como não faz parte do trabalho de um empregado do Facebook produzir informações para os usuários do serviço, faz parte do trabalho de um jornalista do *The New York Times* confeccionar uma reportagem para os seus leitores.

A terceira abordagem estabelece uma analogia entre a atividade das plataformas e a das empresas responsáveis por transmitir sinais de rádio e televisão (“*broadcasters*”) ou por transportar bens de um lugar a outro (“*common carriers*”)<sup>429</sup>. Ela se sustenta não exatamente em um precedente da Suprema Corte, mas em um argumento contrário a ele. O paradigma, nessa abordagem, é o caso *Reno v. ACLU*<sup>430</sup>, precedente em que regras do “*Communications Decency Act*”, uma lei destinada à regulação da publicação de material pornográfico na internet, foram declaradas inconstitucionais. No precedente, como refere Klonick, a Corte “*explicitamente diferenciou a internet da ‘mídia de transmissão’ [‘broadcast media’], porque falta à primeira escassez, invasividade e uma história de regulamentação governamental*”<sup>431</sup>. Assim, as formas de regulamentar a liberdade de expressão utilizadas na mídia tradicional – como criminalizar a conduta de se transmitir material pornográfico a um menor de idade – não podem ser utilizadas indistintamente na internet, mas devem ser analisadas à luz das particularidades desse novo ambiente de comunicação – em que a mesma criminalização pode significar uma restrição muito severa à liberdade de expressão. Por isso, a Corte entendeu que, no específico contexto da internet, “*à CDA [lei] falta a precisão que a Primeira Emenda requiere quando uma lei regula o conteúdo de manifestações*”, porque “*para impedir que menores acessem conteúdo potencialmente danoso, a CDA efetivamente suprime uma quantidade de*

---

<sup>428</sup> Cf. *supra*, Cap. I.2, 2.2.

<sup>429</sup> KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.660 e ss.

<sup>430</sup> *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997), cf. *supra*, nota n.º 72.

<sup>431</sup> “[...] *the Court explicitly differentiated the internet from broadcast media because the former lacks scarcity, invasiveness, and a history of government regulation*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.661.

*manifestações que os adultos têm o direito constitucional de receber e enviar um para o outro*”<sup>432</sup>.

Segundo essa terceira abordagem, modificações ocorridas desde o momento em que o julgamento foi proferido – o caso data de 1997 – fizeram da internet uma ferramenta cada vez menos diferente da “*broadcast media*”, porque as características que foram ditas, à época, como exclusivas da segunda – escassez, invasividade e regulação estatal – agora também se fazem presentes na primeira. Assim, a internet teria se tornado um ambiente em que há escassez – porque o número de espaços de fala, isto é, de plataformas, é limitado –, invasividade – é difícil pensar em algo tão presente em todos os momentos da vida das pessoas – e inúmeras regulamentações governamentais. Esse movimento de aproximação permitiria inserir as plataformas como parte da categoria da qual foram distinguidas no caso *Reno* e, assim, entende-las como instrumentos que transmitem aquilo que as pessoas dizem de um local a outro.

O problema dessa analogia é semelhante ao da anterior, mas dirigido à outra ponta do espectro: se as plataformas não se confundem com a mídia tradicional, porque o conteúdo que veiculam não é o seu, elas também não são meras transmissoras de informações alheias, uma vez que não há dúvidas de que agem modulando o discurso de seus usuários, tanto ao dispor como as pessoas poderão se manifestar na plataforma – via moderação de conteúdo em sentido geral –, quando ao determinar o que elas podem dizer – via moderação de conteúdo em sentido restrito. Qualquer abordagem que identifique as plataformas, por meio de uma analogia, ou com a mídia tradicional, ou com as formas usuais de se transmitir informações, deixa de captar aspectos que lhe são essenciais e, nesse sentido, constrói paralelismos que menos explicam o que é a moderação de conteúdo e mais lhe impõem modelos que se mostram incompatíveis com o que ela é. No caso dessa terceira abordagem, isso significa minimizar a importância que a atuação desses intermediários tem na administração do ambiente em que as pessoas se manifestam, reduzindo-as a fios condutores – algo que não são. Essa distinção – e a impropriedade da analogia – são perceptíveis mesmo porque há serviços na internet – como os de mensagem instantânea e de e-mail – que exercem esse tipo de atividade, mas, exatamente por fazê-lo, não são plataformas.

---

<sup>432</sup> “[...] *the CDA lacks the precision that the First Amendment requires when a statute regulates the content of speech. In order to deny minors access to potentially harmful speech, the CDA effectively suppresses a large amount of speech that adults have a constitutional right to receive and to address to one another*” *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997).

Reuni, nesta seção, três abordagens que concretizam a forma pela qual a moderação de conteúdo das plataformas da internet foi inicialmente analisada. Em síntese, essas perspectivas partem de um método comum, em que traços semelhantes são identificados entre a moderação de conteúdo e categorias jurisprudenciais estabelecidas pela Suprema Corte do Estados Unidos ao interpretar a Primeira Emenda, fazendo-o de modo que seja possível estender as conclusões tiradas pela Corte, por exemplo, quanto à atuação do conselho editorial de um jornal, às plataformas. Esse tronco comum se materializa em abordagens que mobilizam diferentes precedentes e que argumentam, a partir deles, que determinada categoria mais se aproxima do que fazem as plataformas ao moderar conteúdo – e, por isso, fornece as melhores conclusões a esse respeito –, seja as comparando com empresas que detêm cidades-operárias, conselhos editoriais de jornais ou empresas de radiodifusão. Na seção seguinte, apresento as críticas a essa abordagem e a segunda – e mais atual – forma de se pensar a moderação de conteúdo.

### 3.2.2 A moderação de conteúdo como uma forma de governança privada

Como a reconstrução que propus acima expõe, a primeira corrente que analisa a moderação de conteúdo não é imune a críticas, sintetizadas na constatação de que *“nenhuma dessas analogias da moderação privada do direito público à manifestação parece se encaixar precisamente na descrição da natureza do que as plataformas online são, ou nos resultados normativos quanto àquilo que elas querem ser”*<sup>433</sup>. Essa dificuldade em se estabelecerem analogias não chega a ser o cerne do problema desse tipo de abordagem: primeiro, porque é possível que, em algum momento, a moderação de conteúdo venha a ser o tema de um caso na Suprema Corte, eliminando a necessidade desse tipo de paralelismo. Segundo, porque há abordagens filiadas a essa corrente que não se propõem a encontrar uma categoria jurisprudencial para servir como base de uma

---

<sup>433</sup> “[...] none of these analogies to private moderation of the public right of speech seem to precisely meet the descriptive nature of what online platforms are, or the normative results of what we want them to be”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.662.

analogia, mas sim em construir uma específica das plataformas, capaz, inclusive, de fornecer um paradigma regulatório mais adequado ao setor<sup>434</sup>.

O problema, portanto, é mais central: como mencionei no capítulo anterior, ao tentar explicar a moderação de conteúdo a partir de uma perspectiva que não a vê como uma nova instância regulatória, mas apenas como um novo âmbito a ser inserido na relação Estado / indivíduo – subsumindo-o tanto a um antigo, quanto a um novo paradigma regulatório –, perde-se o essencial, porque se deixa de perceber que esses atores privados não apenas constituíram novos espaços de manifestação, mas criaram, eles mesmos, novas instâncias decisórias independentes que, por vezes, têm um impacto muito maior no que as pessoas podem ou não dizer do que os próprios Estados. O problema, nesse sentido, não é que as categorias encontradas na primeira corrente – ou por ela criadas – não se afeiçoam àquilo que as plataformas fazem, mas sim que essa forma de pensar a moderação de conteúdo adota uma perspectiva diádica da regulação da liberdade de expressão, que, como argumentei antes, é simplesmente incapaz de abranger a complexidade regulatória da era digital.

Seja por uma razão ou por outra, Klonick propõe, a partir dessas dificuldades, uma forma diferente de pensar a moderação de conteúdo, construindo uma segunda corrente marcada pela transição de uma visão diádica da regulação da liberdade de expressão para uma perspectiva pluralista.

Em síntese, seu argumento sustenta que a atividade das plataformas de moderar conteúdo é mais bem compreendida “*sob as perspectivas da governança privada e da autorregulação*”<sup>435</sup>. Ela mobiliza o termo “*governança*”, no texto, como um agrupador de “*características que descrevem com precisão a interação entre usuário e plataforma*”<sup>436</sup>, em específico pela forma pela qual as plataformas estabelecem estruturas de exercício de poder sobre aquilo que os seus usuários expressam. Elas, nesse sentido,

---

<sup>434</sup> WU, T. Is The First Amendment Obsolete? **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2017, pp. 19 e ss, disponível em <<https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>>, acesso em 13.6.2019

<sup>435</sup> “*Instead, analysis of online speech is best considered from the perspectives of private governance and self-regulation*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.662.

<sup>436</sup> “*Rather, “new governance model[s]” identify several features that accurately describe the interplay between user and platform*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.617.

“*governam*” aquilo que eles podem dizer, fazendo-o a partir de uma estrutura interna, que tanto gera quanto aplica padrões normativos àquilo que as pessoas dizem de forma relativamente independente. Essa perspectiva tira a efetividade de abordagens que tentam encaixar as plataformas em padrões jurisprudenciais e fornece razões suficientes para justificar abordagens que tentam entender como, afinal de contas, essas estruturas funcionam. Seu argumento para sustentar isso é dividido em duas partes: na primeira, Klonick afirma que essa é uma forma mais *precisa* de se entender as plataformas; na segunda, defende que essa é uma forma *normativamente mais útil* de fazê-lo.

Sob a perspectiva descritiva, falar de governança expressa melhor aquilo que fazem as plataformas, porque, segundo Klonick, elas:

*“[...] têm um corpo centralizado, um conjunto estabelecido de regras, procedimentos ex ante e ex post para adjudicar conteúdos contrários às regras [...]; políticas e regras são modificadas e atualizadas por meio de inputs externos; plataformas são economicamente sujeitas à influência normativa de seus cidadãos-usuários e também agem de forma colaborativa com redes externas como governos e grupos terceiros”<sup>437</sup>.*

Como o trecho mostra, a ideia de governança se encaixa na de plataformas a partir de dois pontos de contato: primeiro, pela existência de uma estrutura interna a essas empresas, semelhante à de um governo, que tanto estabelece quanto adjudica condutas – as duas vertentes em que dividida a moderação de conteúdo em sentido restrito. Segundo, pela existência de uma relativa independência, pela qual as plataformas estabelecem as suas próprias políticas quanto à liberdade de expressão de seus usuários, “*governando-os*”, mas também lidando com o impacto de influências externas, provenientes, por exemplo, de Estados, que são tratadas a partir dos próprios critérios da plataforma.

Assim, quanto ao primeiro ponto de contato, no caso que reconstruí no capítulo anterior e se referia à posição do Facebook quanto a páginas que negavam o Holocausto, via-se com clareza que a plataforma não só tinha um posicionamento a respeito da controvérsia – no caso, o de não excluir essas publicações – quanto o efetivava,

---

<sup>437</sup> “[...] *platforms have a centralized body, an established set of laws or rules, ex ante and ex post procedures for adjudication of content against rules, and democratic values and culture; policies and rules are modified and updated through external input; platforms are economically subject to normative influence of citizen-users and are also collaborative with external networks like government and third-party groups*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.663.

“*governando*” aquilo que os seus usuários podiam dizer. Nesse sentido, e tal como um governo que implanta uma determinada política pública, a plataforma tinha, no caso, uma política de conteúdo, que fazia valer por um conjunto organizado de normativas, que orienta os seus usuários – os “*governados*” – sobre o que podiam dizer, e determina aquilo que os seus moderadores de conteúdo deveriam fazer. Mudada a política – como fez o YouTube, proibindo esse tipo de manifestação em seu ambiente – mudam-se as normativas e, por consequência, a sua aplicação: como relatei, moderadores de conteúdo passaram a excluir esse tipo de publicação da plataforma, ajustando-se mecanismos automatizados para detectá-las<sup>438</sup>.

Quanto ao segundo ponto de contato – e continuando-se no mesmo exemplo –, tal como um governo, o Facebook não simplesmente se sujeitou às influências que lhe eram externas – no caso, as proibições, estabelecidas por alguns países, de que o Holocausto fosse negado – mas a avaliou a partir dos seus próprios critérios, decidindo em que medida lhe era vantajoso adotar ou não essa política. Isso é especialmente claro nessa situação, porque os documentos vazados no *Facebook Files* denotam que a plataforma tinha plena consciência de que negar o Holocausto é ilegal em quatorze países, mas julgou necessário apenas fazer um bloqueio – o geobloqueio – restrito a quatro deles, porque apenas França, Alemanha, Israel e Áustria ofereciam riscos apreciáveis à continuidade de suas atividades, por isso, geravam possibilidades concretamente contrárias a esse “*governo*” privado.

A ideia de governança, nesse sentido, amolda-se bem àquilo que as plataformas fazem quando moderam conteúdo porque ao mesmo tempo em que é capaz de expressar o fato de que elas exercem poder sobre os seus usuários, ela dispensa comprometer-se com modelos pré-constituídos – como o de Estado. O ponto aqui, portanto, não é dizer que as plataformas são Estados – como há textos que cogitam a existência de um “*Facebookstão*”<sup>439</sup> –, mas sim que elas exercem uma atividade de governo semelhante à estatal, apreendendo-a, sob o conceito de governança, de uma maneira flexível o suficiente para abranger as suas especificidades. Ele permite, nesse sentido, um enfoque maior na prática social de governar – ou seja, de exercer o poder –, constituindo-a como um fenômeno difuso socialmente e, por isso, atribuível tanto a

---

<sup>438</sup> Cf. *supra*, nota n.º 160.

<sup>439</sup> Cf., quanto ao ponto, CHANDER, A. Facebookistan. *North Carolina Law Review*, v. 90 n. 5, pp. 1.807/1.844, 2012; e, quanto à posição de Klonick a respeito, KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. *Harvard Law Review*, v. 131, n. 6, 2017, nota n.º 125.



Estados, quanto a corporações – fala-se em governança corporativa – e às próprias plataformas.

Sob a perspectiva normativa, falar em um modelo de governança daquilo que as pessoas dizem nas plataformas da internet é útil porque, ao toma-las como instâncias independentes, é possível compreender melhor, por um lado, a dinâmica de estímulos aos quais elas estão sujeitas e, por outro, a forma como elas respondem a essas influências. Assim, muito daquilo que as plataformas fazem – e muito do que se espera delas, como por exemplo, uma maior eficiência no combate ao discurso de ódio – ganha uma nova luz. Isso porque, se as plataformas não são apenas uma base fática à qual aplicáveis as regulações dos Estados – como defende a posição diádica – mas um núcleo decisório independente, capaz de “*governar*” aquilo que seus usuários podem dizer, então é possível tentar entender as razões por trás das suas decisões, articulando aquilo que lhes move em uma ou outra direção – e, assim, até mesmo influenciá-las com isso. Assim, estabelecê-las como “*governantes*” relativamente independentes permite enxergar os seus pontos de dependência e, assim, compreender melhor em que direção caminham.

Um exemplo, que desenvolvo melhor no capítulo seguinte, pode ajudar a entender isso. Vistas sob essa perspectiva de independência, como “*governantes*” da liberdade de expressão alheia, é possível entender as decisões das plataformas nesse campo a partir do conjunto de influências às quais estão sujeitas. Talvez a mais evidente seja a econômica: plataformas são empresas, e, como tal, buscam o lucro – Facebook, Google e Twitter, por exemplo, são companhias abertas listadas na Bolsa de Nova York. Em particular, suas receitas provêm de mecanismos publicitários, que geram mais dinheiro à medida que mais usuários interagem com mais frequência nos ambientes que administram<sup>440</sup>.

Sob essa perspectiva, é possível lançar uma nova luz sobre a moderação de conteúdo, de modo que, ao invés de se procurar uma explicação para o controle que as plataformas exercem sobre os seus usuários na jurisprudência da Suprema Corte, faz mais sentido tentar estabelecer relações que expliquem essa forma de “*governo*” a partir dos interesses aos quais essas empresas estão sujeitas. No caso, como interessa a essas empresas atrair cada vez mais pessoas à plataforma – e que, uma vez nesse espaço, elas interajam entre si – faz sentido moderar conteúdo para estabelecer um ambiente livre de certos conteúdos que deixam os usuários desconfortáveis, sem, entretanto, afastá-los. Por

---

<sup>440</sup> Cf. *supra*, nota n.º 250.

isso, ao mesmo tempo em que não interessa à plataforma estabelecer um ambiente muito restritivo – porque isso distancia as pessoas e coíbe as suas manifestações – também não é positivo deixar que se estabeleça um ambiente em que conteúdos capazes de impactar pessoas mais sensíveis – como os que contêm violência ou pornografia – sejam livremente disponibilizados. Nesse sentido, a moderação de conteúdo – e, em específico, a adoção de políticas de conteúdo que, via de regra, permitem a manifestação dos usuários e apenas excepcionalmente as restrinjam – é iluminada, pela perspectiva pluralista, como um instrumento de “*governo*” que viabiliza economicamente as plataformas.

#### 4. CONCLUSÃO

Iniciei este capítulo com um objetivo: explicar como a moderação de conteúdo funciona. Fiz isso uma vez que, como tenho deixado claro ao longo deste trabalho, a opacidade desse mecanismo não é só uma característica, mas também é um problema, porque que dificulta a compreensão de situações concretas e impede que problemas e hipóteses de pesquisa sejam desenvolvidos com precisão.

Para resolver isso, propus, de início, uma sistematização quanto à forma como plataformas da internet – a exemplo do Facebook, Twitter e YouTube – regulam a liberdade de expressão via moderação de conteúdo. Para tanto, dois conceitos foram delimitados: o de plataformas da internet e o de moderação de conteúdo.

De acordo com o primeiro, plataformas são serviços e páginas na internet que têm três características – elas *(i)* hospedam, organizam e circulam o conteúdo compartilhado e as interações realizadas por seus usuários; *(ii)* sem que tenham produzido esse material; fazendo-o *(iii)* de acordo com uma infraestrutura própria, que processa os dados enviados pelos usuários para viabilizar essa atividade e lucrar.

De acordo com o segundo, moderação de conteúdo é um termo utilizado tanto sob uma perspectiva geral, quanto sob uma perspectiva específica. Sob uma perspectiva geral, ela é o instrumento utilizado pelas plataformas da internet para exercer algum nível de controle sobre as manifestações de seus usuários, dispondo sobre a forma como elas virão à luz, sem toma-las como suas. Sob uma perspectiva específica, moderar conteúdo é a atividade das plataformas de determinar, dentro da modulação do discurso de seus usuários, qual conteúdo deve ou não ser permitido veicular nos espaços que administram.

Sob a perspectiva específica, a moderação de conteúdo se desdobra em dos tipos de atividades exercidas pelas plataformas. Primeiro, as plataformas estabelecem um conjunto de regras – gerais e específicas, opacas e públicas – que fixam aquilo que é permitido veicular em seus respectivos ambientes, fazendo-o tanto de forma a orientar os seus usuários, quanto de modo a pautar o trabalho dos responsáveis pela análise do conteúdo publicado. Segundo, as plataformas estabelecem procedimentos adjudicatórios, pelos quais detectam, decidem e sancionam a conduta de seus usuários, no sentido de fazer valer as suas políticas de conteúdo no ambiente que administram.

Quanto a esse segundo ponto – foco deste trabalho – propus dividi-lo em três fases diferentes e subsequentes, que esquematizam como essa espécie de procedimento funciona: primeiro, as plataformas detectam o conteúdo, remetendo-o à moderação de conteúdo; depois, tomam uma decisão quanto a se ele viola ou não as normas que estabeleceram; e, por fim, se o caso, aplicam ao conteúdo ou ao próprio usuário uma sanção.

Na primeira fase, em que um conteúdo específico, dentre os que circulam na plataforma, é selecionado ser analisado pela moderação de conteúdo, há quatro classificações possíveis: de acordo com (i) a origem da qual provém a provocação à plataforma para que analise algo (extraplataforma / intraplataforma); (ii) o momento em que o conteúdo objeto da moderação é detectado (*ex ante*, preventiva / *ex post*, repressiva); (iii) o agente que detecta o conteúdo a ser moderado (detecção proativa, da própria plataforma / detecção reativa, de um usuário, que sinaliza uma publicação para revisão); e (iv) ao modo pelo qual o conteúdo é levado à análise da moderação de conteúdo (por meio da ação humana / por meio de mecanismos automatizados).

Quanto à segunda fase, em que uma decisão é firmada a respeito do conteúdo detectado anteriormente, ela pode ser classificada (i) de acordo com a forma pela qual essa decisão é tomada (de forma centralizada / de forma descentralizada); (ii) de acordo com quem toma essas decisões (de forma automatizada / de forma manual, por um moderador humano).

Por fim, pode-se compreender a terceira fase – em que, se o caso, uma sanção é aplicada pela plataforma em razão da violação de uma de suas normas – a partir (i) do objeto da sanção (o conteúdo publicado / o usuário que fez a publicação); e, dentro de cada uma dessas situações, por gradações (ii) que, no caso do conteúdo, vão da sua permanência com certos condicionantes à sua retirada; e, no caso do usuário, vão de proibições a que se publique por um tempo determinado até o banimento.

O quadro abaixo resume essas classificações, abrangendo desde o conceito geral de moderação de conteúdo, até a forma como as sanções podem ser divididas:

**Quadro 11 – A moderação de conteúdo nas plataformas da internet**

<b><u>MODERAÇÃO DE CONTEÚDO EM SENTIDO GERAL</u></b>	<b><u>MODERAÇÃO DE CONTEÚDO EM SENTIDO ESPECÍFICO</u></b>	Moderação de conteúdo como normatização	<b><u>Normas</u></b>	<b>Gerais</b>				<b>Específicas</b>					
			<b>Opacas</b>	- Modelo utilizado no início da moderação de conteúdo, em que diretivas opacas e gerais eram elaboradas para os moderadores.				- Modelo utilizado nos manuais elaborados para os moderadores de conteúdo atualmente.					
			<b>Públicas</b>	- Modelo utilizado nos termos de serviço e nos padrões de comunidade e atualmente.				- Modelo utilizado nos termos de serviço e nos padrões de comunidade atualmente.					
<b><u>MODERAÇÃO DE CONTEÚDO EM SENTIDO GERAL</u></b>	<b><u>MODERAÇÃO DE CONTEÚDO EM SENTIDO ESPECÍFICO</u></b>	Moderação de conteúdo como adjudicação	<b>Fase n.º 1: Detectar</b>				<b>Fase n.º 2: Decidir</b>				<b>Fase n.º 3: Sancionar</b>		
			<b>Critério</b>	<b><u>Origem</u></b>	Extraplatafor ma	<b>Classificações</b>	<b>Critério</b>	<b><u>Modo</u></b>	Centralizado	<b>Classificações</b>	<b>Abrangência</b>	<b><u>Consequências externas</u></b>	
					Intraplatafor ma				Descentralizado				
				<b><u>Momento</u></b>	Ex ante / preventiva			<b><u>Agente decisório</u></b>	Manual			<b><u>Consequências internas</u></b>	Voltadas ao conteúdo
					Ex post / repressiva				Automatizado				Voltadas ao usuário
<b><u>Tipo de Agente</u></b>	Proativa												
<b><u>Modo</u></b>	Reativa												
		Ação humana											
		Ação automatizada											
<b><u>OUTRAS PRÁTICAS</u></b>	Conjunto de práticas de modulação do discurso alheio feito pelas plataformas e que não se enquadram na atuação direta de decidir o que é permitido ou não dizer no ambiente por ela administrado, a exemplo de escolhas estruturais quanto ao formato, local e limites dentro dos quais as pessoas podem se manifestar.												

Na segunda parte do capítulo, apresentei subsídios para construir uma esquematização que refletisse a forma como a literatura especializada vê a moderação de conteúdo. Reconstruí, nesse sentido, as perspectivas que dominam o debate a esse respeito: de um lado, reverberando a visão diádica da regulação da liberdade de expressão – que a vê exclusivamente como uma relação estabelecida entre um Estado regulador e um indivíduo regulado –, há uma primeira corrente, que analisa a moderação de conteúdo a partir da jurisprudência da Suprema Corte dos Estados Unidos. Em síntese, sua estratégia é estabelecer comparações entre a moderação de conteúdo e categorias jurisprudenciais, de forma a identificar a que melhor se aproxima com ela e, dessa forma, revelar o regime jurídico que lhe seria aplicável.

De outro lado, e a partir de uma crítica voltada tanto à qualidade dessas analogias, quanto à sua utilidade, há uma segunda corrente, que reverbera a visão pluralista da regulação da liberdade de expressão – e, por isso, vê na moderação de conteúdo uma instância regulatória em si, mesmo que, em uma de suas partes, não atue o Estado. De acordo com essa perspectiva, a moderação de conteúdo seria melhor entendida tanto descritivamente quanto normativamente como uma forma de governança das plataformas em relação àquilo que dizem os seus usuários, lançando luz sobre a forma como essas estruturas são constituídas e sobre a maneira pela qual elas podem ser influenciadas para que se transformem, por exemplo, em instrumentos mais eficientes para coibir a atuação deletéria de algumas pessoas na internet.

A taxonomia que proponho acima, ao organizar melhor as informações que se encontram dispersas na literatura, em investigações jornalísticas e entrevistas conferidas por pessoas-chave da moderação de conteúdo das plataformas da internet, tem duas utilidades. A primeira, que ficará mais evidente no próximo capítulo, é a de lançar luz sobre a opacidade característica – e problemática – da moderação de conteúdo para viabilizar que problemas e hipóteses de pesquisa sejam construídos com mais precisão. É difícil, nesse sentido, problematizar aquilo que não se conhece e, nesse sentido, tanto este capítulo, quanto o anterior, tiveram como respectivos objetivos construir uma estrutura que tanto justifica estudar a moderação de conteúdo quanto sistematiza o seu funcionamento e o modo como ele tem sido compreendido, abrindo as portas para avançar no sentido de abordagens mais consequentes quanto a esse fenômeno decisório.

A segunda utilidade é mais concreta: a sistematização acima ajuda a entender o que ocorreu em casos nos quais, tal como os que tenho trabalhado ao longo deste trabalho, atuaram as plataformas atuaram para regular aquilo que as pessoas podem

dizer. Isso porque, de início, ela dá uma dimensão da complexidade que há por trás da moderação de conteúdo: o que, à primeira vista, parece uma decisão arbitrária que traz perplexidade a pessoas como Dov Lipman e Tom Egeland, é, na verdade, o resultado de um procedimento profissionalmente organizado, construído a partir de escolhas estrategicamente realizadas pelas plataformas da internet. Sistematizar essas possibilidades e decisões, nesse sentido, esclarece não apenas os fatos por trás de uma posição determinada, mas também em que medida opções feitas pela plataforma quanto à forma como se modera conteúdo geraram situações problemáticas, como as que têm sido citadas ao longo deste trabalho.

O caso Lipman, com o qual este trabalho começou, é um exemplo disso. Aliado o relato que temos disponível – o de um parlamentar que, para testar a coerência do Facebook, trocou as palavras de uma publicação favorável aos israelenses e que tinha sido banida, tornando-a pró-Palestina e, por isso, admissível na plataforma – às classificações feitas acima, é possível esclarecer o que provavelmente aconteceu. Nesse sentido, é possível depreender, com um grau razoável de certeza, que a postagem feita pela página israelense foi (i) detectada por uma provocação interna, *ex post* – ela já tinha sido veiculada no Facebook –, de forma reativa – a questão não é abrangida pelos temas que são objeto de ações proativa das plataformas – e via atuação humana – como, pelo que se sabe, a publicação era inédita, mecanismos automatizados dificilmente poderiam ser configurados para interpretar o texto na forma em que veiculado na página. Depois, a compatibilidade do conteúdo com os Padrões da Comunidade do Facebook foram objeto de uma decisão (ii), tomada de forma centralizada – foram moderadores do Facebook que decidiram retirá-la de toda a plataforma – e manual – pelas mesmas razões que exigiram uma detecção manual. Por fim, aplicou-se uma sanção (iii), tanto sobre conteúdo, retirando-o, quanto sobre o usuário, impondo a proibição de que novas publicações fossem realizadas por três dias.

O cerne do caso, entretanto, não eram propriamente os fatos por trás da exclusão da imagem, mas sim a incoerência dessa postura com a manutenção, pouco tempo depois, de uma publicação semelhante – que provavelmente passou pelo mesmo procedimento, com exceção da fase sancionatória. Também aqui a classificação é útil: tanto a remoção da primeira imagem quanto a manutenção da segunda foram decisões intraplataforma, *ex post*, reativas e humanas, o que significa que a controvérsia se originou em uma falta de padronização nas decisões tomadas pelos moderadores de conteúdo do Facebook, algo que pode ser solucionado tanto com uma melhor especificação normativa –

via regras específicas e opacas direcionadas aos funcionários do Facebook, que toquem particularmente a relação entre palestinos e israelenses –, quanto por mecanismos de revisão automática do conteúdo moderado, algo já instituído pela plataforma, inclusive como uma forma de análise do desempenho contratual de empresas terceirizadas.

Essa mesma lógica pode ser aplicada aos demais casos que mencionei ao longo da primeira parte deste trabalho, e mesmo em outros que podem vir a surgir, em um movimento que pode ajudar a diferenciar as controvérsias que surgem a partir da moderação de conteúdo. Assim, por exemplo, a divergência entre moderadores humanos, que ocasionou o caso Lipman, não é uma causa que deve ser cogitada para analisar as reclamações provenientes de usuários do YouTube, de que documentários sobre o nazismo foram deletados de suas contas por violarem a nova política de discurso de ódio da plataforma<sup>441</sup>. Como, nesse caso, a detecção foi feita de forma automática, o problema que há por trás dessa controvérsia não se traduz em uma questão de treinamento de moderadores, mas sim nas dificuldades que os sistemas utilizados pelas plataformas têm de entender contextos linguísticos: é provável que o software tenha sido configurado para buscar símbolos e palavras associadas ao nazismo – os quais se fazem presentes tanto em vídeos neonazistas, quanto em documentários.

Da mesma forma, os problemas identificados nesses dois casos não se confundem com aqueles que aparecem na controvérsia envolvendo os nomes reais de *drag queens*, cujo cerne não era o modo de detecção, mas sim o tipo de agente que era responsável por isso: como se tratava, no caso, de uma forma de moderar conteúdo baseado em um *flagging* reativo, deveriam ter sido tomados cuidados maiores quanto às denúncias encaminhadas pelos usuários, uma vez que elas podem ser – como o foram – manipuladas. Identificar, portanto, em qual nicho se encaixa um determinado caso, ajuda a entender onde exatamente surgiu a controvérsia que lhe deu origem e, mais importante, como essa controvérsia se articula com outros casos, seja se aproximando daqueles que se inserem na mesma categoria, seja se distanciando dos demais.

Com a conclusão desta seção, encerro a primeira parte deste trabalho. Ela se dedicou, em síntese, a tanto justificar a elaboração de um trabalho científico dedicado à moderação de conteúdo; quanto a explicar como esse mecanismo funciona. Quanto ao primeiro ponto, como argumentei no primeiro capítulo, estudar moderação de conteúdo é algo que faz sentido porque, sem fazê-lo, deixamos de compreender uma parte

---

<sup>441</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.



fundamental dos impactos que a era digital trouxe para a regulação da liberdade de expressão – não a criação apenas de novos *espaços de manifestação*, mas o surgimento de novos *espaços de decisão*, que determinam, com relativa independência – e em paralelo aos Estados – aquilo que as pessoas podem dizer. Quanto ao segundo, objeto deste capítulo, proporcionei uma taxonomia que se dedicou tanto a conceituar, quanto a sistematizar a moderação de conteúdo, lançando luz sobre esse mecanismo de forma a, por um lado, viabilizar uma maior clareza na sua problematização; e, por outro, expor aquilo que está por trás da opacidade que lhe é característica – e problemática – em casos concretos.

Na parte seguinte, proponho um trabalho de problematização e reinterpretação da moderação de conteúdo, tomando como base as sistematizações construídas nestes dois primeiros capítulos e as transformações pelas quais esse mecanismo tem passado recentemente.

## PARTE II | POR UMA REINTERPRETAÇÃO

### CAPÍTULO 3 – PROBLEMATIZANDO A MODERAÇÃO DE CONTEÚDO

#### 1. INTRODUÇÃO

Este é um trabalho dividido em duas partes. Fiz isso para organizá-lo de acordo com os problemas que ele enfrenta. Assim, a primeira parte se dedicou a oferecer subsídios para fazer frente à característica – e problemática – opacidade da moderação de conteúdo das plataformas da internet. Abordei isso em duas etapas: primeiro, oferecendo razões que justifiquem escrever uma dissertação sobre regulação da liberdade de expressão que tematize a moderação de conteúdo (Capítulo 1). Segundo, obtendo das fontes que temos disponíveis – investigações jornalísticas, literatura especializada e entrevistas colhidas com pessoas que trabalham com esse mecanismo – informações que formassem um quadro de conceitos e classificações capazes de explicar como ela funciona (Capítulo 2).

Como expus na conclusão do capítulo anterior, fazer isso tem duas utilidades. A primeira é compreender melhor as controvérsias causadas pelas decisões que, tomadas pelas plataformas da internet, definem aquilo que as pessoas podem ou não dizer na internet. A segunda é permitir que problemas e hipóteses de pesquisa sejam construídos com mais precisão, o que exige uma base de conhecimento que não existe quanto à moderação de conteúdo. Como tenho insistido no decorrer deste trabalho, a moderação de conteúdo é um mecanismo localizado em estruturas empresariais e, por isso, que não está sujeito aos mesmos protocolos de transparência das formas estatais de regular a liberdade de expressão. Por isso, não há uma consolidação única – um manual, um código ou um compilado de precedentes – capaz de fornecer um retrato de como ela funciona. Construir esse retrato, nesse sentido, é um pressuposto para que perguntas e respostas mais consequentes a respeito desse mecanismo sejam feitas, razão pela qual a ele dediquei a primeira parte da dissertação.

Atravessada essa etapa, a segunda parte do trabalho propõe uma problematização mais complexa, derivada tanto das conclusões do capítulo anterior, a respeito de como a moderação de conteúdo funciona, quanto das limitações detectadas nas perspectivas que a explicam. Entretanto, diferente da opacidade, um problema autoevidente – constatável, por exemplo, pela perplexidade dos usuários ao terem

manifestações suas excluídas de plataformas como o Facebook – e que pode ser enfrentado por uma abordagem relativamente simples – a reconstrução do que se tinha disponível a respeito da moderação de conteúdo – o problema e o método que proponho a seguir demandam uma explicação mais detalhada, à qual dedico este capítulo.

Sustento, a seguir, que a moderação de conteúdo em sentido estrito passa por uma transformação tão profunda quanto inexplicável em sua atividade de adjudicar condutas (2). Argumento, a esse respeito, que as plataformas têm deixado paulatinamente para trás um modelo de decidir arbitrário – pelo qual estabeleceram procedimentos que melhor atendiam à necessidade de manter controle sobre o que é veiculado no ambiente que administram – para adotar formatos decisórios próximos ao do Estado, inculcando, nesse procedimento – mesmo que inexistam regulamentações ou precedentes estatais as obrigando a tanto –, mecanismos típicos do devido processo legal.

O problema é que, ao mesmo tempo em que esse movimento tem transformado profundamente a moderação de conteúdo, passando a moldar a forma como decisões quanto àquilo que as pessoas podem dizer na internet são tomadas, ele não tem explicação na literatura disponível. Suas causas e consequências não são esclarecidas seja pelas perspectivas diádicas, dedicadas a compreender o regime jurídico aplicável às plataformas a partir de um diálogo com a Primeira Emenda; seja pelas perspectivas pluralistas disponíveis, que identificam na moderação de conteúdo uma forma de governança privada. Nem uma, nem outra, consegue explicar por que as plataformas, depois de construírem métodos decisórios que garantissem a elas um amplo – e compreensível – controle sobre aquilo que as pessoas dizem nos ambientes que administram, resolveram abrir mão desse poder, absorvendo, em seus modelos procedimentais decisórios, mecanismos do devido processo legal dos Estados.

No final da seção, apresento duas hipóteses para resolver esse problema. Proponho, em síntese, que a transformação identificada nas plataformas é explicada como um movimento que reflete tanto a busca dessas empresas por legitimidade em suas decisões, quando, em quadro maior, o surgimento de uma forma de constitucionalização no mundo digital.

Em seguida, argumentarei, como método (3), a utilização de perspectivas sociológicas jurídicas robustas para abordar o problema (3.1), dentre as quais adoto a Teoria dos Sistemas de Niklas Luhmann (3.2), estabelecendo, concomitantemente, uma estratégia para abordar a questão, dividindo-a em dois capítulos: o primeiro (Capítulo 4), em que o devido processo legal é explicado enquanto um mecanismo estatal; e um segundo

(Capítulo 5), em que ele é abordado como um instrumento utilizado pelas plataformas. No final, encerro o capítulo sintetizando esse planejamento e o vinculando com o objetivo do próximo capítulo (4).

## 2. PROBLEMA E HIPÓTESES

A moderação de conteúdo das plataformas da internet é um tema tanto novo quanto dinâmico: como parte da realidade da era digital, a forma como as plataformas decidem aquilo que os seus usuários podem ou não dizer na internet muda constantemente. À medida que isso acontece, ela desafia os modelos que temos disponíveis para explicá-la, e por isso, exige deles uma complexidade cada vez maior. O problema que apresento nesta parte do trabalho diz respeito a uma dessas mudanças e à forma como ela leva, ao seu limite, as forças das perspectivas que temos disponíveis para entender a moderação de conteúdo.

Na primeira parte, expus a moderação de conteúdo tanto sob uma perspectiva geral – atrelada a todas as decisões tomadas pelas plataformas desde a sua concepção e que impactam, de alguma forma, no modo como os seus usuários se manifestam – quanto sob uma perspectiva específica – constituída pelas decisões tomadas pelas plataformas quanto àquilo que os seus usuários podem ou não dizer. Nesse último ponto, identifiquei duas vertentes de atuação: na primeira, as plataformas dispõem padrões gerais de comportamento; na segunda, elas aplicam esses padrões, construindo um procedimento adjudicatório. É nessa última vertente que irei me concentrar – isto é, o viés adjudicatório da moderação de conteúdo em sentido estrito, ao qual passarei a me referir, a seguir, simplesmente como “*moderação de conteúdo*”.

No capítulo anterior, essa vertente da moderação de conteúdo foi apresentada a partir de uma taxonomia, em que sistematizadas as diferentes formas pelas quais ela ocorre de acordo com as suas três fases: é possível, por exemplo, *detectar* conteúdo por mecanismos automáticos ou humanos; *decidir* quanto a ele por determinações *ex ante* e *ex post*; e o *sancionar* gerando consequências tanto sobre a publicação quanto sobre os usuários que a disponibilizaram. Detectar essa variabilidade é importante, porque ela expõe que não há uma forma única de se definir o que as pessoas podem dizer: a moderação de conteúdo não é uma sequência necessária de fatos que obedece a uma lógica fixa, mas o resultado de uma série contingente de decisões tomadas pelas plataformas, que a amoldam segundo aquilo que entendem ser mais conveniente.

Essas decisões, entretanto, não dizem respeito apenas ao conteúdo que permanece ou não na plataforma, mas atinem, também, à forma como isso é definido. Elas podem, por isso, ser divididas em duas categorias: a das decisões de caráter *material* e a das decisões de caráter *procedimental*.

Por um lado, as plataformas tomam decisões de caráter *material*: elas definem *aquilo que os seus usuários podem ou não publicar*. Isso toma forma concreta quando elas precisam responder a perguntas como, por exemplo, se a imagem de uma criança fugindo de um ataque de napalm viola a política de nudez do Facebook; se a notícia falsa publicada por Alex Jones deve permanecer ativa no YouTube; ou se uma página que afirma que o Holocausto não existiu deve ou não ser retirada daquele ambiente.

Por outro, as plataformas tomam decisões de caráter *procedimental*: elas definem *o procedimento que determina se uma publicação viola ou não as suas políticas*. Concretamente, elas precisam responder a perguntas como, por exemplo, se um usuário será ouvido antes de ter uma publicação sua excluída; se ele poderá saber as razões que levaram à plataforma a agir dessa forma; ou se é possível ou não recorrer dessa decisão.

Como disse antes, a variabilidade de respostas a essas perguntas – exposta na extensão da taxonomia do capítulo anterior – é um dado importante, porque revela o pensamento estratégico por trás da configuração, pelas plataformas, de estruturas de moderação de conteúdo. Elas não adotam uma forma única de decidir: antes, constroem aquela que melhor se afeiçoa à sua atividade, a partir de uma espécie de cardápio de modelos decisórios, algo coerente com o papel essencial que a moderação de conteúdo ocupa na atividade que exercem e o poder que tem sobre ela. Como afirma Gillespie, plataformas não são plataformas se não fazem alguma forma de moderação de conteúdo<sup>442</sup>; a isso, o diagnóstico do capítulo anterior acrescenta que plataformas não são plataformas se não fazem alguma forma *estratégica* de moderação de conteúdo, aqui compreendida a ação de modelá-la segundo aquilo que melhor lhe convêm, ajustando cada detalhe dessa forma de decidir a partir de opções feitas dentre um conjunto de possibilidades de detecção, decisão e sanção. Como exponho a seguir, essa contingência decisória fica ainda mais clara quando distinções entre as decisões de caráter material e procedimental são trazidas a claro.

Assim, para fornecer respostas às questões de caráter *material* – isto é, para definir aquilo que os seus usuários podem ou não dizer no ambiente que administram –, as

---

<sup>442</sup> Cf. *supra*, nota n.º 259.

plataformas construíram assumidamente um modelo inspirado naquilo que os Estados já faziam, baseando-se em um paradigma democrático de liberdade de expressão – concretizado, em seu formato mais conhecido, na Primeira Emenda.

Com isso, não quero dizer que a jurisprudência da Suprema Corte e as decisões das plataformas coincidem, nem que elas tomaram a Primeira Emenda como uma cartilha a ser seguida às últimas consequências. A primeira parte deste trabalho expôs como essas afirmações não têm respaldo na realidade: a Primeira Emenda protege discursos que são muitas vezes vedados nas plataformas<sup>443</sup>. Elas incorporaram, entretanto, na moderação de conteúdo em sentido estrito, o que Klonick chama de uma “*filosofia corporativa da liberdade de expressão*”<sup>444</sup>: um quadro geral de atuação, sustentado no modelo do Estado, segundo a qual, via de regra, os usuários devem ser livres para expressar aquilo que quiserem nesses ambientes. Apenas excepcionalmente, a partir de certas circunstâncias específicas – como as que ameaçam a segurança das pessoas – é que as plataformas devem intervir, modulando o discurso alheio de modo a restringi-lo ou, em casos extremos, excluí-lo daquele ambiente. Como Klonick colhe a partir de entrevistas, esse não foi um alinhamento automático ou acidental, mas conscientemente adotado pelo YouTube – “*No YouTube, Wong procurou pelos valores da companhia em adição aos das normas de liberdade de expressão dos Estados Unidos no desenvolvimento de sua abordagem à moderação de conteúdo*”<sup>445</sup> – no Facebook – “*Como Wong, Hoffman e Willner consideraram que a missão do Facebook – ‘fazer o mundo mais aberto e conectado – e compreenderam que ela está frequentemente alinhada com valores mais amplos de liberdade de expressão e democracia dos Estados Unidos*”<sup>446</sup> – e no Twitter –

---

<sup>443</sup> Cf. *supra*, notas n.º 423 e 424.

<sup>444</sup> “*Twitter’s corporate philosophy of freedom of speech [...]*” (“*Da mesma forma, a filosofia corporativa da liberdade de expressão do Twitter [...]*”). KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **HARVARD LAW REVIEW**, v. 131, n. 6, 2017, p. 1.626.

<sup>445</sup> “*At YouTube, Wong looked to the values of the company in addition to American free speech norms in developing an approach to content moderation*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **HARVARD LAW REVIEW**, v. 131, n. 6, 2017, p. 1.626.

<sup>446</sup> “*Like Wong, Hoffman and Willner considered the mission of Facebook — ‘to make the world more open and connected’ — and found that it often aligned with larger American free speech and democratic values*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **HARVARD LAW REVIEW**, v. 131, n. 6, 2017, p. 1.626.

*“Da mesma forma, a filosofia corporativa do Twitter quanto à liberdade de expressão justificou seu fracasso em moderar conteúdo”*<sup>447</sup>.

A adoção de um quadro geral de defesa democrática da liberdade de expressão, alinhado com o Estado e repetido à exaustão pelos executivos do Facebook<sup>448</sup>, Google<sup>449</sup> e Twitter<sup>450</sup>, entretanto, não é o reflexo de uma postura de identificação mútua de valores – ou, ao menos, essa não é a única razão para fazê-lo –, mas algo que se mostra útil às plataformas. Isso fica claro quando a dinâmica desse quadro geral – estabelecida por (i) uma regra geral de não intervenção; e (ii) uma autorização excepcional para fazê-lo – é contextualizada com a realidade econômica das empresas que estão por trás desses ambientes, observando-se que essas formas de decidir também são explicadas pela forma como essas companhias lucram.

Assim, de um lado – e quanto à vertente desse discurso que estabelece, como regra, a ausência de intervenção –, as plataformas são empresas que lucram à medida que captam usuários e estabelecem estímulos para que eles interajam no ambiente que administram – e, com isso, gerem recursos, seja adquirindo planos diferenciados de acesso, seja atraindo anunciantes<sup>451</sup>. Por isso, não surpreende que elas tenham adotado como regra a liberdade de expressão: aqui, ela é instrumentalizada como uma forma de introduzir,

---

<sup>447</sup> *“Likewise, Twitter’s corporate philosophy of freedom of speech justified its failure to moderate content”*. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.626.

<sup>448</sup> Cf., por exemplo, as declarações de executivos do Facebook ao reestabelecer a foto de Egeberg, no sentido de que o objetivo da plataforma é *“melhorar nossas políticas para garantir que elas promovam a livre expressão e mantenham a comunidade segura”*, *supra*, Cap. I.1, 3.2.2.1 e nota n.º 128.

<sup>449</sup> Cf., por exemplo, a declaração segundo a qual *“If we were to take all potentially offensive content down, we’d be losing valuable speech”* (*“Se nós tirarmos todo o conteúdo potencialmente ofensivo, nós estaríamos perdendo manifestações valiosas”*). DALE, C. Taking a harder look at harassment. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/taking-harder-look-at-harassment.html>>, acesso em 30.7.2019

<sup>450</sup> Cf., por exemplo, o op-ed escrito por Vijaya Gadde, executiva do Twitter, em que ela afirma que a essência da plataforma coincide com a da liberdade de expressão, isto é, com a *“the idea that everyone has a voice and the right to use it”* (*“ideia de que todo mundo tem voz e tem o direito de usá-la”*). GADDE, V. Twitter Executive: Here’s How We’re Trying to Stop Abuse While Preserving Free Speech. **The Washington Post**, 16.4.2015, disponível em <[https://www.washingtonpost.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/?utm\\_term=.e8fd3bcd5193](https://www.washingtonpost.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/?utm_term=.e8fd3bcd5193)>, acesso em 30.7.2019.

<sup>451</sup> Cf. *supra*, nota n.º 250.

dentro desses ambientes, uma lógica de atuação que favorece a permanência e a interação entre os seus usuários – e, portanto, a maximização dos lucros desses agentes.

De outro – e quanto à outra vertente, que permite, excepcionalmente, a intervenção das plataformas naquilo que dizem seus usuários – ela justifica o viés restritivo da moderação de conteúdo, porque, se, por um lado, “*manter todo o conteúdo [publicado] na plataforma gera o risco de fazer com que os usuários se sintam desconfortáveis, perdendo, com isso, visualizações e lucro*”<sup>452</sup>; por outro, se a plataforma “*retira muito conteúdo, ela gera perdas não apenas de oportunidades de interação, mas também a confiança potencial de seus usuários*”<sup>453</sup>. Por isso, esse alinhamento material entre plataformas e Estado só acontece a um nível geral: se a elas interessa ter a liberdade de expressão como uma “*filosofia corporativa*”, que deixa as pessoas livres para se manifestarem, também a elas interessa ter um razoável grau de discricionariedade para permitir e proibir aquilo que lhes convém nesse ambiente, de modo a encontrar um “*ponto ideal*” regulatório, em que nem as pessoas se sintam desconfortáveis com o que veem, e, por isso, saiam da plataforma; nem se sintam constantemente coibidas em suas manifestações, o que poderia fazer com que procurassem outro lugar para se expressar. Esse alinhamento material, portanto, é mais uma convergência de interesses do que um compartilhamento de valores: ele tanto se justifica quanto se baliza pelo empenho das plataformas em atrair e manter o maior número possível de usuários, construindo estímulos à sua interação.

Esse caráter seletivo e estratégico das decisões que determinam a moderação de conteúdo fica ainda mais evidente quando as respostas das perguntas de caráter *material* são cotejadas com as de caráter *procedimental*. Aqui, as plataformas se afastam completamente da inspiração constatada no âmbito material, constituindo procedimentos que nada tem que ver com a forma como o Estado toma decisões a respeito da liberdade de expressão. É como se a atenção dispensada pelas plataformas à Primeira

---

<sup>452</sup> “*Likewise, keeping up all content on a site risks making users uncomfortable and losing page views and revenue*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.627.

<sup>453</sup> “[...] *users will spend more time on the site and advertising revenue will increase. Take down too much content and you lose not only the opportunity for interaction, but also the potential trust of users*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.627.



Emenda não tivesse avançada à Quinta e à Décima Quarta Emendas<sup>454</sup>, nas quais consta a cláusula do devido processo legal<sup>455</sup>. Não existe, como afirma, Danielle Citron um “*devido processo legal tecnológico*”<sup>456</sup>, sequer no sentido do quadro geral que há em relação à Primeira Emenda nas questões materiais. Isso fica claro quando as características que marcam, de forma transversal, todas as fases do procedimento de moderar conteúdo, são colocadas em evidência, lado a lado àquelas que distinguem o modo como o Estado resolve as mesmas controvérsias no âmbito judicial, como tem sido feito por movimentos sociais que exigem uma transformação no modo como as plataformas moderam conteúdo.

O exemplo mais consolidado nesse sentido<sup>457</sup>, e que tomo como base para expor essa situação, é o expresso nos *Santa Clara Principles On Transparency and*

---

<sup>454</sup> A Quinta Emenda estabelece a cláusula de forma geral; a Décima Quarta afirma expressamente que ela é aplicável para os Estados que compõem a federação. Ambas têm a mesma redação, dispondo que o Estado não pode “*deprive any person of life, liberty, or property, without due process of law*” (“*privar ninguém da vida, liberdade ou propriedade, sem o devido processo legal*”). O dispositivo é praticamente idêntico ao art. 5.º, LIV, da Constituição Federal do Brasil (“*ninguém será privado da liberdade ou de seus bens sem o devido processo legal*”).

<sup>455</sup> A cláusula do devido processo legal é desdobrada na jurisprudência dos Estados Unidos em dois vieses. Sob um viés substantivo, ela protege direitos não relacionados ao procedimento, como o direito à liberdade. Nesse sentido, o caso *Lochner v. New York*, 198 U.S. 45 (1905), em que uma lei do Estado de Nova York que limitava o número de horas de trabalho de padeiros foi declarada inconstitucional com fundamento na Décima Quarta Emenda, iniciando a “*Lochner Era*”. Sob um viés procedimental, ela exige dos agentes do Estado a obediência a um conjunto de regras procedimentais antes que atos que impactem a vida, a liberdade ou a propriedade de alguém sejam praticados. *Cf.*, a esse respeito, a consolidação dessas regras – que cito a seguir, nas notas n.º 464, 465 e 467, para traçar paralelos com as atividades das plataformas – em FRIENDLY, H. Some Kind of Hearing. **University of Pennsylvania Law Review**, v. 123, n. 6, pp. 1.267/1.317, 1975. No texto, as referências à cláusula devem ser entendidas como remissões ao seu viés procedimental.

<sup>456</sup> O termo é utilizado para se referir aos problemas de devido processo legal ocasionados pela utilização de ferramentas tecnológicas para que a administração pública realize atividades burocráticas, como conceder benefícios às pessoas. *Cf.* CITRON, D. Technological Due Process. **Washington University Law Review**, v. 85, n. 6, pp. 1.249/1.313, 2008.

<sup>457</sup> Há outros casos em que subsiste esse tipo de exigência, como, por exemplo, na carta-aberta composta pelo editor do jornal em que Egeland trabalhava (*cf. supra*, Cap. I.1, 3.2.2.1). A utilidade dos *Santa Clara Principles* é que eles consolidam essas exigências em um elenco recente de princípios, discutido com os principais atores desse movimento, algo evidenciado pela sua menção no documento em que o Facebook consolida as mudanças recentemente implantadas na moderação de conteúdo da plataforma (*cf. infra*, nota n.º 491).

*Accountability in Content Moderation*<sup>458</sup>. O documento surgiu a partir de duas reuniões, ocorridas no início de 2018, entre organizações – dentre as quais, por exemplo, a União Americana pelas Liberdades Civis –, advogados e acadêmicos que “*apoiam o direito à liberdade de expressão online*” e se propuseram a pensar “*como melhor obter mais transparência e responsabilidade [Accountability] na moderação de conteúdo gerado por usuários, que é realizada cada vez de forma mais agressiva pelas plataformas da internet*”. Em particular, o objetivo era propor soluções a serem adotadas pelas plataformas “*para que elas estabeleçam o devido processo legal [“due process”] para aqueles que se manifestam e são impactados [por suas decisões]*”<sup>459</sup>. As reuniões geraram dois documentos. Primeiro, uma carta-aberta ao Facebook, sustentando que a plataforma “*está muito atrás de seus competidores quando o assunto é garantir aos seus usuários o devido processo legal [“due process”]*”<sup>460</sup>. Segundo, os *Santa Clara Principles*, “*passos iniciais*” a partir dos quais se propôs uma mudança no procedimento utilizado pelas plataformas para moderar conteúdo, de modo a garantir, aos seus usuários, um formato decisório semelhante ao do “*devido processo legal*”:

– Princípio n.º 1, “*Números*”: “*As empresas devem publicar o número de publicações removidas e de contas temporariamente ou permanentemente suspensas em razão de violações de suas diretrizes de conteúdo*”<sup>461</sup>. O objetivo desse princípio é aumentar a transparência na moderação de conteúdo, viabilizando estudos quantitativos e

---

<sup>458</sup> THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 29.7.2019.

<sup>459</sup> “*who support the right to free expression online was convened to consider how best to obtain meaningful transparency and accountability around internet platforms’ increasingly aggressive moderation of user-generated content.*” [...] “*we propose these three principles as initial steps that companies engaged in content moderation should take to provide meaningful due process to impacted speakers [...]*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 29.7.2019.

<sup>460</sup> “*Facebook remains far behind its competitors when it comes to affording its users due process*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. An Open Letter to Mark Zuckerberg: The World’s Freedom of Expression is in Your Hands. Disponível em <<https://santaclaraprinciples.org/open-letter/>>, acesso em 29.7.2019.

<sup>461</sup> “*Companies should publish the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 29.7.2019.

qualitativos a respeito dessas decisões a partir de dados fornecidos pelas próprias plataformas.

– Princípio n.º 2, “Aviso”: “*As empresas devem fornecer um aviso a cada usuário cujo conteúdo é retirado ou cuja conta é suspensa sobre a razão pela qual houve a remoção ou a suspensão*”<sup>462</sup>. O objetivo desse princípio é garantir que o usuário possa participar do procedimento de moderação de conteúdo, em particular recebendo (i) um aviso que o notifique de uma ação da plataforma contra ele; e informações sobre (ii) qual conteúdo é proibido; (iii) como o procedimento de detecção funciona; (iv) quais fundamentos levaram à sanção; e (v) qual foi o resultado de uma publicação que ele reportou.

– Princípio n.º 3, “Apelação”: “*As empresas devem estabelecer uma possibilidade para oportuna apelação de qualquer remoção de conteúdo ou suspensão de conta*”, algo que abrange, no longo prazo “*processos de revisão externa e independente*”<sup>463</sup>. O objetivo desse princípio é garantir que a decisão da plataforma que impacta na liberdade de expressão de um usuário não seja definitiva, mas possa ser revista – pela própria empresa; ou, preferencialmente, por um agente externo.

O teor das exigências feitas pelas empresas à época já é, em si, um sinal de que a realidade da moderação de conteúdo nas plataformas da internet não envolve a aplicação de algo sequer parecido a uma cláusula do devido processo legal; no que segue, torno isso mais concreto analisando qual era a situação quanto a cada um desses princípios no momento em que eles foram apresentados.

O primeiro princípio tem conexão com a ideia de *transparência*. Em síntese, enquanto o paradigma do devido processo legal implica em algum grau de transparência quanto aos procedimentos que o Estado conduz<sup>464</sup> – em particular, há

---

<sup>462</sup> “*Companies should provide notice to each user whose content is taken down or account is suspended about the reason for the removal or suspension*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 29.7.2019. O nome do princípio na origem é “*Notice*”, isto é, literalmente “*Aviso*”, que traduzi como “*Participação*”, por melhor refletir a ideia por trás do princípio.

<sup>463</sup> “*Companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension*”. THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 29.7.2019.

<sup>464</sup> Cf., por exemplo, FRIENDLY, H. Some Kind of Hearing. **University of Pennsylvania Law Review**, v. 123, n. 6, 1975, pp. 1.280/1.281 e 1.291/1.292 (“2. *Notice of the Proposed Action and the Grounds Asserted for It*” e “9. *The Making of a Record*”).

protocolos pelos quais é possível se identificar esses processos e seus resultados – a moderação de conteúdo é marcada pela *opacidade*. Especificamente, no momento em que vieram à luz os *Santa Clara Principles*, não havia dados consolidados e disponíveis ao público que viabilizassem pesquisas quantitativas e qualitativas a respeito da moderação de conteúdo, nem que permitissem compreender qual era a real extensão da atuação das plataformas sobre o conteúdo veiculado nos ambientes que administravam. No limite, antes do biênio 2018 / 2019 – em que as transformações que relato a seguir ocorreram –, plataformas a exemplo do YouTube forneciam apenas relatórios parciais, que compreendiam apenas a retirada de conteúdo realizada pelas plataformas por ordens governamentais, representativas apenas de uma fração do que passa pelos procedimentos organizados por essas empresas e, no limite, refletem uma compreensão tão limitada quanto a do Balkin de 2014.

O *segundo princípio* tem conexão com a ideia de *participação*. Em síntese, o devido processo legal implica a possibilidade de que as partes que serão impactadas por uma decisão proferida pelo Estado participem da forma como ela é construída<sup>465</sup>. Refiro-me, aqui, a um conjunto de garantias que dizem respeito tanto a momentos anteriores ao processo – como o conhecimento prévio das regras que o regem e dos critérios que ele utiliza para proferir uma decisão (respectivamente itens *(ii)* e *(iii)*, acima) – quanto àquilo que acontece no processo – como o conhecimento de que há um processo, de qual fase ele se encontra, e de quais são os fundamentos de uma decisão que foi proferida nele (itens *(i)*, *(iv)* e *(v)*, acima).

Inexiste isso nas plataformas, marcadas que são pela *ausência de participação*: como a perplexidade experimentada por pessoas como Tom Egeland demonstra – que acessou a plataforma e descobriu, por uma mensagem genérica, que uma publicação sua tinha sido excluída e que sanções lhe foram aplicadas –, os usuários são espectadores, e não partícipes do procedimento de moderação de conteúdo. Isso se revela em, pelo menos, três níveis diferentes: primeiro, na ausência de normativas que definam e publicizem como decisões são tomadas pela moderação de conteúdo. Não há uma sistematização normativa processual – tal como existem, nos Estados, códigos processuais e como emergiram, sob a perspectiva material, consolidações como os Padrões da

---

<sup>465</sup> Cf., por exemplo, FRIENDLY, H. Some Kind of Hearing. *University of Pennsylvania Law Review*, v. 123, n. 6, 1975, pp. 1.281, 1.282/1.287 (“3. An Opportunity to Present Reasons Why the Proposed Action Should Not Be Taken”, “5. To Know the Evidence Against One”).

Comunidade do Facebook – que explique e delimite os passos que são percorridos desde a detecção do conteúdo até a sua exclusão. Como a investigação da primeira parte revelou, as informações que temos a esse respeito advêm, basicamente, de vazamentos, entrevistas, e conclusões tiradas de controvérsias que vêm a público. Além disso, como expliquei com mais detalhes na primeira atividade afeta à moderação de conteúdo em sentido estrito<sup>466</sup>, a maior variabilidade normativa utilizada pelas plataformas não significou a exclusão de modelos anteriores, de modo que ainda há normas de teor material ocultas aos usuários.

Segundo: não há, via de regra, mecanismos que permitam saber que existe um procedimento de moderação de conteúdo em andamento e, muito menos, vias que possibilitem acompanhá-lo, tanto na qualidade de “*parte*” – ou seja, de alguém que publicou algo que está sob análise –, quanto na qualidade de “*denunciante*” – ou seja, de alguém que reportou uma publicação alheia feita na plataforma.

Terceiro: não há uma forma sistemática de se conhecer as razões que fundamentam as decisões tomadas pelas plataformas. Isso não significa que nunca sabemos esses fundamentos – como no caso do Holocausto, eles podem aparecer de forma acidental, em uma entrevista, uma publicação no Twitter ou em uma mensagem à imprensa –, mas sim que não há uma sistemática que, como o que exige o devido processo legal, garanta sempre a quem é impactado por uma decisão o conhecimento das razões que a fundamentaram. A perplexidade de Lipman tem a ver com isso: ninguém sabe por que a primeira publicação foi excluída; mas também ninguém sabe exatamente por que ela foi reestabelecida – a mensagem genérica do Facebook se resumia a reconhecer um “*erro*”.

O *terceiro princípio* tem conexão com a ideia de *recorribilidade* – uma especificação da ideia de participação, tratada no princípio anterior, que é veiculada em um item à parte dos *Santa Clara Principles*. Enquanto no devido processo legal há sistemáticas pelas quais a controvérsia possa ser analisada novamente, por um julgador diferente e com poder suficiente para reverter a decisão anterior<sup>467</sup>, a moderação de conteúdo é caracterizada também aqui – exceção feita ao YouTube, que estabeleceu um mecanismo para tanto em 2010<sup>468</sup> – pelo contrário, ou seja, pela *irrecorribilidade*. É sintomático, a esse respeito, que as fases nas quais consolidei o procedimento adjudicatório

---

<sup>466</sup> Cf. *supra*, Cap. I.2, 2.3.1.

<sup>467</sup> Cf., por exemplo, FRIENDLY, H. Some Kind of Hearing. **University of Pennsylvania Law Review**, v. 123, n. 6, 1975, pp. 1.279/1.280 e 1.294/1.295 (“*I. An Unbiased Tribunal*”, “*II. Judicial Review*”).

<sup>468</sup> WU, A. Strike you’re out! Or maybe not? **YouTube Official Blog**, 2.7.2010, disponível em <<https://youtube.googleblog.com/2010/07/strike-youre-out-or-maybe-not.html>>, acesso em 30.7.2019.

no capítulo anterior se encerram com a sanção: tomada a decisão, inexistem vias internas que permitam ao usuário sequer apresentar razões tendentes a revertê-la.

Mais uma vez, isso não significa que as decisões da plataforma nunca são revertidas: os dois casos que reconstruí no primeiro capítulo e nos quais houve uma mudança de posicionamento indicam exatamente isso, na medida em que os usuários tiveram que recorrer a meios externos à plataforma para conseguirem ser ouvidos. No caso Lipman, o parlamentar israelense construiu uma estratégia para mostrar a incoerência do Facebook e depois publicou uma carta aberta à plataforma, que, pressionada, decidiu mudar a decisão; no caso Egeland, todas as tentativas internas de se reverter a decisão – as publicações feitas por outros usuários – redundaram em novas exclusões, cenário que foi alterado apenas após a repercussão fora da plataforma ter atingido níveis significativos. O que inexistente, nesse sentido, são sistemáticas internas de recorribilidade pelas quais os usuários afetados por decisões proferidas pela plataforma possam acioná-la, expondo razões capazes de influenciar uma nova análise.

Juntas, *opacidade, ausência de participação e irrecorribilidade* – características da moderação de conteúdo – opõe-se, respectivamente à *transparência, participação e recorribilidade* – elementos básicos de qualquer sistema procedimental de decisão inspirado, mesmo que da forma mais genérica, em uma cláusula de devido processo legal. Analogamente à adoção de um paradigma semelhante ao do Estado, no caso das questões materiais, o distanciamento que se observa nas questões processuais não é um posicionamento automatizado, nem uma divergência de valores, mas uma decisão consciente, que reflete uma opção útil às plataformas.

Isso porque se a moderação de conteúdo é algo necessário e essencial às plataformas, como um instrumento de controle sobre aquilo que dizem os seus usuários, então nada é mais desinteressante para elas do que adotar um modelo decisório próximo ao do Estado. Nesse sentido, instituir um procedimento transparente e que garanta participação e recorribilidade significa, em última análise, diminuir o poder que as plataformas têm sobre o que os seus usuários dizem: expor, por exemplo, dados que revelem a quantidade e a qualidade das decisões proferidas pela moderação de conteúdo, abre um flanco para que essas empresas sejam acusadas de conferir mais atenção a uma área da moderação de conteúdo – como a dos direitos autorais, que pode sujeitá-las a prejuízos econômicos significativos<sup>469</sup> – em detrimento de outras, em que esse risco não é

---

<sup>469</sup> Cf. *supra*, nota n.º 97.

tão claro – mas que podem ser considerados socialmente mais importantes, como a repressão ao discurso de ódio. Revelar a fundamentação por trás das decisões tomadas significa abrir caminho para que elas sejam contestadas – por exemplo, viabilizando que pessoas que tiveram conteúdos excluídos ou permitidos comparem entre si os resultados dessas decisões – tal como o teste feito por Lipman – e, a partir deles, constatem incoerências na atuação das plataformas que de outra forma permaneceriam ocultas, pressionando-as a decidir em um ou outro sentido. Ainda, permitir que as pessoas recorram de uma decisão significa não só gerar uma possibilidade de que se constate um erro da plataforma – algo que fica ainda mais provável quando o responsável por analisar o recurso não faz parte de seus quadros –, como também aumenta o trabalho da moderação de conteúdo, algo sensível à vista do problema de escala que relatei antes<sup>470</sup>.

Portanto, o quadro que se faz presente ao menos até pouco tempo na moderação de conteúdo faz sentido dentro das abordagens que temos disponíveis: materialmente, as plataformas adotam uma perspectiva estatal-democrática, de defesa da liberdade de expressão, porque isso lhes interessa – na medida em que sua atividade depende, essencialmente, de que cada vez mais pessoas se manifestem nos ambientes que administram; processualmente, as plataformas adotam uma perspectiva que é tudo menos estatal-democrática, de procedimentos opacos, sem participação nem recorribilidade, porque isso também lhes interessa – na medida em que a ausência desses mecanismos gera menos restrições às suas decisões e lhes confere mais liberdade e poder para decidir aquilo que os seus usuários podem ou não podem dizer, de uma forma menos custosa. Conjugados esses interesses e perspectivas, eles geram um modelo curioso, mas justificável: é como se as plataformas tivessem implantado um sistema híbrido, que diz garantir os direitos da Constituição dos Estados Unidos tomando decisões como um Tribunal da Arábia Saudita.

O problema é que esse quadro passa por transformações que têm desafiado, com sucesso, as perspectivas que temos disponíveis para compreender a moderação de conteúdo. Em síntese: cada vez mais, as plataformas têm implantado mudanças radicais no procedimento que utilizam para apurar se um usuário violou ou não as suas regras, aproximando-o justamente do paradigma estatal do devido processo legal – e ninguém sabe explicar exatamente o porquê. É como se, depois de décadas de interesse na Primeira Emenda, as redes técnicas constituídas pelas plataformas com as melhores mentes da

---

<sup>470</sup> Cf. *supra*, Cap. I.2, 2.4.2.1.

tecnologia e do direito – que foram mobilizadas para construir um dos mais complexos mecanismos decisórios do mundo –, tivessem subitamente se dado conta que a Quinta e da Décima Quarta Emendas existem.

O YouTube, por exemplo, publicou em 2018 o primeiro relatório de transparência quanto à remoção de conteúdo da plataforma<sup>471</sup> – em que reunidas informações sobre o tipo e volume de materiais removidos – e, após quase uma década com as mesmas políticas de conteúdo, resolveu muda-las em 2019, por entender que “*o mais importante é ter uma aplicação consistente, políticas claras e transparências sobre o impacto de uma penalidade*”<sup>472</sup>. Com isso, foram incluídas nas Regras do YouTube informações quanto ao formato e consequências das sanções aplicadas pela moderação de conteúdo, que passará a operar de uma forma diferente. Primeiro: a plataforma está “*deixando nossas notificações de e-mail e desktop mais claras, e elas irão informar mais detalhes sobre qual política foi violada*”, consolidando-as em um local só, de forma que o usuário “*tenha todas as informações importantes sobre a sanção disponíveis com facilidade*”<sup>473</sup> – incluindo as razões que a fundamentam, tipo de sanção, data em que ela começa a ter seus efeitos e data em que eles findam. Segundo: facilitando uma política em que foi pioneira<sup>474</sup>, a plataforma está “*incluindo [a possibilidade] de apelar da decisão [que aplicou uma punição] se você entender que ela foi um erro*”<sup>475</sup>, de modo que cada

---

<sup>471</sup> GOOGLE. Aplicação das diretrizes da comunidade do YouTube. **Google Transparency Report**, 2018, disponível em <<https://transparencyreport.google.com/youtube-policy/removals>>, acesso em 30.7.2019.

<sup>472</sup> “[...] *consistent enforcement, clear policies, and transparency about the impact of a strike are most important*”. YOUTUBE. Making our strikes system clear and consistent. **YouTube Creator Blog**, 19.2.2019, disponível em <<https://youtube-creators.googleblog.com/2019/02/making-our-strikes-system-clear-and.html>>, acesso em 29.7.2019. A esse respeito, cf. GARTENBERG, C. YouTube updates channel strike system with one-time warning and more consistent punishments. **The Verge**, 19.2.2019, disponível em <<https://www.theverge.com/2019/2/19/18231382/youtube-strike-system-warning-punishment-community-guidelines>>, acesso em 29.7.2019.

<sup>473</sup> “[...] *we’re making our email and desktop notifications clearer, and they will provide more details on which policy was violated [...] have all the important information about a strike available at a glance*”. YOUTUBE. **Making our strikes system clear and consistent**. 19.2.2019, disponível em <<https://youtube-creators.googleblog.com/2019/02/making-our-strikes-system-clear-and.html>>, acesso em 29.7.2019.

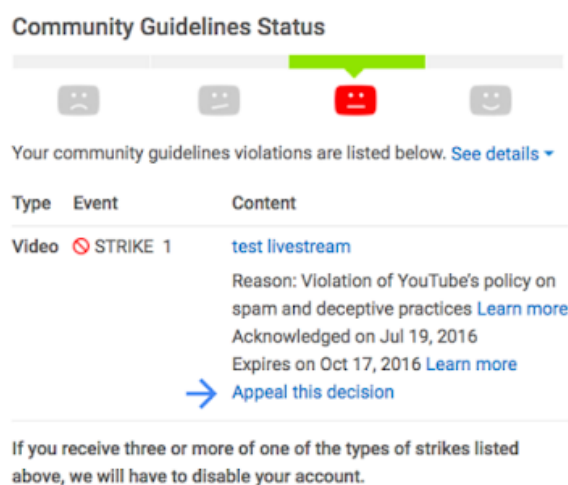
<sup>474</sup> WU, A. Strike you’re out! Or maybe not? **YouTube Official Blog**, 2.7.2010, disponível em <<https://youtube.googleblog.com/2010/07/strike-youre-out-or-maybe-not.html>>, acesso em 30.7.2019.

<sup>475</sup> “[...] *including appealing the decision in case you think it was a mistake*”. YOUTUBE. **Making our strikes system clear and consistent**. 19.2.2019, disponível em <<https://youtube-creators.googleblog.com/2019/02/making-our-strikes-system-clear-and.html>>, acesso em 29.7.2019.



usuário pode recorrer, pela própria plataforma, da decisão que ela tomou. Terceiro: essas medidas foram consolidadas em uma seção à parte das Regras do YouTube, que estabelecem, com clareza, o caminho procedimental a ser adotado para fazê-las valer. A imagem abaixo, por exemplo, mostra o novo painel do usuário, em que informações sobre a retirada de conteúdo são fornecidas; bem como explica o procedimento de apelação. Ela foi retirada da nova seção das regras da plataforma, que referi acima.

**Figura 19 – Transparência e recorribilidade no YouTube**

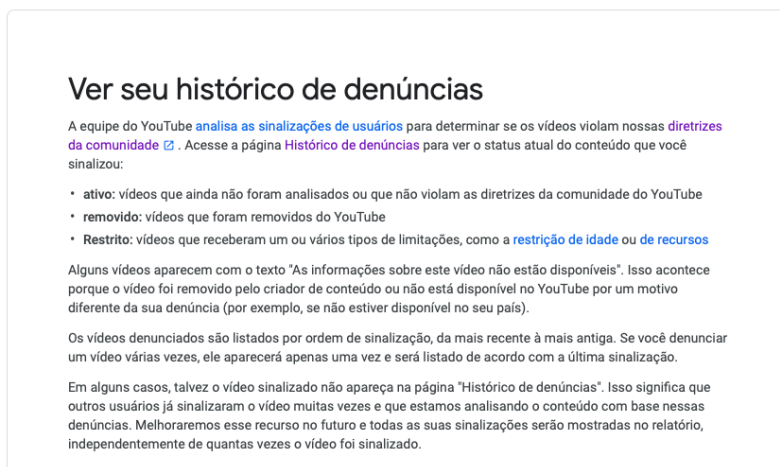


Fonte: YouTube<sup>476</sup>

Além disso, agora os usuários do YouTube que oferecem uma denúncia quanto ao conteúdo de uma publicação na plataforma podem acompanhar o seu resultado, por um painel próprio que expõe, para cada um dos casos, as providências tomadas pela moderação de conteúdo a respeito. É possível ver, nesse sentido, se o vídeo permaneceu ativo, foi removido ou se foram impostas restrições – como filtros por idade – à sua visualização:

<sup>476</sup> GOOGLE. Contestar ações das diretrizes da comunidade. **Ajuda do YouTube**, disponível em <[https://support.google.com/youtube/answer/185111?hl=pt-BR&ref\\_topic=2803138](https://support.google.com/youtube/answer/185111?hl=pt-BR&ref_topic=2803138)>, acesso em 29.7.2019.

**Figura 20 – Informações sobre o histórico de denúncias**



Fonte: YouTube<sup>477</sup>

O Twitter também passou a disponibilizar, em 2018, um relatório mais amplo de transparência, específico quanto à moderação de conteúdo<sup>478</sup>, incluindo o número de contas que foram denunciadas e, dentre elas, aquelas que sofreram algum tipo de ação da plataforma – de restrições de acesso ao conteúdo à exclusão. Sob a perspectiva da participação, e à semelhança do que fez o YouTube, o Twitter passou a dedicar um trecho de suas regras – ao qual é remetido o usuário toda vez que as viola<sup>479</sup> – para esclarecer quais são as sanções e os procedimentos que devem ser adotados pelos usuários ao serem notificados de que uma publicação foi considerada incompatível com as regras da comunidade:

*“Restrição à visibilidade do Tweet: o conteúdo fica menos visível no Twitter, nos resultados de busca, nas respostas e nas timelines. A restrição à visibilidade de um Tweet depende de várias indicações relacionadas à natureza da interação e à qualidade do conteúdo.*

<sup>477</sup> GOOGLE. Ver seu histórico de denúncias. **Ajuda do YouTube**, disponível em <[https://support.google.com/youtube/answer/7687979?hl=pt-BR&ref\\_topic=2803138](https://support.google.com/youtube/answer/7687979?hl=pt-BR&ref_topic=2803138)>, acesso em 29.7.2019.

<sup>478</sup> TWITTER. Twitter Rules enforcement. **Twitter Transparency Report**, 2018, disponível em <<https://transparency.twitter.com/en/twitter-rules-enforcement.html>>; cf., a esse respeito, o anúncio feito em TWITTER PUBLIC POLICY. Evolving our Twitter Transparency Report: expanded data and insights. **Twitter Blog**, 12.12.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2018/evolving-our-twitter-transparency-report.html](https://blog.twitter.com/en_us/topics/company/2018/evolving-our-twitter-transparency-report.html)>, acesso em 29.7.2019.

<sup>479</sup> GASCA, D., TOIZER, S. More clarity on reported Tweets and enforcement. **Twitter Blog**, 17.10.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/product/2018/more-clarity-on-reported-tweets-and-enforcement.html](https://blog.twitter.com/en_us/topics/product/2018/more-clarity-on-reported-tweets-and-enforcement.html)>, acesso em 29.7.2019.

*Solicitação de remoção do Tweet: quando concluímos que um Tweet violou as Regras do Twitter, exigimos que o violador o remova para que possa tweetar novamente. Enviamos ao violador uma notificação por e-mail identificando o Tweet e quais políticas foram violadas. O violador precisará, então, passar pelo processo de remoção do Tweet ou recorrer à análise do Twitter se acreditar que cometemos algum equívoco.*

*Ocultação de um Tweet ofensivo enquanto a remoção é aguardada: do momento em que o Twitter toma uma medida corretiva até quando a pessoa remove o Tweet, impedimos que esse Tweet seja visto publicamente e substituímos o conteúdo original por um aviso informando que o Tweet não está mais disponível porque violou nossas Regras. Além disso, esse anúncio ficará disponível por 14 dias após a remoção do Tweet.”<sup>480</sup>*

A plataforma, que tinha estabelecido mecanismos de apelação de suas decisões em 2018<sup>481</sup>, implantou uma nova forma de fazê-lo em 2019, de um modo mais rápido e transparente<sup>482</sup>. Como a imagem abaixo indica – disposta de forma sequencial, da esquerda para a direita, com os passos do procedimento –, o Twitter agora indica as razões pelas quais uma determinada publicação de um usuário foi suspensa; e viabiliza, no mesmo aviso, que seja feito um recurso. Há, ainda, um espaço para que razões sejam apresentadas à plataforma a respeito. O exemplo apresentado pelo Twitter é o de uma publicação supostamente agressiva, mas que se tratava, na realidade, de uma brincadeira sobre videogames entre amigos:

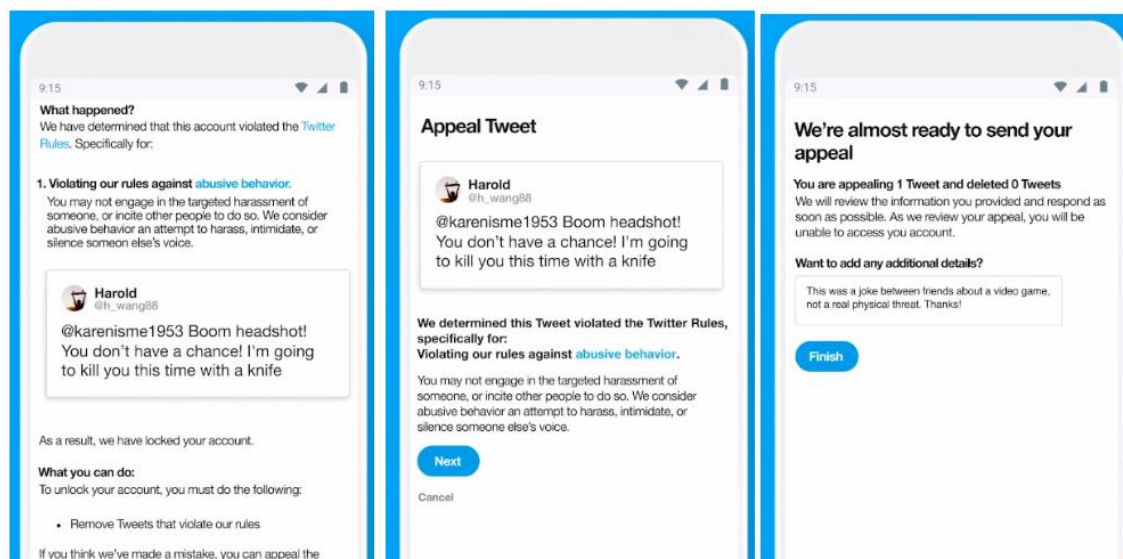
---

<sup>480</sup> TWITTER. **Diretrizes e políticas gerais – Nossas opções de medidas corretivas**. Disponível em <<https://help.twitter.com/pt/rules-and-policies/enforcement-options>>, acesso em 29.7.2019.

<sup>481</sup> ROTH, Y., HARVEY, D. How Twitter is fighting spam and malicious automation. **Twitter Blog**, 26.6.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html](https://blog.twitter.com/en_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html)>, acesso em 30.7.2019.

<sup>482</sup> A informação foi veiculada na conta Twitter Safety, no Twitter; cf. TWITTER SAFETY (Twitter). Sem título. Sem localização, 2 de abril de 2019. **Twitter**: TwitterSafety. Disponível em <<https://twitter.com/TwitterSafety/status/1113139073303089152>>, acesso em 29.7.2019.

Figura 21 – Procedimento de Apelação no Twitter



Fonte: Twitter<sup>483</sup>

Além disso, o Twitter passou a estabelecer de forma mais transparente o procedimento adotado pela plataforma a partir do momento em que um usuário reporta um conteúdo à moderação de conteúdo<sup>484</sup>.

A mudança de paradigma mais relevante, entretanto, ocorreu na maior plataforma da internet – o Facebook. Assim como o Twitter e o YouTube, a plataforma começou em 2018 um movimento de transformação do procedimento de moderação de conteúdo que mudou completamente o modo como decisões a respeito do que permanece na plataforma são tomadas. Nesse sentido, o ano começou com o anúncio<sup>485</sup> de que a

<sup>483</sup> As imagens foram obtidas utilizando-se de recurso que permite extrair uma cópia da tela do computador (“screenshot”), quanto à animação disponibilizada em TWITTER SAFETY (Twitter). Sem título. Sem localização, 2 de abril de 2019. **Twitter:** TwitterSafety. Disponível em <<https://twitter.com/TwitterSafety/status/1113139073303089152>>, acesso em 29.7.2019.

<sup>484</sup> “We’re refreshing rule pages to have more information, including examples, step-by-step instructions about how to report, and details on what happens when we take action” (“Nós estamos renovando as páginas em que estão as novas regras para que elas tenham mais informação, incluindo exemplos, instruções passo-a-passo sobre como denunciar, e detalhes sobre o que acontece quando nós agimos”). HARVEY, D. Making our rules easier to understand. **Twitter Blog**, 7.6.2019, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2019/rules-refresh.html](https://blog.twitter.com/en_us/topics/company/2019/rules-refresh.html)>, acesso em 29.7.2019.

<sup>485</sup> ZUCKERBERG, M. (Mark Zuckerberg). Sem título. Sem local, 24 de abril de 2018. **Facebook:** zuck. Disponível em <<https://www.facebook.com/zuck/posts/10104874769784071>>; e BICKERT, M. Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process. **Facebook Newsroom**,

plataforma tornaria público aos seus usuários os seus Padrões de Comunidade – que orientam as decisões tomadas pelos moderadores – e instituiria um procedimento de apelação quanto ao conteúdo que eles excluía, direcionando-o para a própria plataforma.

Em seguida, o Facebook avançou e estabeleceu uma sistemática de relatórios periódicos quanto à execução dos Padrões de Comunidade pela moderação de conteúdo, produzindo um primeiro documento nesse sentido já maio de 2018<sup>486</sup>, seguido do anúncio, em novembro, de um novo patamar de transparência quanto a esses dados, viabilizando pesquisas quantitativas a respeito; da ampliação do procedimento de apelação, para abarcar praticamente todas as decisões da moderação de conteúdo; e, por fim, da formação de um corpo decisório independente de sua estrutura empresarial – a “*Suprema Corte do Facebook*”<sup>487</sup> – que, com poder sobre as decisões tomadas na moderação de conteúdo da plataforma, tem o objetivo “*defender o princípio de conferir voz às pessoas, ao mesmo tempo em que reconhece a necessidade real de manter as pessoas seguras*”<sup>488</sup>. O projeto foi à frente e, em 28 de janeiro de 2019, a plataforma anunciou a formação, ao longo de 2019, do “*Comitê de Supervisão para Decisões de Conteúdo*”, um órgão independente para “*analisar as decisões de conteúdo do Facebook*”, “*Reverter decisões do Facebook quando necessário*” e “*ser uma autoridade independente externa ao Facebook*”<sup>489</sup> – uma espécie de corte arbitral internacional da liberdade de expressão. O

---

24.4.2018, disponível em <<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>>, acessos em 30.7.2019.

<sup>486</sup> FACEBOOK. Community Standards Enforcement Report. **Facebook Transparency**, 2018, disponível em <<https://transparency.facebook.com/community-standards-enforcement#hate-speech>>, acesso em 30.7.2019.

<sup>487</sup> A ideia – e o termo – foram utilizados por Zuckerberg (cf., por exemplo, KLEIN, E. Mark Zuckerberg on Facebook’s hardest year, and what comes next. **Vox**, 2.4.2018, disponível em <<https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge>>), e por especialistas (cf. KLONICK, K., KADRI, T. How to Make Facebook’s ‘Supreme Court’ Work. **The New York Times**, 17.11.2018, disponível em <<https://www.nytimes.com/2018/11/17/opinion/facebook-supreme-court-speech.html>>). Como consta no infográfico, o nome utilizado nos últimos anúncios é “*Oversight Board*”, algo como um Conselho de Supervisão, acessos em 30.7.2019.

<sup>488</sup> ZUCKERBERG, M. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

<sup>489</sup> CLEGG, N. Charting a Course for an Oversight Board for Content Decisions. **Facebook Newsroom**, 28.1.2019, disponível em <<https://newsroom.fb.com/news/2019/01/oversight-board/>>, acesso em 30.7.2019.

infográfico abaixo, retirado desse anúncio, fornece um quadro geral sobre esse novo órgão, expondo a função, composição e plano de trabalhos para implantação do Conselho:

**Figura 22 – Comitê de Supervisão para Decisões de Conteúdo**

## An Oversight Board to Review Facebook's Content Decisions


When reviewing content on Facebook, some decisions go to the core of how we balance safety and free expression. We believe these decisions are too consequential for Facebook to make alone. An oversight board will review these decisions.

---

**What will the board do?**

- Provide oversight of Facebook's content decisions
- Reverse Facebook's decisions when necessary
- Be an independent authority outside of Facebook


---



**Who will be on the board?**

A global group with expertise in diverse disciplines and a commitment to safety and free expression.

They will be obligated to the people who use Facebook—not Facebook the company.



**What issues will they discuss?**

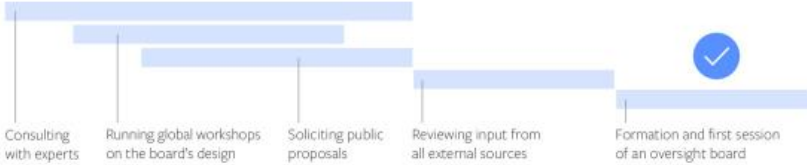
Our most difficult and contested decisions about taking down or leaving up content.

Some issues could be artistic expression, suspected hate speech or bullying.

---

**How will the board be formed?**

Facebook has begun consulting with experts, and will continue to provide updates through a multi-phased process in 2019.




---

**How will the board and Facebook inform the public moving forward?**

- Provide regular updates on the board's design
- Publish the charter and decision making process
- Announce all members
- Publish its decisions

---

**facebook**

© 2019 Facebook, Inc. All rights reserved. Materials, including images, are the property of Facebook and/or its licensors and may not be reproduced without Facebook's prior express permission. Trademarks are the property of their respective owners.

Fonte: Facebook<sup>490</sup>

Junto com a divulgação do formato do corpo decisório, o Facebook abriu uma consulta pública em que representados 88 países, via 28 workshops e mesas-redondas, colhendo informações com mais de 2.000 pessoas<sup>491</sup> e resultando em um documento<sup>492</sup> em que consolidadas as propostas a serem colocadas em prática nas próximas etapas. Elas se reuniram em três eixos principais: o estabelecimento de mecanismos de independência, que garantam que o Facebook não irá intervir nas decisões do colegiado; a construção de normas procedimentais – a respeito, por exemplo, da seleção de casos, da forma como o Conselho irá comunicar suas decisões e da possibilidade de se consultar terceiros com conhecimento técnico; e as regras de constituição desse Conselho, fazendo-o representativo da diversidade de usuários que compõem o Facebook. Conferindo corpo à proposta, em dezembro de 2019, a plataforma anunciou que constituiu um fundo independente (*trust*) de 130 milhões de dólares que financiará as atividades dos seis anos iniciais do painel – que devem se iniciar ainda em 2020 –, e atuará sem vinculação à empresa, atendendo diretamente aos custos do Conselho de Governança<sup>493</sup>.

Adiante neste trabalho – em específico, no Capítulo 5 – irei abordar esse caso com mais detalhes –, mas o que trouxe acima já é o suficiente para expor que o Facebook tem levado a sério a proposta, que caminha para, em breve, ser implantada na forma de um órgão independente que julgará aquilo que os usuários da plataforma – atualmente aproximadamente 2,27 bilhões de pessoas<sup>494</sup> – podem ou não dizer na internet.

Para o que interessa para este momento do trabalho, o fato é que, uma vez jungidas as constatações acima quanto às três plataformas que tenho utilizado como base para a análise da moderação de conteúdo, é possível visualizar que há, em curso, uma

---

<sup>490</sup> CLEGG, N. Charting a Course for an Oversight Board for Content Decisions. **Facebook Newsroom**, 28.1.2019, disponível em <<https://newsroom.fb.com/news/2019/01/oversight-board/>>, acesso em 30.7.2019.

<sup>491</sup> FACEBOOK. Global Feedback and Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, disponível em <<https://newsroom.fb.com/news/2019/06/global-feedback-on-oversight-board/>>, acesso em 29.7.2019.

<sup>492</sup> FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 29.7.2019.

<sup>493</sup> HARRIS, B. An Update on Building a Global Oversight Board. **Facebook Newsroom**, 12.12.2019, disponível em <<https://about.fb.com/news/2019/12/oversight-board-update/>>, acesso em 8.1.2020.

<sup>494</sup> Cf. *supra*, nota n.º 17.

profunda transformação nas decisões de caráter procedimental tomadas pelas plataformas da internet – isto é, no procedimento adotado para definir se uma publicação viola ou não as suas políticas internas.

Primeiro: a moderação de conteúdo *não é mais um procedimento tão opaco*. Como a divulgação de dados estatísticos específicos quanto à moderação de conteúdo – que dizem respeito, por exemplo, ao número de publicações excluídas ou de perfis suspensos e ao volume de denúncias que aportam periodicamente – de plataformas como o Facebook, Twitter e YouTube, cada vez mais há avanços no sentido de se poder fazer análises quantitativas e qualitativas<sup>495</sup> a partir dessas informações que, pressupõe-se, são fidedignas.

Segundo: a moderação de conteúdo *não é mais um procedimento marcado pela ausência de participação*. Há um claro movimento de mobilização do usuário, para que ele ocupe uma posição ativa nesse procedimento. O que era uma forma de decidir que causava perplexidade, caracterizada pela ausência de normativas sobre o seu procedimento e de mecanismos que permitissem acompanhar o seu trâmite e obter as razões que embasavam suas decisões tem, cada vez mais, avançado nesses aspectos, transformando-os. Como os exemplos acima demonstram, há, agora, seções à parte nas consolidações normativas das plataformas, indicando os passos do procedimento que julga as publicações; mecanismos que permitem acompanhar o trâmite e os resultados da atuação da moderação de conteúdo em relação a determinada conta ou publicação que foi reportada; e uma exposição mínima dos motivos que levaram um conteúdo a ser retirado ou uma conta de usuário a ser suspensa, indicando-se, ao menos, o item nas normativas da comunidade que foi considerado violado.

Terceiro: a moderação de conteúdo *não é mais um procedimento caracterizado pela irrecurribilidade*. Antes – e com a já mencionada exceção do YouTube –, não havia sequer uma aproximação do modelo estatal quanto à ideia de recorribilidade. Agora, as três plataformas que referi acima adotaram amplos mecanismos de recorribilidade, no interior dos quais inseriram a oportunidade de manifestação do usuário a respeito do conteúdo que foi publicado e considerado violador das regras das

---

<sup>495</sup> Cf., como um exemplo dessa possibilidade, a análise estatística feita antes desse movimento de abertura a partir de quatro meses de dados coletados quanto às políticas de nudez e de nomes reais do Facebook, em ANDERSON, J., STENDER, M., WEST, S., YORK, J. Unfriending Censorship: Insights from four months of crowdsourced data on social media censorship. **Online Censorship**, 31.3.2016, disponível em <<https://www.onlinecensorship.org/en/news-and-analysis/44>>, acesso em 30.7.2019.



plataformas. Esse movimento se desdobrou tanto na implantação desses mecanismos – caso do Facebook e do Twitter – quanto no aperfeiçoamento de vias estabelecidas anteriormente – como ocorreu no YouTube. Em todos esses casos, passou a haver uma espécie de “*apelação*”, em que, após ser notificado da decisão e dos seus fundamentos – também essa uma novidade nesse novo momento da moderação de conteúdo –, o usuário dispõe de um espaço próprio para arrazoar seu recurso, apresentando-o à plataforma – ou, em breve, no caso do Facebook, a um colegiado externo – e solicitando a reversão do que foi decidido antes.

O quadro abaixo sintetiza esse movimento que relatei acima a partir dos *Santa Clara Principles*, indicando a transformação que ocorreu nas plataformas no sentido de aproximar o procedimento de moderação de conteúdo – indicado em vermelho – a um quadro geral associado à cláusula do devido processo legal – indicado em verde, com a cor amarela indicando situações de parcial adequação. Ele esclarece que, com uma exceção – a sistemática de apelações do YouTube –, todas as demais modificações ocorreram no biênio 2018 / 2019, constituindo um movimento relativamente rápido e convergente em um mesmo sentido:

**Quadro 12 – A moderação de conteúdo em transformação a partir dos *Santa Clara Principles***

<u><i>Santa Clara Principles</i></u>	<u>Plataforma</u>	<u>Situação</u>	
<p><b>Princípio n.º 1</b></p> <p>“<i>Números</i>”</p> <p>A plataforma disponibiliza informações específicas quanto às publicações que são objeto da moderação de conteúdo?</p>	Facebook	<b>Antes de 15.5.2018:</b>	<u>Não</u> . Havia relatórios de transparência, mas eles exibiam informações parciais e que não eram particularizadas a respeito da moderação de conteúdo.
		<b>Depois de 15.5.2018:</b>	<u>Sim</u> . A partir da data, os relatórios de transparência passaram a contar com uma seção específica com essas informações, denominada “ <i>Community Standards Enforcement Report</i> ” <sup>496</sup> .
	Twitter	<b>Antes de 12.12.2018:</b>	<u>Não</u> . Havia relatórios de transparência desde 2012, mas eles exibiam informações parciais e que não eram particularizadas a respeito da moderação de conteúdo.
		<b>Depois de 12.12.2018:</b>	<u>Sim</u> . A partir do 13.º Relatório Bianual de Transparência, o Twitter passou a disponibilizar informações específicas sobre a atuação da plataforma sobre as publicações e contas de seus usuários <sup>497</sup> .
	YouTube	<b>Antes de 23.4.2018:</b>	<u>Não</u> . Havia relatórios de transparência, mas eles exibiam informações parciais e que não eram particularizadas a respeito da moderação de conteúdo.
		<b>Depois de 23.4.2018:</b>	<u>Sim</u> . A partir da data, o Google estabeleceu a publicação trimestral de relatórios que refletem como o YouTube aplica suas regras <sup>498</sup> .

<sup>496</sup> FACEBOOK. Community Standards Enforcement Report. **Facebook Transparency**, 2018, disponível em <<https://transparency.facebook.com/community-standards-enforcement#hate-speech>>, acesso em 30.7.2019.

<sup>497</sup> TWITTER. Twitter Rules enforcement. **Twitter Transparency Report**, 2018, disponível em <<https://transparency.twitter.com/en/twitter-rules-enforcement.html>>, acesso em 30.7.2019.

<sup>498</sup> GOOGLE. Aplicação das diretrizes da comunidade do YouTube. **Google Transparency Report**, 2018, disponível em <<https://transparencyreport.google.com/youtube-policy/removals>>, acesso em 30.7.2019.

<p><b>Princípio n.º 2</b></p> <p>“Participação”</p> <p>A plataforma oportuniza mecanismos de participação dos usuários na moderação de conteúdo, notificando-lhes quanto à detecção, fundamentação e decisão tomada?</p>	<p><b>Facebook</b></p>	<p><b>Antes de 24.4.2018:</b></p>	<p><u>Não</u>. Inexistiam informações específicas quanto à violação que permitissem ao usuário participar do processo de moderação de conteúdo.</p>
		<p><b>Depois de 24.4.2018:</b></p>	<p><u>Sim</u>. A plataforma implantou, junto com o sistema de apelações, um mecanismo que disponibiliza informações específicas quanto à política violada pelo usuário na publicação<sup>499</sup>.</p>
	<p><b>Twitter</b></p>	<p><b>Antes de 23.8.2018:</b></p>	<p><u>Não</u>. Inexistiam informações específicas quanto à violação que permitissem ao usuário participar do processo de moderação de conteúdo.</p>
		<p><b>Depois de 23.8.2018:</b></p>	<p><u>Sim</u>. A partir de 1.11.2017, o usuário que tem a sua conta suspensa recebe um e-mail, com detalhes sobre qual regra violou; a partir de 27.3.2018, o e-mail também é enviado para quem denunciou a publicação, informando o resultado da atuação da plataforma<sup>500</sup>.</p>
	<p><b>YouTube</b></p>	<p><b>Antes de 19.2.2019:</b></p>	<p><u>Em parte</u>. Havia a possibilidade de participação, mas as informações a respeito da detecção, fundamentação e decisão tomada eram limitadas.</p>
		<p><b>Depois de 19.2.2019:</b></p>	<p><u>Sim</u>. O YouTube implantou uma série de medidas, como notificações no e-mail e no aplicativo de que uma política foi violada pelo usuário, detalhando de forma mais precisa a regra violada<sup>501</sup>.</p>

<sup>499</sup> BICKERT, M. Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process. **Facebook Newsroom**, 24.4.2018, disponível em <<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>>.

<sup>500</sup> TWITTER SAFETY (Twitter). Sem título. Sem localização, 1 de novembro de 2017. **Twitter**: TwitterSafety. Disponível em <<https://twitter.com/TwitterSafety/status/925827487677870080>>, acesso em 30.7.2019; TWITTER SAFETY (Twitter). Sem título. Sem localização, 27 de março de 2018. **Twitter**: TwitterSafety. Disponível em <<https://twitter.com/TwitterSafety/status/978768701481656320>>, acesso em 30.7.2019.

<sup>501</sup> YOUTUBE. **Making our strikes system clear and consistent**. 19.2.2019, disponível em <<https://youtube-creators.googleblog.com/2019/02/making-our-strikes-system-clear-and.html>>, acesso em 29.7.2019.

<p><b>Princípio n.º 3</b></p> <p>“Apelação”</p> <p>Há um mecanismo que permita ao usuário recorrer da decisão da plataforma, apresentando suas razões preferencialmente a um órgão de revisão externo?</p>	<p><b>Facebook</b></p>	<p><b>Antes de 24.4.2018:</b></p>	<p><u>Não</u>. As decisões da plataforma eram definitivas.</p>
		<p><b>Depois de 24.4.2018:</b></p>	<p><u>Em parte</u>. A plataforma implantou um sistema de apelações, inicialmente direcionado para publicações relacionadas à nudez, atividade sexual, discurso de ódio e violência explícita<sup>502</sup>. Está em formação um corpo externo que irá avaliar parte desses recursos<sup>503</sup>.</p>
	<p><b>Twitter</b></p>	<p><b>Antes de 26.6.2018:</b></p>	<p><u>Não</u>. As decisões da plataforma eram definitivas.</p>
		<p><b>Depois de 26.6.2018:</b></p>	<p><u>Em parte</u>. A plataforma implantou um sistema de apelações, inicialmente por formulários<sup>504</sup> que depois foi ampliado para viabilizar recursos no próprio aplicativo<sup>505</sup>. É a própria plataforma que julga os recursos.</p>
	<p><b>YouTube</b></p>	<p><b>Antes de 2.7.2010:</b></p>	<p><u>Não</u>. As decisões da plataforma eram definitivas.</p>
		<p><b>Depois de 2.7.2010:</b></p>	<p><u>Em parte</u>. A plataforma implantou uma forma de apelar das decisões da moderação de conteúdo, que, com adaptações, é ativa até os dias de hoje<sup>506</sup>. É a própria plataforma que julga os recursos.</p>

<sup>502</sup> ZUCKERBERG, M. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

<sup>503</sup> CLEGG, N. Charting a Course for an Oversight Board for Content Decisions. **Facebook Newsroom**, 28.1.2019, disponível em <<https://newsroom.fb.com/news/2019/01/oversight-board/>>, acesso em 30.7.2019.

<sup>504</sup> ROTH, Y., HARVEY, D. How Twitter is fighting spam and malicious automation. **Twitter Blog**, 26.6.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html](https://blog.twitter.com/en_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html)>, acesso em 30.7.2019; trata-se da primeira referência oficial à existência do sistema de apelações, que é descrito, em um texto de 18.12.2017 (TWITTER SAFETY. Enforcing new rules to reduce hateful conduct and abusive behavior. **Twitter Blog**, 18.12.2017, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2017/safetypoliciesdec2017.html](https://blog.twitter.com/en_us/topics/company/2017/safetypoliciesdec2017.html)>, acesso em 29.7.2019) como em desenvolvimento.

<sup>505</sup> TWITTER SAFETY (Twitter). Sem título. Sem localização, 2 de abril de 2019. **Twitter**: TwitterSafety. Disponível em <<https://twitter.com/TwitterSafety/status/1113139073303089152>>, acesso em 29.7.2019

<sup>506</sup> WU, A. Strike you're out! Or maybe not? **YouTube Official Blog**, 2.7.2010, disponível em <<https://youtube.googleblog.com/2010/07/strike-youre-out-or-maybe-not.html>>, acesso em 30.7.2019.

É importante ressaltar que, como ocorreu quanto aos aspectos materiais – em que absorvido um “*quadro geral*”, inspirado na Primeira Emenda –, também quanto aos aspectos procedimentais não se caminha para a replicação integral de um modelo decisório estatal nas plataformas. Assim como há manifestações que são resguardadas pela Primeira Emenda, mas vedadas pelas plataformas, há regras procedimentais que decorrem da cláusula do devido processo legal nos Estados – como a de ser representado por um advogado<sup>507</sup> – mas que não têm perspectiva de serem aplicadas nas plataformas. Entretanto, da mesma forma que essa absorção limitada a um quadro geral não obscurece a inspiração que decorre da Primeira Emenda quanto às decisões materiais, a ausência de uma coincidência integral entre o que fazem os Estados e o que ocorre na moderação de conteúdo não prejudica a constatação de que as plataformas passam por uma transformação. Ela se manifesta, em resumo, em uma transição de um modelo procedimental que era tudo menos baseado nas práticas do Estado, para um que absorve o essencial exatamente desse formato decisório específico e contingente, adotando uma forma de tomar decisões que, sustentada em ideias como transparência, participação e recorribilidade, amolda-se – assim como a Primeira Emenda se referia à liberdade de expressão – a um quadro geral refletido na cláusula de respeito ao devido processo legal.

O problema é que se havia o que justificasse o quadro anterior, em que uma replicação material da ideia de liberdade de expressão era associada a um distanciamento processual do devido processo legal, não há o que explique o movimento atual das plataformas. Faz sentido utilizar a Primeira Emenda como uma base para moderar conteúdo: ela tanto viabiliza uma maior permanência e interação entre os usuários da plataforma, quando permite que a plataforma exerça, excepcionalmente, poder sobre as manifestações que impactam negativamente nesses efeitos, garantindo que esses fatores – que, no final das contas, são os que geram lucro para essas empresas – sejam maximizados. Faz sentido, também, distanciar-se do Estado como um paradigma de procedimento decisório: moderar conteúdo de uma forma opaca e sem oportunizar manifestação aos usuários garante mais poder e discricionariedade à plataforma em suas decisões, permitindo modulá-las – e mesmo alterá-las – de acordo com o que melhor lhe convém. Mas por que as plataformas decidiram mudar o seu modelo decisório, guinando-o a um formato sustentado no que obriga os Estados? Que razões existem para que, depois de

---

<sup>507</sup> FRIENDLY, H. Some Kind of Hearing. *University of Pennsylvania Law Review*, v. 123, n. 6, 1975, pp. 1.287/1.291 (“7. Counsel”).

desenvolverem um sistema próprio de moderação de conteúdo – construído para que tivessem o total controle sobre os seus resultados –, elas agora se empenham em transformá-lo em algo mais próximo do modelo decisório estatal, perdendo, com isso o domínio que têm sobre o ambiente que administram? Por que, afinal de contas, essas empresas estão abrindo mão do seu poder?

Uma busca por essa resposta se depara, de início, com a constatação de que esse era um movimento há muito exigido pelos atores desse debate. Como os casos que reconstruí ao longo deste trabalho expõem, sempre foi objeto de certo incômodo das pessoas o fato de que as decisões sobre aquilo que podem ou não dizer nas plataformas fossem tomadas por um procedimento arbitrário. A forma que melhor resume essa questão, e que, inclusive, utilizei como base para detectar esse movimento, são os *Santa Clara Principles*, que consolidaram, à época, a percepção de organizações, acadêmicos e advogados no sentido de que era preciso mudar a forma como as plataformas moderam conteúdo, direcionando-o para o devido processo legal.

É superficial, entretanto, extrair desse movimento a razão por trás da transformação que relatei acima. No contexto das plataformas, dizer que elas mudaram simplesmente porque foram expostas a pressões de organizações não governamentais é um argumento, no mínimo, incompleto. Isso porque adotar essa explicação como a única necessária para explicar o movimento significa dizer que, para esclarecê-lo, basta adotar uma narrativa otimista – em parte induzida pelas próprias plataformas<sup>508</sup> –, de que, atentas à sua responsabilidade social e movidas por pressões de grupos organizados, empresas como o Facebook teriam decidido modificar a moderação de conteúdo, adequando-a a um modelo que melhor atendesse àquilo que esperam seus usuários. Essa transformação seria, portanto, simultaneamente um testemunho do sucesso desses movimentos sociais e algo que confere esteio à propaganda dessas companhias, de que não são meras estruturas

---

<sup>508</sup> “*I’ve always cared about this and that’s why the first words of our mission have always been ‘give people the power’. I believe the world is better when more people have a voice to share their experiences, and when traditional gatekeepers like governments and media companies don’t control what ideas can be expressed*” (“*Eu sempre me importei sobre isso e é por isso que as primeiras palavras de nossa missão sempre foi ‘dê às pessoas o poder’. Eu acredito que o mundo é melhor quando mais pessoas têm voz para compartilhar suas experiências e quando os gatekeepers tradicionais como governos e companhias de mídia não controlam quais ideias podem ser expressadas*”). ZUCKERBERG, M. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

empresariais, mas comunidades que, como indica o slogan do Facebook, servem de “*utilidade social que conecta você às pessoas ao seu redor*”. Da mudança, portanto, se extrairia a evidência do quanto essas companhias dependem e se moldam pelo que exigem os seus usuários, mesmo que isso signifique abrir mão do poder que lhes permite controlar o que é nelas veiculado.

A questão é que esse paralelismo entre demanda dos usuários e mudança nas plataformas pode ser enganoso, ao sugerir uma causalidade única onde pode haver só coexistência, correlação ou, no mínimo, uma causalidade mais complexa entre fatores desconhecidos – que permanecem ocultos quando, aceitando aquilo que dizem essas companhias, entendemos essas modificações como um simples atendimento àquilo que os membros dessas comunidades exigiram.

Nesse sentido, e por mais atrativa que seja essa narrativa, há indicativos de que, por trás dela, há algo mais complexo que deixamos de compreender quando a aceitamos sem questionar. Por um lado, como as seleções estratégicas de aspectos materiais e procedimentais que relatei acima expõem, a formatação da moderação de conteúdo não é um processo inocente, mas um modelo decisório constituído, em cada um de seus detalhes, pela cúpula dessas plataformas, traduzindo um interesse interno de manter o mais pleno controle sobre esse aspecto de sua atividade empresarial. Isso faz sentido: plataformas podem ser responsabilizadas tanto por aquilo que seus usuários publicam – nem todo país tem, como os Estados Unidos, uma política de imunidade em relação àquilo que é veiculado por terceiros nos sites dessas companhias<sup>509</sup> – quanto por

---

<sup>509</sup> Cf., a esse respeito, o papel central que a Seção 230 da Lei de Telecomunicações dos Estados Unidos têm na moderação de conteúdo ao inviabilizar a responsabilização de plataformas como o Facebook por aquilo que é disponibilizado nela por terceiros – uma regra, entretanto, que não está necessariamente presente em todas as jurisdições que a plataforma atua, algo que pode gerar estímulos no sentido de adotar políticas mais restritivas de moderação de conteúdo; nesse sentido, KLONICK, K. *The New Governors: The People, Rules, And Processes Governing Online Speech*. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1604 e ss; GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, pp. 29 e ss (Kindle, n. p.); e KOSEFF, J. **The Twenty-Six Words that Created the Internet**. Ithaca: Cornell University Press, 2019, pp. 145 e ss. No caso do Brasil, o dispositivo que pode ser considerado correspondente é o art. 19, *caput*, da Lei n.º 12.965 (Marco Civil da Internet): “*Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica, não tomar as providências para, no*

aquilo que fazem ou deixam de fazer em relação a esse conteúdo<sup>510</sup>. É por isso que é preciso ter cautela com uma leitura segundo a qual essas companhias, depois de mobilizar um corpo especializado de advogados, estruturar um formato próprio de análise de conteúdo e operá-lo por anos segundo uma – justificada – lógica de opacidade, ausência de participação e irrecorribilidade, tenham agora percebido que a forma mais adequada de se chegar a uma decisão é a que já existia muito antes de conceberem seus modelos de moderação de conteúdo.

Nesse sentido, se o descompasso entre uma base material associada à Primeira Emenda e uma base processual fundamentalmente distinta do devido processo legal não foi uma obra do acaso – mas o produto de uma forma estratégica de pensar essas plataformas – é razoável supor que a recente mudança nesse paradigma também tenha sido muito mais que um gesto de simpatia a usuários que exigiam isso, sobretudo à vista da forma como as plataformas lidam seletivamente com as pressões que constantemente sofrem de grupos sociais<sup>511</sup>. Tanto essa dualidade entre exigência dos usuários e atendimento das plataformas, quanto a narrativa mutuamente benéfica que decorre dela merecem, no mínimo, desconfiança: parece haver, por trás desse movimento todo, alguma espécie de cálculo estratégico vendido como atenção ao usuário que passa despercebido por esse modo de interpretar a mudança.

Ainda, mesmo que a interpretação de que essa transformação seja o resultado de uma dinâmica de exigência do usuário e atendimento da plataforma esgote o modo como os atores desse processo a entendam, isso não significa que essa seja a visão mais completa do que ocorre, atualmente, com o modelo regulatório privado da liberdade

---

*âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente, ressalvadas as disposições legais em contrário”.*

<sup>510</sup> Cf. *supra*, na nota n.º 350, as obrigações às quais vinculadas as plataformas para denunciar a pornografia infantil.

<sup>511</sup> Cf., por exemplo, a controvérsia a respeito da transparência com que o Facebook lida com os anúncios em sua plataforma, em específico os associados a campanhas políticas. Nesse sentido, mesmo após alterar suas políticas a respeito, a plataforma ainda resiste em liberar dados estatísticos que seriam fundamentais para pesquisas de ciência política a respeito, como expõem INGRAM, D. Facebook to keep wraps on political ads data despite researchers' demands. **Reuters**, 22.6.2017, disponível em <<https://www.reuters.com/article/us-usa-politics-facebook/facebook-to-keep-wraps-on-political-ads-data-despite-researchers-demands-idUSKBN19D1CN>>; e FACEBOOK. Q & A on Transparency for Electoral and Issue Ads. **Facebook Newsroom**, 24.5.2018, disponível em <<https://newsroom.fb.com/news/2018/05/q-and-a-on-ads-transparency/>>, acessos em 30.7.2019.



de expressão na internet. Como as controvérsias que têm surgido no modelo atual – a exemplo da exclusão acidental de vídeos educativos sobre o Holocausto como se fossem propaganda nazista<sup>512</sup> – demonstram, o domínio que as plataformas exercem sobre a moderação de conteúdo é limitado, e não significa que elas ou mesmo os seus usuários sejam oniscientes quanto a como essa forma de decidir funciona e se transforma. Assim como entrevistar um parlamentar quanto às suas intenções relativamente a uma legislação que auxiliou a aprovar não é um dado suficiente para entender as suas causas e consequências, colher aquilo que os executivos do Facebook esperam atingir com as mudanças que implantaram na moderação de conteúdo é algo insuficiente para compreendê-las. Por mais que esse tenha sido um método importante para definir como essas plataformas funcionam, utilizado neste trabalho por meio da literatura reconstruída no capítulo anterior, ele tem limitações, porque nada garante que os responsáveis por operacionalizar esse tipo de transformação tenham completa consciência daquilo que ocorre nessas plataformas. Por isso, mais do que colher novas entrevistas, essa parece ser o momento de se tomar distância da mudança diagnosticada acima, analisando-a não como um conjunto de intencionalidades provenientes de plataformas e usuários, mas como um movimento contingente de um modelo decisório.

Colocada de lado a solução de entender a transformação acima como o resultado de pressões dos usuários das plataformas, o caminho seguinte para se ter uma análise mais distanciada do problema – no sentido de que menos vinculada às percepções dos atores que atuaram nesse movimento – seria mobilizar as formas de pensar a moderação de conteúdo que temos disponíveis e aplicá-las à situação, tentando explicá-la. Essa, entretanto, é uma abordagem malsucedida: tanto a versão diádica controvertida por Klonick, quanto a pluralista defendida por ela mostram-se insuficientes para fornecer uma explicação razoável do que está acontecendo.

Isso porque, de um lado, quanto às perspectivas diádicas – que enxergam a moderação de conteúdo por meio de categorias jurisprudenciais vinculadas à Primeira Emenda –, utilizá-las para enfrentar esse tipo de problema só reforça as conclusões, postas no primeiro e segundo capítulo, de que elas colaboram pouco para compreender essa forma de regular a liberdade de expressão. Primeiro, porque as analogias construídas por essa perspectiva não se amoldam às especificidades da moderação de conteúdo, gerando comparações que deixam de abranger aspectos essenciais à sua atividade – em específico,

---

<sup>512</sup> Cf. *supra*, Cap. I.2, 2.4.2.1.

a sua posição intermediária que nem produz o próprio conteúdo, nem apenas conduz o conteúdo de terceiros. Segundo, porque mesmo que se produzisse uma analogia perfeita – ou mesmo que houvesse um precedente aplicável a essas plataformas – isso explicaria muito pouco da transformação detectada acima. Perspectivas diádicas entendem as plataformas sob o quadro dos Estados, ou seja, pela forma como elas se encaixam ou não em determinados regimes jurídicos que lhe são alheios. Por isso, elas não são percebidas como nichos decisórios autônomos, capazes, por exemplo, de mudar a forma como tomam decisões – cerne do problema colocado. No mais, a ausência de regulamentações ou decisões conhecidas que tenham imposto às plataformas um modelo decisório determinado<sup>513</sup> é um indicativo relevante de que essa é uma transformação que não adveio da conformação dessas empresas a um regime jurídico estatal que precisa ser desvendado, mas sim de um movimento interno e relativamente independente desses agentes.

De outro, se as perspectivas pluralistas fornecem um caminho mais frutífero que as diádicas, elas não conseguem ir longe o suficiente para explicar esse último movimento das plataformas. Leituras como a de Klonick se habilitam a melhor compreender o problema, porque percebem as plataformas como instâncias independentes e capazes de construir normatizações próprias – e, portanto, modificáveis – de acordo com o seu interesse. É possível vê-las, nesse sentido, como instâncias que “*governam*” a liberdade de expressão de seus usuários e, por isso, podem alterar as regras do jogo, como o fizeram nesse movimento, transformando o modo como decisões sobre o conteúdo publicado por seus usuários são produzidas.

Dizer, entretanto, que as plataformas *podem* fazer isso não significa explicar o *porquê* delas o terem feito. A pergunta que fiz acima permanece: se interessa mais às plataformas manter um modelo decisório que lhes garante o poder sobre aquilo que é

---

<sup>513</sup> Diferente do que ocorre com a forma pela qual as plataformas tratam a privacidade das pessoas, algo que é normatizado tanto em âmbito nacionais – a exemplo da Lei n.º 13.709/2018, a Lei Geral de Proteção de Dados Pessoais no Brasil – quanto supranacional – como no Regulamento Geral sobre a Proteção de Dados, que vige na União Europeia e no Espaço Econômico Europeu –, há poucas normas estatais específicas que incidam sobre a moderação de conteúdo e que possam ter alguma relação com esse movimento. O exemplo mais importante – e que, mesmo assim, é restrito geograficamente e, por isso, incapaz de explicar o movimento detectado neste capítulo, que vale de modo indistinto para a atuação das plataformas em todos os países – é o da *Netzwerkdurchsetzungsgesetz* vom 1. September 2017 (BGBl. I S. 3352), a “*Lei do Facebook*”, que estabeleceu, por exemplo, disposições quanto ao prazo para a retirada de conteúdo da plataforma (§ 3.º, (2)) e ao procedimento a ser adotado pela plataforma quando receber reclamações a respeito do conteúdo disponibilizado (§ 2.º, (1)). *Cf. supra*, notas n.º 166 e 167.

veiculado na internet, por que mudar? Isso, o pluralismo parece ainda não conseguir responder: ele toma um caminho mais viável, ao possibilitar hipóteses que tomam as plataformas como espaços de decisão independente – e, por isso, que não dependam de uma determinação estatal para explicar uma transformação –, mas não vai longe o suficiente para concretizar uma teoria capaz de ir aos pormenores desse movimento. Klonick aventa, nesse sentido, justamente a possibilidade dessa transformação – isto é, que as “*plataformas voluntariamente assumam um compromisso com a noção de ‘devido processo legal tecnológico’*” –, descartando-a, entretanto, porque reconhece que, no quadro teórico que se tinha à mão, “*essas plataformas privadas têm pouca motivação para abrir mão do poder como em uma empresa voltada ao bem público, ou para adotar as regras e as ideias de transparência requeridas no devido processo legal tecnológico de Citron*”<sup>514</sup>. O problema é que, contrariamente a essa percepção – e à capacidade explicativa da perspectiva pluralista que está por trás dela –, foi justamente esse o caminho tomado pelas plataformas.

A proposta da segunda parte deste trabalho, nesse sentido, é compreender as razões por trás desse movimento, que fez as plataformas migrarem de modelos decisórios quase arbitrários para regramentos que, ao estabelecerem mecanismos de transparência, participação e recorribilidade, propõem uma forma de decidir cada vez mais próxima à do Estado. O objetivo, com isso, é compreender melhor os elementos que temos disponíveis a esse respeito e que, atualmente, constroem um quadro que reflete apenas parte do que acontece: parece haver uma espécie de *tradeoff* oculto, segundo o qual quanto mais poder cedem as plataformas, adotando procedimentos menos arbitrários, mais elas obtêm algo que permanece desconhecido para nós, mas é tão importante para elas que as faz ceder em um ponto estratégico da forma como administram a sua atividade. Minha intenção, em síntese, é propor um modelo que dê sentido a essa decisão contingente e, dessa forma, explique melhor como se transforma o mecanismo que determina, todos os dias, aquilo que as pessoas podem dizer na internet.

---

<sup>514</sup> “*Another avenue would be for platforms to voluntarily take up a commitment to a notion of ‘technological due process.’ In this groundbreaking model for best practices in agency use of technology, Citron advocates for a model that understands the trade-offs of ‘automation and human discretion,’ protects individuals’ rights to notice and hearings and gives transparency to rulemaking and adjudication*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.668/1.669.

Minha hipótese, a esse respeito, é de que a transformação pela qual passa a moderação de conteúdo tem dois significados. Primeiro, o movimento das plataformas de adotar procedimentos decisórios mais próximos ao paradigma do devido processo legal não é apenas uma forma de atender às exigências de seus usuários, mas a solução para um problema que essas companhias enfrentam ao tomar esse tipo de decisão: a falta de legitimidade<sup>515</sup>. Meu argumento, nesse sentido, é de que diferente do modelo anterior, em que o foco das políticas de moderação de conteúdo era manter o absoluto controle de empresas como o Facebook quanto àquilo que percorria suas plataformas, o modelo atual se propõe a abrir mão desse poder em nome de um procedimento que leva as pessoas a aceitarem as decisões tomadas pelas plataformas, porque legítimas, mesmo que lhe sejam desfavoráveis. Dessa forma, e, assim como ocorre nos processos judiciais, estabelecer um procedimento decisório baseado em ideias como transparência, participação e contraditório em plataformas na internet tem uma função chave: conferir, à decisão que é produto desse procedimento, um corpo de legitimidade, que se traduz em uma aceitabilidade do desfavorável, neutralizando o potencial destrutivo do conflito gerado a partir da divergência da qual ela emerge.

Segundo, sustento que essa transformação significa algo a mais: um indício importante de um movimento maior, designado por Gunther Teubner como a constitucionalização do mundo digital. Nesse sentido, e desde que feita a análise da moderação de conteúdo a partir de uma perspectiva do pluralismo jurídico que confere um significado não-estatal tanto à ideia de direito, quanto à de constituição, argumento ser possível ler essa mudança como parte de uma transformação mais ampla que, viabilizada pela internet, permite perceber não só a formação de novos polos decisórios não-estatais – como a moderação de conteúdo –, mas também a constitucionalização dessas forma de decidir – pela estruturação de formatos procedimentais responsáveis por imunizar essas decisões – tanto das estruturas corporativas em que estão inseridos, quanto daquilo que lhe é externo, como a influência estatal.

### 3. MÉTODO

#### *3.1 Por que sociologia?*

---

<sup>515</sup> AMMORI, M. The ‘New’ New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, 2014, pp. 2.276/2.277.

Na seção anterior, delimito o problema da segunda parte deste trabalho: meu objetivo, em síntese, é compreender as transformações pelas quais passa a moderação de conteúdo das plataformas da internet para descobrir quais são as razões e consequências por trás do movimento de companhias como o Facebook, o Twitter e o Google, que, mesmo após desenvolverem métodos decisórios independentes quanto àquilo que as pessoas podem dizer em suas plataformas, decidiram adotar modelos próximos ao do estatal para determinar o que as pessoas podem ou não dizer em suas plataformas. Também expus que, à primeira vista, os modelos que temos disponíveis para compreender a moderação de conteúdo não fornecem respostas viáveis a esse problema: por um lado, o viés diádico enfoca a questão sob a perspectiva estatal – que não é útil, porque, ao que tudo indica, esse movimento surge e se desenvolve nas próprias plataformas e não como o resultado da atuação de um ator externo, como o Estado; por outro, a perspectiva pluralista que temos disponível não é madura o suficiente para lançar luz ao *tradeoff* que parece oculto nessa situação. Como, então, responder essa pergunta? Que método adotar para resolver esse problema?

Na segunda parte deste trabalho, sugiro inserir a transformação pela qual passa a moderação de conteúdo em um quadro teórico mais amplo. Em específico, minha proposta é utilizar a sociologia, e, em particular, a Teoria dos Sistemas de Niklas Luhmann, para abordar esse problema. Fazer isso exige duas justificativas metodológicas: primeiro, é preciso oferecer razões que sustentem deixar de lado, por um momento, abordagens enfocadas na obtenção de mais informações quanto à moderação de conteúdo – entrevistando seus principais atores, por exemplo –, conferindo prioridade a se estruturar, a partir de uma teoria robusta, um quadro teórico que permita organizar as que já temos sob uma perspectiva sociológica. Segundo, é preciso oferecer razões que justifiquem escolher, dentro desse universo sociológico, a Teoria dos Sistemas. Abordo, nesta seção, a primeira justificativa e dedico, à segunda, a seção seguinte.

Colocadas as abordagens diádicas de lado – pelas razões que expus no primeiro e segundo capítulos, e também na seção anterior –, a visão pluralista da moderação de conteúdo tem se concentrado muito mais em obter informações a respeito de como esse mecanismo decisório funciona do que organizá-las. Isso é compreensível, porque, como insisti na primeira parte, diferente de outras formas de regulação da liberdade de expressão – como as utilizados pelo Estado –, que estão submetidas a protocolos de transparência, a moderação de conteúdo é uma forma de decidir

essencialmente opaca. O texto de Klonick, que utilizei acima<sup>516</sup>, é um exemplo disso: a parte mais extensa do artigo é dedicada a explicar, a partir de entrevistas com executivos de companhias de tecnologia, como se constituiu e como funciona, atualmente, a moderação de conteúdo nas plataformas da internet. Por isso, mesmo que o cerne do texto seja apresenta-las como as “*novas governantes*” da liberdade de expressão, essa é uma ideia que acaba relegada ao final do artigo. Ver a moderação de conteúdo como uma forma de governança privada apreende com mais precisão – descritiva e normativa – o que acontece nas plataformas, mas falta desenvolver essa explicação melhor, desdobrando-a em uma estrutura que consiga ser capaz de compreender os vetores de transformações como a que descrevi acima. Como disse na seção anterior, é sintomático dessa constatação que Klonick chegue a cogitar a adoção de um modelo de devido processo legal nas plataformas, mas logo descarte essa possibilidade por não ver razão para as plataformas o fazerem<sup>517</sup>.

Por isso, irei me concentrar, na segunda parte deste trabalho, em robustecer os instrumentos que temos disponíveis para abordar a moderação de conteúdo sob uma perspectiva pluralista – focando, portanto, mais na abordagem que consta no final do artigo de Klonick do que a de seu início –, desenvolvendo-os e aplicando-os nas informações que temos disponíveis a respeito da moderação de conteúdo, e que consolidei nos dois primeiros capítulos da dissertação.

Concretamente, minha proposta, portanto, não é dar continuidade àquilo que predominou nas abordagens anteriores – conduzindo, por exemplo, novas entrevistas com executivos das plataformas –, mas, antes, entender como as informações colhidas esse respeito podem ser melhor organizadas por uma estrutura teórica mais robusta. Faço isso tanto porque essa é uma abordagem ainda pouco explorada quanto à moderação de conteúdo – o texto mais atual e completo a esse respeito, de Klonick, não avança o suficiente para resolver o problema proposto acima – quanto porque a forma como a moderação de conteúdo tem sido aprofundada – sem essa estrutura teórica robusta por trás – tem limitações.

Essas limitações se manifestam, por exemplo, na dificuldade em se identificar, em eventuais respostas que resultariam de uma abordagem direta aos

---

<sup>516</sup> Refiro-me a KLOINICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, pp. 1.598/1.670.

<sup>517</sup> Cf. *supra*, nota n.º 514.

executivos dessas plataformas, o que reflete a realidade por trás do movimento que pretendo investigar e o que reflete um discurso que é interessante para esse tipo de empresa e para os seus agentes divulgar. Dito de outra forma: pode não convir, para uma plataforma como o Facebook, expor a estratégia por trás da adoção dessa forma diferente de moderar conteúdo. Não é difícil cogitar o porquê: como apresentei na seção anterior, já existe um discurso oficial. Dizer que essas plataformas se mostraram sensíveis às exigências de seus usuários e, por isso, resolveram adotar um paradigma que lhes confere voz em um procedimento antes arbitrário é uma posição confortável para todos os atores dessa transformação. Por um lado, as organizações responsáveis por essa pressão apresentam esse resultado como um sucesso; por outro, as plataformas expõem, por meio dele, a sua sensibilidade àquilo que os usuários lhes apresentam como demanda, vendendo-se como comunidades e não empresas. Nesse sentido, e, se essa é a posição oficial dessas empresas<sup>518</sup>, não há razão para pensar que ela seria mitigada – mas, antes, reforçada – por uma abordagem que procura colhê-la dos atores envolvidos nesse processo.

Ainda, esse tipo de abordagem pressupõe que os atores responsáveis por esse movimento tenham plena consciência do que fazem, guindando a plataforma em um ou outro sentido a partir de um plano escondido em alguma gaveta que poderia ser revelado com alguma insistência. Isso, entretanto, pode não ser verdade: mesmo que se pudesse contar com a disposição desses atores, é superficial compreender um movimento dessa magnitude como uma simples reunião de intencionalidades que, localizadas nas mentes dos principais executivos dessas empresas, poderiam ser acessadas com um grau suficiente de persuasão. Nesse sentido, minha pretensão ao tentar resolver o problema que posicionei na seção anterior não é entender o que se passa na cabeça dos executivos do Facebook, mas sim compreender a lógica por trás do caminho que tem sido percorrido por um modelo decisório. Ele pode, nesse sentido, refletir de algum modo um plano estabelecido por essas pessoas, mas considera-las oniscientes ao ponto em que a pesquisa se reduza a acessá-las e tomar aquilo que dizem como verdade é uma abordagem no mínimo, incompleta.

A proposta desta parte do trabalho, nesse sentido, é enfrentar essas dificuldades a partir de uma abordagem teórica, que mais do que tentar colher novos elementos a respeito da transformação pela qual passa a moderação de conteúdo, insere o que já temos disponível – e que reconstruí nos capítulos anteriores – em um quadro teórico

---

<sup>518</sup> Cf. *supra*, nota n.º 450.

mais amplo e capaz de articular aquilo que acontece com a moderação de conteúdo, aventando e sustentando hipóteses a respeito de suas transformações. Minha proposta, nesse sentido, é adotar uma perspectiva teórica já estruturada que mitigue as limitações do foco identificadas até o momento quanto à moderação de conteúdo, proporcionando respostas que dependam menos, por um lado, da disposição de fontes determinadas – como executivos dessas empresas – de refletir aquilo que realmente ocorre nessas plataformas; e, por outro, da onisciência dessas pessoas sobre os processos em que estão envolvidas.

Para realizar isso, proponho fazer uma abordagem sociológica do problema que apresentei na seção anterior. Por abordagem sociológica quero expressar, em sentido amplo, a circunstância de que, diferente daquilo que se aproxima de um viés diádico, minha intenção não é vislumbrar essas questões sob o viés interno, de um sistema jurídico, descobrindo quais normas são aplicáveis a uma determinada situação – no caso, a moderação de conteúdo –, mas, antes, vê-la por uma dinâmica externa, como um núcleo produtor de normatizações, algo que se alinha com aquilo que argumentei no primeiro capítulo, no sentido de que perspectivas que tentam subsumir a moderação de conteúdo a um regime jurídico específico – via, por exemplo, analogias –, não conseguem compreender, em toda a sua extensão, os impactos da era digital para a regulação da liberdade de expressão.

Nesse sentido, a sociologia fornece tanto uma perspectiva externa desse fenômeno – e, assim, alinhada com o pluralismo de Klonick e do Balkin de 2018 – quanto instrumentais teóricos mais amplos e contextualizados que a governança privada, esboçada por Klonick em seu texto, para analisar esse tipo de problema. Por “*mais amplos*”, refiro-me à circunstância de que a sociologia trabalha em um nível acima ao da ideia de governança privada, na medida em que opera com conceitos mais fundamentais – como, por exemplo, os de ação e comunicação – que, inclusive, podem fornecer uma perspectiva mais estruturada mesmo a respeito do que seria governança privada.

Por “*mais contextualizados*”, refiro-me à circunstância de que justamente porque os instrumentos da sociologia são mais amplos – eles não se dedicam a entender apenas aquilo que ocorre em plataformas da internet, mas são vocacionados a fornecer interpretações a respeito de fatos sociais, termo que utilizo em sentido amplo e variável de acordo com a teoria adotada, o que abordo melhor na seção seguinte –, eles podem ajudar a compreender essa transformação de uma maneira contextualizada, em paralelo com outras questões que não dizem respeito necessariamente às plataformas e à regulação da liberdade de expressão, mas que podem ajudar a entender as suas transformações. Ela viabiliza o



acesso, nesse sentido, a uma espécie de “*caixa de ferramentas*” – composta de conceitos, classificações, formas de inserir transformações em contextos sociais mais amplos, etc. – capaz de preencher a ausência de formas teóricas mais robustas para analisar a moderação de conteúdo e suas transformações, trazendo, para o problema que proponho enfrentar nesta parte do trabalho, um instrumental alinhado com a perspectiva que melhor explica a moderação de conteúdo – a pluralista –, robustecendo-a com uma perspectiva mais ampla e contextualizada.

### *3.2 Por que Teoria dos Sistemas?*

Dediquei a seção anterior a justificar as bases do método que sugiro para enfrentar o problema. Em síntese, ele é sustentado na proposta de que os elementos já disponíveis a respeito da moderação de conteúdo sejam entendidos sob uma perspectiva sociológica robusta, capaz de articulá-los, aventando e testando hipóteses a respeito de como e por que ela se transforma. Mas qual perspectiva sociológica? Sob qual paradigma particular a moderação de conteúdo deve ser vista? Para responder a essas perguntas, é preciso expor, mesmo que em linhas gerais, as estratégias que proponho para resolver o problema.

Minha proposta, em síntese, é utilizar a sociologia jurídica funcionalista para abordar tanto o processo decisório estatal quanto a moderação de conteúdo. A estratégia que proponho, nesse sentido, é superar as limitações das análises que temos disponíveis por meio de uma perspectiva funcionalista que permita entender, primeiro – e independentemente da leitura que os próprios atores têm a esse respeito –, qual é o papel que esse formato específico de decidir – resumidamente identificado sob a ideia de um devido processo legal – tem em sua origem: as decisões tomadas no âmbito do Estado. Com isso, a intenção é isolar a questão de particularidades locais – a regulamentação, caso a caso, desse tipo de procedimento –, extraíndo, a partir de um modelo geral, a função que esse formato decisório tem no seu âmbito de origem. Essa é a vantagem da análise sociológica funcionalista, que a justifica como método: a capacidade de, sob uma perspectiva externa, responder a perguntas quanto ao papel que institutos jurídicos – como, por exemplo, o devido processo legal – têm em fenômenos sociais – a exemplo da tomada de decisões vinculantes – relacionando-os sem o compromisso, inerente às perspectivas internas, de se atingir uma decisão quanto ao direito aplicável ao caso concreto.

Segundo, minha expectativa é conseguir levar esse diagnóstico do Estado às plataformas, identificando como essa função inicialmente desempenhada no âmbito decisório estatal esclarece as mudanças pelas quais passa a moderação de conteúdo, obtendo conclusões que não dependam de percepções pessoais de atores envolvidos nesse processo – tais como as captadas pelas pesquisas disponíveis. Como exponho adiante, essa transição do Estado à plataforma é um movimento a ser feito com cautela, porque as elaborações funcionalistas que temos disponíveis quanto ao devido processo legal foram produzidas a partir de contextos estatais, anteriores mesmo às versões iniciais de movimentos como o do pluralismo jurídico, em que questões quanto à existência de uma produção jurídica em contextos diferentes do estatal foram enfrentadas. Nesse sentido, meu objetivo não é pressupor que a mesma função exercida pelo devido processo legal em contextos estatais é por ele desempenhado nas plataformas, mas sim que a primeira oferece evidências a partir das quais a segunda pode ser melhor investigada.

Dentre as análises sociológicas funcionalistas, adoto, neste trabalho, a Teoria dos Sistemas de Niklas Luhmann. Faço isso porque, primeiro, a Teoria dos Sistemas é uma perspectiva sociológica sustentada no funcionalismo estrutural<sup>519</sup>, e, por isso, é uma ferramenta que viabiliza a estratégia que propus acima, ao permitir tomar as características dos procedimentos decisórios jurisdicionais como contingentes, analisando-as não como consequências de um ordenamento jurídico específico, mas como engrenagens que têm funções externamente identificáveis e que, uma vez isoladas, podem esclarecer as razões pelas quais essas mesmas estruturas são mobilizadas em contextos diferentes do estatal. Como afirma Luhmann, ao estabelecer os marcos metodológicos de sua sociologia jurídica:

*“Inicialmente, o funcionalismo tenta capturar, para além dos manifestos efeitos pretendidos ou que decorrem da ação, as suas funções latentes, para ampliar os horizontes de orientação da ação.. Não se trata apenas de trazer à tona [os efeitos] subentendidos ou esquecidos. Em vez disso, o interesse sociológico é direcionado para a "latência necessária", isto é, às relações funcionais que o ator não pode apreender em sua consciência - seja porque a ação perdeu a sua motivação; seja que se envolveu em contradições; seja porque a possibilidade de conferir atenção a isso tenha sido superada”<sup>520</sup>.*

---

<sup>519</sup> Cf. *supra*, nota n.º 48.

<sup>520</sup> “Zunächst sucht der Funktionalismus neben den manifesten, also bezweckten oder doch mitbedachten Wirkungen auch latent Funktionen zu erfassen, also den Orientierungshorizont des Handelns zu erweitern.

Segundo, porque uma preocupação central da Teoria dos Sistemas é aplicar essa perspectiva funcionalista ao direito, explicando, em específico, o papel desempenhado por elementos característicos de sua estrutura – dentre os quais justamente o procedimento decisório judicial<sup>521</sup>. É, portanto, dessa possibilidade de se extrair da Teoria conclusões já existentes e que resultaram da aplicação de suas ferramentas no fenômeno decisório jurisdicional – explicando funcionalmente o papel que o devido processo legal tem no direito moderno – que espero estabelecer hipóteses quanto ao papel que essas mesmas estruturas têm em contextos decisórios privados, como é o caso da moderação de conteúdo.

Terceiro, adotar a Teoria dos Sistemas é uma estratégia justificável porque enfrentar o problema que proponho neste trabalho exige uma análise de transformações que ocorrem em um espaço de manifestação naturalmente transnacional: as comunidades criadas por plataformas como o Facebook, o Twitter e o YouTube na internet. Como a era digital não significou apenas uma modificação no espaço de manifestação, mas se traduziu na criação de um novo espaço de decisão, exigem-se, agora, ferramentas teóricas capazes de lidar com um formato de criar regras, apurar condutas e fazer julgamentos que é essencialmente desvinculado da ação estatal.

Por isso, se, por um lado, a análise sustentada por este trabalho precisa uma sociologia jurídica que tenha uma robusta estrutura funcionalista – que, com isso, ofereça uma alternativa às abordagens que mencionei antes; por outro, ela demanda ferramentas capazes de lidar com esse novo fenômeno decisório não como algo à margem do direito tomado em sua perspectiva estatal – algo que redundaria na visão diádica –, mas levando a sério a independência dessa nova forma de julgar aquilo que as pessoas podem ou não dizer nesse espaço de manifestação. Para tanto, não basta o funcionalismo – e as conclusões que ele já fornece quanto ao papel desempenhado pelas características do procedimento decisório jurisdicional: é preciso acoplá-lo com alguma modalidade de entendimento plural do fenômeno jurídico, que, ao dissociar direito e Estado, viabilize a

---

*Dabei geht es nicht nur darum, Übersehenes oder Vergessenes ins Bewußtsein zu rücken. Vielmehr richtet sich das soziologische Interesse auch und gerade auf »notwendige Latenz«, das heißt auf funktionale Beziehungen, die der Handelnde nicht in sein Bewußtsein aufnehmen kann - sei es, weil das Handeln dadurch seine Motivierbarkeit verlöre; sei es, daß es sich in Widersprüche verwickelte; sei es, daß das Potential für Aufmerksamkeit dadurch überfordert werden würde''.* LUHMANN, N. **Ausdifferenzierung des Rechts: Beiträge zur Rechtssoziologie und Rechtstheorie**. Frankfurt: Suhrkamp, 1999, p. 281.

<sup>521</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 55/129.

utilização dessas ferramentas sociológicas em espaços transnacionais – como o que produz as decisões da moderação de conteúdo.

Essa transição entre uma sociologia jurídica funcionalista, mas que vê no direito um fenômeno estatal, e uma sociologia jurídica que permanece funcionalista, mas que absorve elementos do pluralismo jurídico, é justamente o que marca a Teoria dos Sistemas nos últimos anos de seu desenvolvimento. Isso porque, ainda que sob pressupostos inovadores, essa é uma teoria que nasce, com Luhmann, ainda na década de 70, em um momento anterior ao pluralismo jurídico. Com isso, ela foi construída ainda sobre uma base essencialmente vinculada ao Estado, isto é, a partir de uma associação, mesmo que implícita, entre os fenômenos jurídicos e estatais. Isso desencadeou, na Teoria dos Sistemas, exatamente o mesmo movimento teórico que é necessário fazer, neste trabalho, para enfrentar o problema que descrevi acima: com a formação de novos espaços decisórios desvinculados do Estado, um dos principais intérpretes de Luhmann, em específico quanto às relações privadas, Gunther Teubner, viu-se na necessidade de operar justamente essa transição<sup>522</sup> entre um direito essencialmente vinculado ao fenômeno estatal para a possibilidade de uma criação jurídica<sup>523</sup> – e mesmo constitucional<sup>524</sup> – não-estatal.

Por isso, a Teoria dos Sistemas detém, enquanto sociologia jurídica, uma estrutura marcada por uma trajetória em que tanto reteve os seus aspectos funcionalistas – capazes, portanto, de fornecer interpretações quanto ao papel de institutos jurídicos como o devido processo legal – quanto inovou pela apreensão do pluralismo jurídico, nos termos em que reconstruída por Teubner – viabilizando, assim, que novos fenômenos decisórios como o que me proponho a estudar não sejam vistos como periféricos, mas também constitutivos do direito. Ela ficou, em síntese, com o melhor dos dois mundos – e, por isso,

---

<sup>522</sup> Cf., TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, pp. 255/290, 1996; e TEUBNER, G. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne, in SCHIMDT, E., WEYERS, H. (ed.) **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, pp. 191/214, 1995.

<sup>523</sup> Cf., TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, pp. 255/290, 1996; e TEUBNER, G. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne, in SCHIMDT, E., WEYERS, H. (ed.) **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, pp. 191/214, 1995.

<sup>524</sup> Cf., também introdutoriamente, TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, pp. 1/28, 2003.

junge as ferramentas sociológicas adequadas para a análise de um trabalho que propõe fazer uma análise funcionalista de um fenômeno decisório transnacional.

#### 4. CONCLUSÃO

Neste capítulo, meu objetivo foi apresentar, em síntese, três diferentes elementos que irão conduzir a segunda parte deste trabalho: um problema, duas hipóteses e um método. Superada a opacidade típica da moderação de conteúdo pelo primeiro e segundo capítulos desta dissertação, eles compõem o seu segundo objetivo.

O problema ao qual se dedica esta segunda parte resume-se em compreender as razões por trás da transformação que ocorre na forma como as principais plataformas da internet moderam conteúdo. Como exemplos extraídos do Facebook, Twitter e YouTube expõem, cada vez mais o procedimento adotado pelas plataformas para moderar conteúdo se afasta do modelo arbitrário que o marcou, fazendo-o no sentido de adotar mecanismos de transparência, participação e recorribilidade que lhe aproximam do devido processo legal utilizado pelos Estados para produzir decisões vinculantes. A questão é entender por que isso tem acontecido, uma vez que as interpretações que temos disponíveis a respeito da moderação de conteúdo – tanto as diádicas quanto as pluralistas – não conseguem explicar esse movimento.

Em síntese, minhas hipóteses a esse respeito são de que essa transformação se explica, primeiro, como uma forma pela qual as plataformas obtêm cada vez mais legitimidade para tomar esse tipo de decisão, mesmo que isso signifique abrir mão do poder envolvido; e, segundo, que esse é um movimento inserido em uma transformação mais ampla, a da constitucionalização do mundo digital, capaz de prover esses nichos decisórios com autonomia tanto em relação às estruturas corporativas das quais provêm, quanto daquilo que lhes é externo.

Como método capaz de resolver esse problema e provar essas hipóteses, adoto a aplicação, à moderação de conteúdo, da perspectiva funcionalista específica da Teoria dos Sistemas de Niklas Luhmann. Minha intenção, com ela, é superar as limitações das abordagens anteriores e inserir a moderação de conteúdo em um esquema teórico robusto, capaz de articular suas características, aventar hipóteses a seu respeito e demonstrá-las.

Estruturalmente, faço isso, nos dois capítulos seguintes, correspondentes às etapas do meu argumento. Na primeira etapa, correspondente ao Capítulo 4, analiso como

o devido processo legal funciona nos Estados, conferindo-lhes, por meio de um procedimento, legitimidade decisória. Faço isso reconstruindo o argumento de Niklas Luhmann no texto a *Legitimação pelo Procedimento*. Na segunda etapa, correspondente ao Capítulo 5, utilizo essa relação estabelecida por Luhmann entre procedimento e legitimidade nos Estados para investigar, a partir do quadro teórico construído por Gunther Teubner – que faz a transição da Teoria dos Sistemas para uma perspectiva capaz de abordar a moderação de conteúdo das plataformas da internet –, as razões por trás da adoção do devido processo legal em um contexto não-estatal.

## CAPÍTULO 4 – DO PROCEDIMENTO À LEGITIMIDADE

### 1. INTRODUÇÃO

No capítulo anterior, apresentei o problema da segunda parte da dissertação. Em síntese, seu objetivo é investigar o funcionamento da moderação de conteúdo das plataformas da internet para entender por que o procedimento adotado por essas empresas para decidir aquilo que os seus usuários podem dizer tem se transformado e, dessa forma, expor com mais precisão como decisões que moldam a regulação da liberdade de expressão na era digital são tomadas. A transformação que permanece inexplicada é sintetizada na constatação de que, nos últimos dois anos, as plataformas da internet têm implantado, em seus procedimentos de moderação de conteúdo, mecanismos de transparência, participação e recorribilidade que os têm distanciado do modelo arbitrário anterior, aproximando-os do paradigma do devido processo legal pelo qual os Estados tomam decisões vinculantes – mesmo que isso tenha significado perder, ao menos em parte, o poder que têm sobre o que as pessoas dizem nos ambientes que administram.

Também expus, no capítulo anterior, minha estratégia para abordar esse problema. Ela é dividida em dois passos: primeiro, utilizar a sociologia jurídica funcionalista e, em específico, a Teoria dos Sistemas de Niklas Luhmann, para investigar o papel desempenhado pelos procedimentos na estrutura política dos Estados. Segundo, trazer esse modelo dos Estados às plataformas, utilizando ferramentas teóricas – em particular, a adaptação feita por Gunther Teubner da Teoria dos Sistemas ao pluralismo jurídico – para identificar, a partir de uma análise funcionalista, o papel desempenhado pelos procedimentos na moderação de conteúdo.

Este capítulo se dedica ao primeiro desses dois passos. Meu objetivo, com ele, é responder à seguinte pergunta: qual é a função que os procedimentos têm na tomada de decisões na estrutura política dos Estados? Para responder a essa pergunta, reconstruo o raciocínio desenvolvido por Niklas Luhmann no livro *Legitimação pelo Procedimento* (*Legitimation durch Verfahren*), em que é conferido um novo significado “*para os modernos sistemas políticos [...] [às] ideias jurídicas centrais do processo legal, com a ajuda de meios sociológicos e, principalmente, com meios da teoria dos sistemas*”<sup>525</sup>. De

---

<sup>525</sup> “Dieses Buch versucht, die für moderne politische Systeme zentrale Rechtsidee des Verfahrens mit sozialwissenschaftlichen und speziell mit systemtheoretischen Mitteln nachzukonstruieren”. LUHMANN, N.

acordo com esse argumento, o procedimento é um instrumento criado para legitimar decisões, porque cria condições para que elas sejam construídas neutralizando decepções e viabilizando que as pessoas as aceitem, mesmo que lhes sejam desfavoráveis. Há, portanto, de acordo com Luhmann, uma relação entre legitimidade e procedimento, que justifica que os Estados adotem essa forma de decidir.

Estruturalmente, o capítulo está centrado em uma reconstrução do texto de Luhmann (2); seguida de uma conclusão (3), que sistematiza os resultados dessa abordagem e expõe as suas limitações, colocando em evidência a necessidade de que seja dado um passo teórico adicional para que a abordagem funcionalista da Teoria dos Sistemas supere a relação necessária construída por Luhmann entre direito e Estado e, assim, torne-se útil para entender as transformações pelas quais passam as plataformas da internet.

## 2. LEGITIMAÇÃO PELO PROCEDIMENTO

### 2.1 Introdução

A seção que segue é dedicada a entender, a partir do funcionalismo estrutural – e, em específico, da Teoria dos Sistemas de Niklas Luhmann – qual é o papel desempenhado pelos procedimentos na estrutura dos Estados. Para fazer isso, minha proposta é reconstruir o argumento desenvolvido por Luhmann em um texto específico, a *Legitimação pelo Procedimento*, que é dedicado a responder a essa pergunta. Esse não é o livro mais recente de Luhmann em que a existência e a função dos procedimentos jurídicos são tratadas: esse tema também é abordado, por exemplo, em *O Direito da Sociedade*<sup>526</sup>,

---

**Legitimation durch Verfahren.** Frankfurt: Suhrkamp, 1983, p. 1; LUHMANN, N. **Legitimação pelo Procedimento.** Brasília: Ed. Universidade de Brasília, 1980, p. 9. As traduções foram realizadas a partir do original em alemão da 2.<sup>a</sup> edição do texto, reimpresso em 1983, em cotejo com a sua versão em português. Por isso, ambas serão referenciadas no decorrer do capítulo.

<sup>526</sup> A primeira edição de *Legitimation durch Verfahren* data de 1969; a segunda, é de 1975. *O Direito da Sociedade (Das Recht der Gesellschaft)* é um texto de 1993. Mencionei o último como um marco exemplificativo; há outros textos de Luhmann em que a função sociológica do procedimento é tratada; cf., por exemplo, no contexto da positividade do direito, LUHMANN, N. **Rechtssoziologie.** Opladen: Westdeutscher Verlag, 1983, pp. 234 e ss; em um contexto sistêmico pós-autopoietico, LUHMANN, N. **Soziale Systeme: Grundriß einer allgemeinen Theorie.** Frankfurt: Suhrkamp, 1991, pp. 597 e ss (a primeira edição é de 1984); como parte da proposta epistemológica da Teoria dos Sistemas, LUHMANN, N.



publicado vinte anos depois e que, por isso, traz uma Teoria dos Sistemas mais madura. Diferente do que acontece em a *Legitimação pelo Procedimento*, entretanto, em *O Direito da Sociedade* a existência de procedimentos é tratada mais em função do papel que eles exercem na estrutura do direito<sup>527</sup> do que na do Estado – o que torna a primeira obra mais interessante para o objetivo que tenho nesta parte do trabalho.

Dividi a apresentação do argumento em duas seções: na primeira (2.2), exponho o problema e a hipótese às quais Luhmann dedica o texto. Isso é feito em dois passos: primeiro (2.2.1), delimitando tanto o tema do trabalho – o papel que procedimentos têm na estrutura do Estado – quanto as deficiências que caracterizam as abordagens disponíveis a esse respeito; segundo (2.2.2), expondo a hipótese de que a função social dos procedimentos – dentre os quais o processo adjudicatório – é a de conferir legitimidade às decisões que produz, traduzindo-se, por parte dos atores envolvidos, em uma maior aceitabilidade de seus resultados – mesmo que desfavoráveis –, neutralizando o potencial destrutivo do conflito gerado a partir da divergência. Na segunda seção (2.3), reconstruo a abordagem utilizada por Luhmann para demonstrar essa hipótese, em particular identificando nos três principais procedimentos da sociedade moderna – os processos judiciais (2.3.1), e os processos eleitorais e legislativos (2.3.2) – uma estrutura comum que sustenta a tomada de decisão legítima na sociedade moderna.

## 2.2 Problema e hipótese: a legitimação pelo procedimento na teoria

### 2.2.1 O problema

O primeiro passo para reconstruir o argumento de Luhmann no texto a *Legitimação pelo Procedimento* é identificar o problema que ele se propõe a resolver. Farei isso, a seguir, em dois momentos: primeiro, delimitando-o a partir do tema do livro – o papel que procedimentos têm na estrutura do Estado – e, segundo, expondo as deficiências que as respostas disponíveis têm em abordá-lo, o que, segundo Luhmann, abre espaço para uma abordagem diferente a esse respeito – e que é o objeto da seção seguinte.

---

**Soziologische Aufklärung 6: Die Soziologie und der Mensch.** Opladen: Westdeutscher Verlag, 1995, pp. 243 e ss; e como parte da última versão da Teoria de Luhmann, LUHMANN, N. **Die Gesellschaft der Gesellschaft.** Frankfurt: Suhrkamp, 1998, v. 1, pp. 468 e ss.

<sup>527</sup> LUHMANN, N. **Das Recht der Gesellschaft.** Frankfurt: Suhrkamp, 1995, pp. 103 e ss.

Luhmann esclarece, no prefácio, o seu objetivo: “reconstruir para os modernos sistemas políticos as ideias jurídicas do processo legal com a ajuda de meios sociológicos e, principalmente, com meios da teoria dos sistemas”<sup>528</sup>. Seu texto se concentra, portanto, em fornecer um aporte sociológico à forma como os procedimentos jurídicos são compreendidos no interior dos sistemas políticos modernos. Há dois pressupostos implícitos nessa proposta: primeiro, o de que esses procedimentos têm um papel relevante na estruturação política da sociedade moderna – caso contrário não haveria sentido em compreendê-los a partir da função que exercem nela; segundo, que há problemas nas abordagens disponíveis para entender esse papel.

Luhmann detecta, nesse sentido, que a existência de procedimentos juridicamente regulados é um dos “atributos mais extraordinários do sistema político das sociedades modernas”<sup>529</sup>. Eles estão por toda parte dessa estrutura, “praticados por formas tão diversas pelas autoridades, tribunais, parlamentos, em eleições, planejamentos, aplicações jurídicas, decisões de adjudicação, etc.”<sup>530</sup>. Essa onipresença do procedimento como um elemento que constitui a sociedade política moderna contrasta com a ausência de uma “teoria unificadora de todos os procedimentos”<sup>531</sup>: não é difícil descrever, mesmo intuitivamente, espécies procedimentais – a eleição, por exemplo, é uma sequência de fases organizadas para que se escolha um representante; o processo judicial se volta à produção de uma decisão jurisdicional, etc. –, mas explicar aquilo que unifica essas situações diferentes e corriqueiras da estrutura da sociedade moderna sob a rubrica comum de

---

<sup>528</sup> “Dieses Buch versucht, die für moderne politische Systeme zentrale Rechtsidee des Verfahrens mit sozialwissenschaftlichen und speziell mit systemtheoretischen Mitteln nachzukonstruieren”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 1; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 9.

<sup>529</sup> “Rechtlich geordnete Verfahren der Entscheidungsfindung gehören zu den auffälligsten Merkmalen des politischen Systems moderner Gesellschaften”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 11; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 17.

<sup>530</sup> “[...] die in so vielgestaltiger Form durch Behörden, Gerichte, Parlamente, bei Wahlen, Planungen, Rechtsanwendungen, Zuteilungsentscheidungen usw. praktiziert werden”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 11; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 17.

<sup>531</sup> “[...] der das Entstehen einer einheitlichen Theorie aller rechtlich geregelten Verfahren verhindert hat [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 11; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 17.

“*procedimento*” é um desafio diferente e – surpreendentemente – inexplorado. Com isso, essa categoria permanece sob uma elaboração teórica baseada ou em “*premissas não formuladas e transcrições de objetivos vinculados a determinados valores*”<sup>532</sup> ou, juridicamente, em abordagens dogmáticas a respeito do direito processual – civil, penal, eleitoral, etc.. Dessa forma, ou o procedimento é tratado de forma incipiente, ou, quando é sistematizado em uma abordagem teórica, volta-se a “*conceber empiricamente o comportamento fático das partes processuais e de o contrastar com as normas de interesse maior*”<sup>533</sup> – prestando-se a dizer, por exemplo, se uma eleição ou um processo é válido ou nulo.

Mesmo que propondo uma perspectiva diferente, as abordagens anteriores são importantes para o argumento de Luhmann porque é a partir de suas deficiências que ele se propõe a entender o problema de uma forma diferente. Essas leituras se concentram, segundo ele, ao redor dos “*procedimentos que conseguiram uma importância especial, [porque] estruturalmente decisivos nos sistemas políticos atuais: o procedimento de eleição política, o procedimento parlamentar da legislação e o processo judicial*”<sup>534</sup>, identificando os objetivos utilizados como justificativas para a implantação e a continuidade no uso desses mecanismos decisórios na sociedade moderna. O problema, segundo Luhmann, é que essas justificativas apresentam obscuridades que tornam a higidez desses fundamentos duvidosa, abrindo caminho para uma análise sociológica que os coloque à prova.

No caso da eleição política, o discurso indica que seu objetivo é a “*ocupação das instâncias politicamente decisórias com pessoas especialmente capazes e*

---

<sup>532</sup> “*Auf Grund solcher Erwartungen kann man so etwas wie eine klassische Konzeption des Verfahrens vermuten, aber sie steckt mehr in unformulierten Prämissen und wertgebundenen Zweckumschreibungen als in einer ausgearbeiteten, kritikfähigen Theorie*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 12; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 17.

<sup>533</sup> “*Dabei kann es nicht lediglich darum gehen, das faktische Verhalten der Verfahrensbeteiligten empirisch zu erfassen und gegen die Norm in Kontrast zu setzen*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 12; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 18.

<sup>534</sup> “*Wir wählen dafür Verfahren, die im heutigen politischen System besondere, strukturtragende Bedeutung gewonnen haben: das Verfahren der politischen Wahl, das parr lamentarische Gesetzgebungsverfahren und den gerichtlichen Prozeß*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 13; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 18.

que tomarão corretamente decisões, isto é, de acordo com a vontade do povo que, neste sentido, podem, portanto representar verdadeiramente”<sup>535</sup>. Escolhem-se pessoas por um procedimento eleitoral porque ele garante que elas serão as melhores para tomar decisões em nome das outras. Sob a perspectiva procedimental, isso se traduz no estabelecimento de regras que orientam a organização da eleição, conformando a sequência fática que leva a um resultado decisório – no caso, a escolha de alguém a um cargo público – de modo que ela garanta a ampla concorrência para o cargo, o sigilo da votação, que os votos sejam contados de maneira correta, que as pessoas possam manifestar livremente sua vontade, etc.. Esses elementos procedimentais são colmatados por “*conjecturas ideológicas*”<sup>536</sup>, como, por exemplo, a de que há uma vontade geral e não expressa em eleitores isolados, bem informados e racionalmente orientados e que é colhida pelo procedimento eleitoral, gerando uma relação de causalidade entre existir o procedimento e se ter um determinado resultado, no caso, a “*ocupação das instâncias politicamente decisórias com pessoas especialmente capazes*”. Essa relação de causalidade é importante, porque dela decorre uma justificativa para a existência desse procedimento: fazemos processos eleitorais orientados pela ampla concorrência *porque*, dessa forma, escolhemos pessoas de acordo com a vontade do povo e que tomarão as decisões corretas.

Luhmann sustenta, entretanto, que essas são hipóteses que advêm, no limite, da “*experiência*”: elas são “*empiricamente insustentáveis*”<sup>537</sup>, porque “*permanece obscura a forma como esse objetivo pode ser assim atingido*”<sup>538</sup>. É difícil de vislumbrar, nesse sentido, como seria possível sequer provar a relação de causalidade que existe entre se estabelecer um procedimento eleitoral e se escolher a pessoa que tanto representa a

---

<sup>535</sup> “[...] *die Besetzung der politisch entscheidenden Instanzen mit besonders befähigten Personen, die richtig, das heißt nach Maßgabe des Volkswillens, entscheiden werden, die in diesem Sinne also wahr repräsentieren können*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 12/13; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 19.

<sup>536</sup> “[...] *ideologische Zusatzannahmen* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 14; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 19.

<sup>537</sup> “[...] *empirisch unhaltbar* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 14; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 19.

<sup>538</sup> “*Wie jenes Ziel auf diesem Wege erreicht werden kann, bleibt zunächst unklar*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 14; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 19.

vontade do povo, quanto é capaz de tomar as decisões corretas. Isso gera dois problemas: primeiro, enfraquece essa causalidade como algo capaz de justificar a existência do procedimento; e, segundo, torna impossível se sustentar, a partir dela, uma perspectiva geral quanto ao sentido da existência de procedimentos na estrutura política da sociedade moderna. Como afirma Luhmann, “*com fundamento nessas hipóteses clássicas, não se podia desenvolver uma teoria do procedimento*”<sup>539</sup>.

No caso dos procedimentos parlamentares, também há um sentido que os justifica oficialmente: leis são debatidas em instituições por representantes para que se chegue à melhor solução para situações não programadas – isto é, que não estavam antes previstas na legislação. Assim, os parlamentares são reunidos em ambientes que garantam a sua livre manifestação, para que, atrelados apenas à sua própria consciência, possam propor e ponderar, para determinados problemas sociais, soluções que tomam a forma de proposições legislativas. Procedimentalmente, isso é assegurado por regras que tanto garantem que os legisladores podem dizer aquilo que pensam a respeito do assunto debatido – as imunidades parlamentares –, quanto fixam como ilícitas influências externas capazes de impactar nessa expressão da consciência própria que caracteriza essa atividade, blindando-a da “*coaçoão de um grupo político, [d]a permuta de vantagens concretas e [do] entrelaçamento de papéis*”<sup>540</sup>. Como, quando isso acontece, geram-se situações em que o resultado da deliberação pode não refletir o posicionamento individual de cada um a respeito da solução que melhor se afeiçoa ao problema “*não programado*”, criam-se mecanismos procedimentais capazes de afastar esse perigo e reestabelecer, assim, o vínculo particular que existe entre consciência e manifestação parlamentar. Também aqui há um fundo ideológico que dá corpo a esse procedimento, baseado na ideia de que “*a concorrência de opiniões é suficiente para garantir a verdade*”<sup>541</sup>: para que as melhores soluções para os problemas não normatizados sejam atingidas, basta criar as condições

---

<sup>539</sup> “*Auf Grund jener klassischen Annahmen war keine Theorie des Verfahrens zu entwickeln*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 14; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 19.

<sup>540</sup> “*Deshalb sind Weisungsgebundenheit, Fraktionszwang, Tausch konkreter Vorteile, Rollenverflechtungen suspekt, wenn nicht verboten*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 15; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 20.

<sup>541</sup> “[...] *daß Konkurrenz der Meinungen genüge, um Wahrheit zu sichern*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 16; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 20.

para que exista um debate livre entre os representantes do povo que, ponderando as respostas possíveis, certamente irão escolher a mais apropriada.

Como na escolha eleitoral, essa relação causal entre se ter um procedimento conformado de determinada maneira – no caso, imunizando a manifestação parlamentar de interferências externas – e se obter, por consequência, um resultado específico – as melhores soluções normativas para os problemas não programados da sociedade – é frágil. Por que dessa livre ponderação de posicionamentos deveria emergir a melhor solução para um problema? Em particular: como é possível garantir, por um procedimento parlamentar, que a melhor solução será atingida, se virtualmente qualquer coisa pode ser objeto de uma legislação? Como é possível “*harmonizar este objetivo com aqueles meios*”<sup>542</sup>?

No caso dos procedimentos judiciais, o discurso justificador toma a forma de uma perspectiva que imputa ao processo judicial tanto a busca pela verdade – isto é, a descoberta do que “*é legalmente válido e é legal no caso específico*”<sup>543</sup> – quanto a possibilidade de conferir uma proteção jurídica às partes, que são tuteladas por garantias procedimentais frente ao arbítrio do Estado responsável pela decisão. Os mesmos elementos dos casos anteriores repetem-se aqui. Novamente, há a construção de uma relação de causalidade que vincula o procedimento e suas regras – concretizadas, por exemplo, na exigência de que se tenha um juiz imparcial, que ele fundamente suas decisões, etc. – a resultados específicos, como a obtenção da verdade e a tutela do indivíduo contra o arbítrio. Assim, a “*cláusula do ‘due process’ da constituição americana tem o seu problema relativo na força da burocracia e nos perigos de seu abuso*”<sup>544</sup>. Também aqui, essa é uma relação frágil e que dificilmente se presta a fundamentar, sob uma perspectiva funcional, a existência do procedimento: como explicar, por exemplo, a decisão que é válida, porque rigorosamente sujeita às etapas do procedimento, mas incorreta? Como afirma Luhmann, “*não há suficientes instruções para aqueles que*

---

<sup>542</sup> “[...] *wie dieses Ziel und jene Mittel harmonieren*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 15; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 20.

<sup>543</sup> “[...] *was als Recht gilt und im Einzelfall Rechtens ist*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 17; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 21

<sup>544</sup> “[...] *die ‘due process’-Klausel der amerikanischen Verfassung hat ihr Bezugsproblem in der Macht der Bürokratie und den Gefahren ihres Mißbrauchs*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 19; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 22.

*decidem [...] para dar como resultado decisões justas ou injustas, por amor da verdade ou por amor da paz?*”. Inexiste maneira de se garantir, àqueles que são objeto da decisão produzida, que o procedimento ocasionará a verdade ou o justo, ao ponto de esses serem elementos únicos capazes de justificá-lo empiricamente.

Os três casos diferenciam-se de acordo com o âmbito de atuação do procedimento – eleitoral, legislativo ou jurisdicional – mas têm semelhanças na forma como o explicam: em todos, há um discurso que justifica a existência desse instrumento de decidir a partir de uma relação de causalidade, que justifica as suas regras a partir de um resultado julgado positivo para a estrutura política moderna. Elegem-se pessoas por um procedimento que garante a ampla competição porque isso gera escolhas melhores; propostas legislativas são discutidas por um procedimento que garante a manifestação da própria consciência do parlamentar porque isso leva a melhores soluções de problemas não-programados; decisões jurisdicionais são proferidas após procedimentos em que o juiz é imparcial porque assim eles produzem a verdade e protegem as partes.

Luhmann consolida esses discursos justificadores sob a ideia de que eles constroem uma relação causal entre procedimentos – sejam eles eleitorais, legislativos ou jurisdicionais – e uma ideia geral de verdade, entendida como um ideal superior àqueles que detêm o poder e que pode ser atingido, desde que as regras desse jogo procedimental sejam observadas<sup>545</sup>. Pelos procedimentos, descobre-se a verdade de quem deve ser escolhido em uma eleição – acima daquilo que os eleitores individualmente querem; a verdade de qual solução deve ser acolhida para um problema social não programado – acima daquilo que os parlamentares individualmente querem; e a verdade de qual decisão deve ser tomada em um processo judicial – acima daquilo que as partes individualmente requerem. O problema é que, mesmo que o procedimento seja seguido à risca, por papéis sociais individualizados que cumpram, passo a passo, aquilo que a sua cartilha determina – como os de eleitor, parlamentar e juiz –, *ele não é uma garantia, por si só, de que a verdade será atingida*. Há eleições justas que escolhem pessoas inapropriadas, leis válidas que não resolvem problemas sociais e decisões jurisdicionais submetidas à cartilha processual dogmática que erram ao aplicar o direito – e todas essas hipóteses são considerados resultados normais que podem advir desse formato de decidir.

A verdade, nesse sentido, não pode ser o elemento que justifica e confere sentido ao procedimento, baseando-o como se fosse uma constante que sempre pudesse ser

---

<sup>545</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 20.

encontrada nele, simplesmente porque ela nem sempre ocorre. Ela, que é tomada por essas perspectivas iniciais como um ponto de apoio fixo, é, na realidade, parte da contingência do procedimento – como um resultado tanto possível quanto desnecessário, que pode ou não acontecer<sup>546</sup>. Tanto que faz parte da ideia de procedimento a existência de mecanismos destinados justamente a corrigir esses erros: pelo *recall*, a eleição de alguém inapropriado para o cargo é revertida; pela revogação de uma lei, uma regra que não é adequada deixa de gerar efeitos; pelo recurso, uma decisão jurídica incorreta é retificada. Isso expõe a necessidade de que uma forma diferente de abordar a ideia de procedimento seja construída, em particular não mais o apreendendo a partir da verdade que ele causa – porque, por vezes, esse resultado não existe.

Para Luhmann, a chave dessa nova perspectiva é uma visão sociológica particular, que não “*concebe a verdade já não mais como um valor, mas, mais precisamente, como um mecanismo social que permite algo determinado*”<sup>547</sup>. A teoria geral do procedimento que ele propõe é diferente, portanto, porque, por um lado, não é uma forma mais abstrata da dogmática processual: o tipo de pergunta que a norteia não é se alguém tem o direito ao contraditório em uma situação processual específica, mas sim porque existe algo como o direito ao contraditório na estrutura política moderna. Por outro, ela desvenda o “*sentido de regulações jurídicas*” procedimentais, isto é, as razões por trás de se ter um direito posto dessa ou daquela maneira, sem se deixar enganar pela relação discursivamente construída entre processo e verdade: ao se perguntar porque existe algo como o direito ao contraditório, ela não se deixa levar pela ideia de que assim se resguarda melhor uma garantia de alguém ou porque essa é a forma de se produzir uma decisão mais justa.

É, portanto, da fragilidade tanto dos vínculos causais estabelecidos entre a existência e o regramento dos procedimentos e a obtenção da verdade, quanto da própria

---

<sup>546</sup> Em toda a reconstrução teórica que segue neste e nos próximos capítulos, adotarei, como linha de princípio, a postura de acessar a Teoria dos Sistemas na medida do necessário, evitando digressões que não tenham pertinência com a solução do problema da dissertação. Assim, mobilizarei a seguir o conceito de contingência apenas na medida em que ele auxilia a compreensão da função do procedimento, evitando intencionalmente trazê-lo, por exemplo, como um ponto central de distanciamento entre o funcionalismo de Parsons e de Luhmann (cf. LUHMANN, N. **Aufsätze und Reden**. Stuttgart: Philipp Reclam, 2001, pp. 10 e ss).

<sup>547</sup> “[...] *die Wahrheit nicht länger nur als Wert, sondern genauer als einen sozialen Mechanismus begreift* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 23; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 25.



ideia de verdade como um resultado necessário do procedimento, que Luhmann extrai a justificativa para o seu trabalho: sua proposta é realizar um aporte capaz de revelar o sentido de existência de procedimentos no sistema político da sociedade moderna, esclarecendo-o não pela dogmática jurídica ou por associações causais com uma verdade necessária, mas por uma abordagem sociológica de cunho funcionalista. A pergunta que perpassa o livro, por isso, é a seguinte: qual é a *função* do procedimento?

### 2.2.2 A hipótese

A seção anterior pode ser resumida em uma constatação: a forma como os procedimentos que constituem a estrutura política da sociedade moderna são compreendidos é questionável. Isso porque eles são percebidos, implantados e justificados a partir de sua habilidade de solucionar problemas que nem sempre são capazes de resolver. Os procedimentos existem, de acordo com essa compreensão inicial, porque conseguem produzir verdades – as melhores escolhas em eleições, as soluções apropriadas para problemas não normatizados e as decisões juridicamente corretas no exercício da jurisdição. Acontece que essa relação causal entre, de um lado, os procedimentos, e, de outro, as consequências que eles dizem oferecer, é frágil – e isso coloca em dúvida a explicação que os justifica a partir desses resultados. Mas se os procedimentos não fazem aquilo que prometem fazer, então qual é a função deles na sociedade?

Para responder a essa pergunta, a proposta de Luhmann é mobilizar o funcionalismo sociológico como um instrumento capaz de explicar o papel desempenhado pelos procedimentos na estrutura política moderna. Como ele afirma, “*a análise funcionalista é uma técnica de descoberta de problemas já resolvidos*”<sup>548</sup>: ela não transita dos problemas às soluções, mas, inversamente, detecta soluções e se propõe a descobrir qual é o problema que a justifica. Ela traduz, nesse sentido, a pergunta do parágrafo anterior em algo mais específico: qual é o problema que é resolvido pelos procedimentos? Que problema há na estrutura política da sociedade moderna para que ela precise de procedimentos para funcionar?

---

<sup>548</sup> “*Funktionale Analyse ist eine Technik der Entdeckung schon gelöster Probleme*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 6; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 13.

Luhmann oferece, em três passos, uma hipótese para responder a essa pergunta: primeiro, ele investiga a relação entre verdade e procedimento; segundo, constrói um paralelo entre verdade e poder; e, terceiro, pergunta-se qual é a relação que existe entre o poder e o procedimento. No que segue, reconstruo esse caminho.

O ponto de partida para a construção dessa hipótese é a relação entre procedimento e verdade. Como a seção anterior expôs, Luhmann utiliza o termo “*verdade*” para unificar os resultados causais que, de acordo com a perspectiva inicial, decorreriam dos procedimentos políticos: a escolha do melhor representante, a solução legislativa apropriada e a decisão jurisdicional correta. Esses procedimentos, entretanto, não são capazes de produzir esses resultados: em sociedades modernas, verdades não decorrem de procedimentos eleitorais, legislativos ou jurídicos. Elas resultam, antes, da aplicação de uma metodologia científica própria, diferenciada e independente de posições pessoais de eleitores, legisladores e juízes.

Um exemplo pode ajudar a esclarecer esse ponto. Suponha que é preciso definir qual é a decisão apropriada – isto é, a “*verdade*” – que se deve tomar diante da existência do alto número de mortes no trânsito. Como descobrir uma solução para esse problema? A resposta, em uma sociedade moderna, é clara: a partir de uma metodologia científica, que toma essa pergunta como um problema de pesquisa, avança hipóteses e as comprova colhendo e sistematizando, por exemplo, informações que estabeleçam vínculos de correlação e/ou causalidade entre o número de mortes nas estradas e a existência de sanções mais rígidas para quem viola as regras de trânsito. Essa análise, que obedece a um procedimento científico, produz uma verdade no sentido preciso da palavra: uma conclusão tecnicamente fundamentada de que, por exemplo, sanções mais duras devem – ou não devem – ser impostas a quem dirige embriagado. Esse resultado – e esse procedimento – são diferentes do processo legislativo, típico do âmbito político da sociedade: nada garante que, ao assegurar as condições adequadas ao debate, o resultado das deliberações parlamentares seja uma política pública apropriada para resolver um problema social. Como qualquer decisão pode ser transformada em uma lei, é possível que a deliberação legislativa resulte exatamente no contrário do que decorre do procedimento científico – uma lei que, por exemplo, flexibiliza no lugar de endurecer as punições de trânsito.

O procedimento político, portanto, produz uma decisão, também essa política. A verdade científica, de outro forma, é o resultado de algo que ocorre em um âmbito à parte da política – a ciência –, e que tem, por isso, uma função própria: “*a*

*transmissão reduzida de complexidade*”<sup>549</sup>. Para entender o que isso significa, é preciso dar um passo atrás e compreender alguns dos pressupostos que permeiam essa posição de Luhmann.

Luhmann associa a ideia de sociedade moderna à de *complexidade*. Essa constatação, que aparece de forma embrionária na *Legitimação pelo Procedimento*<sup>550</sup>, pode ser resumida na afirmação de que *sociedades modernas são sociedades complexas*. Com a palavra “*complexidade*”, Luhmann designa a característica de que, nesse tipo de sociedade, há “*mais possibilidades do que se podem realizar*”<sup>551</sup>: o mundo moderno, diz ele, é complexo porque os seus participantes vivem uma realidade em que há mais cursos de ação do que é possível adotar. Por isso, sempre é preciso *escolher*: diante da “*multiplicidade de possíveis experiências e ações*”<sup>552</sup>, o indivíduo precisa eleger uma dentre elas para seguir, lidando com uma pressão seletiva que caracteriza a “*situação existencial*”<sup>553</sup> da modernidade. Essa, entretanto, não é uma situação que se deixa resolver facilmente: justamente em razão da complexidade, escolher um curso de ação é uma tarefa difícil, porque o indivíduo detém uma capacidade limitada de assimilar informações. Ele não consegue apreender tudo que é necessário, de todas as opções disponíveis, para conseguir selecionar aquela que é a mais apropriada. Por isso, as pessoas são compelidas a utilizar mecanismos sociais que atenuem a complexidade, reduzindo o número de possibilidades de ação e a pressão seletiva, *acessando escolhas que já foram feitas antes*

---

<sup>549</sup> “*Was Wahrheit im sozialen Verkehr leistet, ist Übertragung reduzierter Komplexität*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 23; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 25. Para seguir a linha desenvolvida no capítulo, optei por uma abordagem que reconstrói o tratamento que Luhmann confere à ideia de complexidade em *Legitimação pelo Procedimento*; uma versão mais abstrata – mas coerente à anterior – dessa ideia consta, em sua versão teórica final, em LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 134/145; *cf.*, também, LUHMANN, N. **Soziologische Aufklärung 2: Aufsätze zur Theorie der Gesellschaft**. Wiesbaden: Springer, 1975, pp. 206/208.

<sup>550</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 23.

<sup>551</sup> “*Unter Komplexität wollen wir verstehen, daß es stets mehr Möglichkeiten gibt, als aktualisiert werden können*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 31.

<sup>552</sup> “[...] *eine Fülle von Möglichkeiten des Erlebens und Handelns* [...]”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 31.

<sup>553</sup> “*In dieser Daseinslage* [...]”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 31.

por alguém. Elas absorvem o trabalho de outra pessoa como se fosse o seu, e, assim, conseguem decidir.

É aqui que Luhmann localiza a verdade<sup>554</sup>: ela é um mecanismo que reduz a complexidade, porque disponibiliza às pessoas as seleções feitas por um procedimento científico, que adstringe as escolhas dos indivíduos a um número menor de possibilidades e, assim, ajuda-os a decidir. É como se as pessoas pudessem, por meio da verdade, iniciar o processo de decisão a partir de um ponto adiante, ao qual são levadas por um trabalho anterior, feito por um terceiro em que confiam<sup>555</sup>.

Mais uma vez, um exemplo pode ser útil para ilustrar esse ponto. Um jovem que atinja o final do ensino médio precisa decidir qual faculdade irá cursar. Ele enfrenta, nesse momento, a realidade da complexidade da sociedade moderna, porque se vê diante de múltiplas possibilidades, que constituem mais cursos de ação do que é possível adotar: seu campo de opções pode variar de Medicina às Artes Cênicas. Não é possível, entretanto, fazer todos os cursos disponíveis: ele precisa, assim, *escolher*. O problema é que, justamente em razão dessa multiplicidade de possibilidades, é impossível deter conhecimento suficiente para orientar adequadamente essa escolha: não há como se saber sozinho, por exemplo, qual é a remuneração de cada uma dessas profissões para construir expectativas que orientem melhor a sua decisão. A solução para isso é acessar análises científicas que, por metodologias de colheita e sistematização de informações, forneçam estimativas quanto ao valor que cada um desses profissionais recebe ao longo da vida. A verdade, aqui concretizada pelos resultados dessas análises, diminui a complexidade porque ao fornecer subsídios para a escolha do indivíduo, reduz as possibilidades às quais ele está exposto, por exemplo, àquelas que igualam o nível de remuneração que o jovem entende adequado para viver. A verdade transmite, nesse sentido, uma seleção prévia – feita de acordo com um procedimento científico –, capaz de amenizar a pressão seletiva e

---

<sup>554</sup> Como fiz quanto à ideia de complexidade, reconstruí o conceito de verdade a partir de *Legitimação pelo Procedimento*. Como aconteceu com os demais conceitos da Teoria dos Sistemas, eles foram trabalhados por Luhmann no decorrer das fases seguintes de desenvolvimento de seu pensamento, culminando, no caso da verdade, no que consta em LUHMANN, N. **Die Wissenschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1992, pp. 167 e ss.

<sup>555</sup> A confiança é, em si, um mecanismo de redução de complexidade à semelhança da verdade, mas que, para o contexto dos procedimentos, tem uma importância menor – razão pela qual optei por dedicar mais atenção à segunda. Uma abordagem completa quanto à confiança como um mecanismo de redução de complexidade pode ser encontrada em LUHMANN, N. **Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität**. Stuttgart: Lucius & Lucius, 2014, pp. 30 e ss.

viabilizar a escolha, mesmo nas condições adversas às quais a sociedade moderna expõe o agente. Fica, em síntese, mais fácil decidir.

Esse papel redutor de complexidade do mecanismo verdade atua também nos procedimentos<sup>556</sup>: eles também se baseiam em afirmações produzidas cientificamente como pontos de partida que transmitem seleções anteriores, reduzindo o número de escolhas possíveis e permitindo, assim, que decisões aconteçam. Procedimentos eleitorais, por exemplo, são organizados a partir de análises estatísticas que preveem o percentual de comparecimento das pessoas, balizando os recursos necessários para que sejam realizados com sucesso; procedimentos legislativos têm os seus resultados restritos àquilo que é cientificamente possível fazer – ninguém cogita editar uma lei que obrigue alguém a estar em dois lugares ao mesmo tempo; procedimentos jurisdicionais geram decisões coerentes com fatos científicos, como, por exemplo, um exame genético que demonstre que o autor é filho do réu. A verdade científica, portanto, não é o *resultado* dos procedimentos políticos, mas é algo que *permite* a esses procedimentos decidir algo, assim como *permite* ao jovem chegar a uma conclusão quanto ao curso que irá se candidatar.

A partir dessa relação entre procedimento e verdade, Luhmann avança para o segundo passo, tentando identificar, no âmbito do político, mecanismos semelhantes aos da verdade. O raciocínio por trás desse movimento é o seguinte: se, na ciência, ao procedimento científico corresponde a verdade – um mecanismo redutor de complexidade –, então, no âmbito do político, aos procedimentos eleitorais, legislativos e jurisdicionais, é possível que também corresponda um mecanismo funcionalmente equivalente, que também atue como um redutor de complexidade.

Luhmann aventa, como hipótese, que esse mecanismo é o poder<sup>557</sup>. Para ele, “*quem tem o poder pode motivar outros a adotar suas decisões como premissas de procedimento, [...] a aceitar como compulsiva uma seleção dentro do âmbito de possíveis*

---

<sup>556</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 24.

<sup>557</sup> Tanto poder quanto verdade seriam incorporados, nos textos seguintes de Luhmann, à ideia comum de meios de comunicação simbolicamente generalizados, definidos como mecanismos sociais que maximizam a possibilidade de êxito da comunicação, porque tornam provável a aceitação de alguém quanto a uma comunicação que lhe é direcionada. *Cf.*, quanto ao tema em geral, LUHMANN, N. Einführende Bemerkungen zu einer Theorie symbolisch generalisierter Kommunikationsmedien. **Zeitschrift für Soziologie**, v. 3, n. 3, pp. 236/255, 1974; e, em específico, respectivamente, quanto ao poder e à verdade, LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 328 e ss e 339 e ss; e em particular quanto ao primeiro, *cf.* LUHMANN, N. **Macht**. Stuttgart: Lucius & Lucius, 2003, pp. 4 e ss.

*alternativas de comportamento*”<sup>558</sup>. Assim, pelo poder, o pai diz ao filho que lugares ele poderá ir; o guarda sinaliza ao carro o caminho que ele deverá percorrer; e o juiz diz ao réu quais dias ele deverá comparecer no fórum. Exercer o poder, nesse sentido, é restringir possibilidades: quem antes tinha como viável adotar um amplo conjunto de ações – algo típico da complexidade da sociedade moderna –, vê-se na contingência de fazer uma escolha entre um campo menor de possibilidades – ou mesmo não fazer escolha nenhuma –, resultado de uma seleção prévia de um terceiro que lhe é imposta. O poder atua, assim, tal como a verdade: reduzindo a complexidade, por restringir as possibilidades daquilo que é viável.

Para entender melhor como verdade e poder atuam de forma paralela, mas em âmbitos distintos, é útil voltar ao exemplo do jovem que deve escolher uma carreira. A complexidade do momento, concretizada na multiplicidade de cursos universitários elegíveis, pode ser reduzida pelo mecanismo verdade, que lhe transporta seleções feitas anteriormente, esclarecendo-lhe, por exemplo, quais cursos oferecem as melhores remunerações. Existe, entretanto, outra forma pela qual o escopo de possibilidades pode ser reduzido: é possível que seus pais exerçam o poder que têm, e digam que só admitem que o filho faça Direito, Medicina e Engenharia. Como acontece no caso da verdade científica – que ajudava o jovem a se decidir separando os cursos de acordo com o salário –, o poder tem a capacidade de reduzir a complexidade, porque ele transmite uma seleção prévia, feita por um terceiro, que estreita o campo de possibilidades disponíveis para a escolha e, assim, reduz a pressão seletiva: o que antes era um campo imensurável de possibilidades, que, iam do Direito às Artes Cênicas, agora se reduz a três cursos de ação.

Esse paralelismo entre verdade e poder, identificados como mecanismos redutores de complexidade, permite avançar ao terceiro passo do argumento: entender a relação que existe entre o poder e os procedimentos da estrutura política da sociedade moderna. O raciocínio por trás desse movimento é o seguinte: se os procedimentos científicos viabilizam a verdade, qual é a relação que existe entre os procedimentos políticos e o poder?

A hipótese de Luhmann a esse respeito começa com a constatação de que verdade e poder são mecanismos com funções equivalentes, mas que têm diferenças entre

---

<sup>558</sup> “*Wer Macht besitzt, kann andere motivieren, seine Entscheidungen als Verhaltensprämissen zu übernehmen, also eine Selektion aus einem Bereich möglicher Verhaltensalternativen als bindend zu akzeptieren*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 25; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 26.

si. A verdade, que é o resultado do procedimento científico, é aceita como uma razão que orienta as pessoas porque ela expõe como o mundo é. O jovem aceita a afirmação de que médicos ganham mais do que artistas porque reconhece que isso reflete uma situação que existe e que, a princípio, não faz sentido questionar – exceto sobre as mesmas bases que levaram à construção dessa verdade, argumentando, por exemplo, que os dados sistematizados não são suficientes para sustentar essa conclusão.

Com o poder, entretanto, é diferente: ele manifesta uma seleção de um terceiro, que decide que algo deve ser de uma determinada forma, e não de outra – mesmo que ambas sejam possíveis. No exemplo da escolha universitária, os pais *decidem* que o filho irá seguir pelas carreiras mais tradicionais, mas eles também poderiam ter decidido obriga-lo a fazer outros cursos, que não Medicina, Direito e Engenharia. O poder, como um mecanismo redutor de complexidade, *exige um ônus diferente da verdade para funcionar*: deve haver *razões*, para além da decisão em si, que o sustentem nessa posição. Faz sentido perguntar: *por que eu devo obedecer? Por que eu devo aceitar a seleção de um terceiro como uma premissa para a minha decisão?*

Uma vez que tipo de pergunta é levado das relações particulares para a estrutura política estatal da sociedade moderna – isto é, vai do nível “*por que eu tenho que obedecer*” para “*por que as pessoas, em geral, obedecem?*” – ela ganha uma dimensão diferente de dificuldade. É fácil ver por que o jovem pode acabar obedecendo os pais: é possível que ele os respeite, entenda que querem o melhor para ele, ou mesmo tenha sido avisado de que ficará sem o suporte financeiro necessário para iniciar o curso. Mas o que produz “*o fenômeno espantoso e surpreendente de uma aceitação geral das decisões do governo*”<sup>559</sup>? Por que as seleções realizadas pelo poder político são tomadas como premissas para as escolhas das pessoas? Por que, afinal de contas, alguém obedece ao direito?

O primeiro passo para compreender a resposta de Luhmann é entender como esse tipo de problema foi tratado. Por trás das perguntas do parágrafo anterior está a questão de se definir *em que sentido o direito de um Estado pode ser considerado legítimo*. Legitimidade, no sentido em que aplicado aqui, é a “*disposição generalizada para aceitar*

---

<sup>559</sup> “[...] *das erstaunliche Phänomen eines durchgängigen Akzeptierens staatlicher Entscheidungen herbeiführen und garantieren*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 30

*decisões de conteúdo ainda não definido, dentro de certos limites de tolerância*”<sup>560</sup>. Uma estrutura legítima é aquela que “*é capaz de produzir uma prontidão generalizada para aceitação de suas decisões*” e o direito legítimo, nessa direção, é aquele que as pessoas estão dispostas a aceitar, isto é, cujas seleções – concretizadas, por exemplo, em regras jurídicas que excluem certos comportamentos do campo do lícito – são adotadas por elas como pressupostos para as suas próprias escolhas<sup>561</sup>. A questão, portanto, resume-se a entender o que faz dessa estrutura chamada direito, na sociedade moderna, algo legítimo<sup>562</sup>.

Há duas respostas iniciais para essa pergunta: a coação e o consenso. Pela primeira, as pessoas aceitam tomar as decisões feitas pelo Estado como premissas para as suas próprias escolhas porque, caso contrário, seriam forçadas de qualquer maneira a fazer isso. Legítimo é aquilo que tem meios para fazer valer as suas decisões, mesmo que contra a vontade de alguém: quem não paga as dívidas, tem a propriedade confiscada e vendida para saldá-las; alguém que invade a propriedade alheia pode ser impelido para fora dela pela polícia; aquele que se recusa a comparecer no juízo será conduzido sob vara.

Pela segunda resposta, as pessoas agiriam dessa forma porque concordam com a decisão tomada pelo Estado e, por isso, aceitam-na. As regras que lhes são impostas, nesse sentido, seriam extensões de um consenso implícito, seja quanto às próprias regras, sejam quanto a valores compartilhados socialmente que lhes dão fundamento: todos concordamos com a existência do Estado, o que implica na necessidade de que impostos sejam pagos. Por isso, as pessoas aceitam – e, portanto, tomam como uma baliza para as suas decisões – a regra de que uma parte da renda mensal deve ser colocada à disposição do Estado. Sob essa perspectiva, legítimo é aquilo que incorpora, na seleção, um valor aceito pelas pessoas como verdade.

---

<sup>560</sup> “[...] *eine generalisierte Bereitschaft, inhaltlich noch unbestimmte Entscheidungen innerhalb gewisser Toleranzgrenzen hinzunehmen*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 30.

<sup>561</sup> LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 3. As citações das páginas 1 a 5 referem-se a uma apresentação feita por Tércio Sampaio Ferraz Jr. à versão em português. Cf., no mesmo sentido, VILLAS BÔAS FILHO, O. V. **Teoria dos Sistemas e o Direito Brasileiro**. São Paulo: Saraiva, 2009, pp. 173/176.

<sup>562</sup> LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 3



Segundo Luhmann, ambas as explicações são frágeis. A primeira, porque “nenhum sistema político pode se apoiar apenas sobre uma força física de coação”<sup>563</sup>. Não é viável que o Estado aja posicionando mecanismos de controle a cada momento da vida das pessoas; na realidade, há inúmeras oportunidades em que elas simplesmente poderiam agir de uma forma diferente daquela que lhes é imposta, sem que isso resultasse em alguma forma de coação, mas não o fazem. A prontidão generalizada à qual refere Luhmann, nesse sentido, acontece mesmo sem uma vigilância generalizada, algo que fica particularmente claro na sociedade moderna, em que há um aumento do número de cursos de ação possível – consequência do aumento da complexidade – sem que isso se traduza, necessariamente, na falta de controle dos Estados quanto àquilo que as pessoas fazem.

A segunda, porque faz pouco sentido pressupor a existência de um consenso geral que permeia toda a atividade política e fundamenta os atos de poder do Estado. A ideia de que, em comunidades políticas, as pessoas concordem quanto a todas as regras que lhe são impostas é desconectada com a realidade de divergências que caracteriza a sociedade moderna e que aparece, inclusive, nos próprios procedimentos, em que discordâncias quanto a escolhas eleitorais, legislativas e procedimentais vem à tona. Nesse sentido, a complexidade que caracteriza a sociedade moderna como um campo de multiplicidade de possibilidades de ação do indivíduo, também a marca com uma multiplicidade de valores inconciliáveis, que tornam o consenso altamente improvável<sup>564</sup>. Como afirma Luhmann, “consenso e coação [...] constituem ‘recursos escassos’ do sistema político”<sup>565</sup> e, por isso, não se prestam a justificar a aceitabilidade generalizada das decisões políticas.

Se nem coação nem consenso explicam a legitimidade do direito, então o que pode fazê-lo? A hipótese de Luhmann é que essa é a função do procedimento: *procedimentos conferem legitimidade ao direito*. Eles criam uma “disposição generalizada

---

<sup>563</sup> “[...] daß kein politisches System sich allein auf physische Zwangsgewalt stützen könne [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 27/28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 29.

<sup>564</sup> Quanto a esse ponto, cf. a explicação quanto a essa impossibilidade feita por Luhmann, ao esclarecer que é um erro pensar que mesmo a própria comunicação seria baseada em um consenso, em LUHMANN, N. **Soziologische Aufklärung 6: Die Soziologie und der Mensch**. Opladen: Westdeutscher Verlag, 1995, pp. 119 e ss.

<sup>565</sup> “Konsens und Zwang sind beides ‘knappe Ressourcen’ des politischen Systems”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 30.

*para aceitar decisões de conteúdo ainda não definido, dentro de certos limites de tolerância*<sup>566</sup>, que substitui a coação e o consenso, viabilizando que as pessoas *aceitem* o poder na sociedade moderna, mesmo que não estejam *coagidas* a fazê-lo a todo momento, e mesmo que não *concordem* sempre com ele.

Mas como eles fazem isso? Essa é uma pergunta importante: assim como a diferenciação de papéis e a adoção de uma metodologia científica própria – que traduz questões do mundo em problemas de pesquisa, aventa hipóteses e as testa – gera, no âmbito da ciência, a verdade, não basta dizer que o procedimento ocasiona legitimidade – é preciso, antes, expor como isso é feito. Na seção seguinte, reconstruo a forma como Luhmann particulariza isso para cada um dos três procedimentos que ele menciona acima<sup>567</sup>, utilizando isso para demonstrar sua hipótese.

Luhmann argumenta, sob uma perspectiva geral, que procedimentos geram legitimidade porque eles *trazem aqueles que serão impactados por uma decisão a uma posição em que eles aprendem a aceita-la*. Procedimentos, nesse sentido, são técnicas que adstringem as divergências das pessoas a contextos específicos, em que elas são controladas, passo a passo, *“criando-se condições para a aceitação de uma decisão final”*<sup>568</sup>. Eles limitam o conflito, fazendo-o caber nesse ambiente institucional em que há regras e papéis que o conduzem do início até uma decisão final, que inevitavelmente irá decepcionar uma das partes – mas que, por também tê-la construído, estará preparada para recebe-la, mesmo que contrária aos seus interesses. O procedimento, nesse sentido, não cria consensos – as partes que iniciaram uma ação judicial discordando irão termina-la com a mesma posição –, mas limita divergências, e, mais importante, imuniza a decisão final que toma parte no conflito, porque *ela se torna esperada*. A legitimidade que se dizia ser o resultado de um compartilhamento de valores, capazes de construir um consenso quanto ao exercício do poder e, por isso, torna-lo aceitável, passa a ser a legitimidade produzida por

---

<sup>566</sup> “[...] *eine generalisierte Bereitschaft, inhaltlich noch unbestimmte Entscheidungen innerhalb gewisser Toleranzgrenzen hinzunehmen*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 30.

<sup>567</sup> Luhmann dedica um quarto e breve capítulo ao procedimento decisório em processos administrativos, que é explicado a partir de uma remissão aos três paradigmas que menciono neste capítulo – judicial, eleitoral e legislativo. Considerando essa posição à margem dos demais e que não há conclusões que auxiliem no objetivo deste trabalho, deixei, por razões de brevidade, de reconstruir esse capítulo. Cf. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 201/218.

<sup>568</sup> LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 4.

procedimentos que reconhecem o pluralismo de valores que torna impossível um compartilhamento consensual, mas, antes, normalizam as decepções que são inevitáveis, impedindo que elas se generalizem pelo tecido social.

Não é, entretanto, qualquer tipo de procedimento que é capaz de fazer isso: esse é um termo utilizado por Luhmann em um sentido específico, que, para designar o instrumento de legitimação do direito na estrutura política da sociedade moderna, diferencia-se da forma como é utilizado corriqueiramente. Entender isso é importante, porque as características que Luhmann agrega a essa ideia de procedimento são aquelas que particularizam esse modelo estatal de legitimidade e que, portanto, precisam estar presentes em outros contextos – como os das plataformas – para que se possa falar, ali, também de estruturas capazes de criar legitimidade. Além disso, entender a especificidade do procedimento ajuda a entender de que forma ele gera, na estrutura política, legitimidade.

A perspectiva tradicional define procedimentos como rituais construídos por uma sequência fixa de ações, situações ou relações, das quais decorre um resultado previamente fixado. A decisão, nesse sentido, não é construída pelo ritual, mas por mecanismos externos e apenas formalizada por um encadear de momentos que não propriamente a legitima, mas apenas chancela uma escolha prévia que independe de como se comportam os atores do processo. Um exemplo disso são as eleições feitas para cargos em que há um acordo prévio de respeito à ordem de antiguidade, como é comum no Poder Judiciário<sup>569</sup>: o que há, nesses casos, não é uma escolha legitimada por um procedimento, mas a chancela de um critério externo, para onde as justificativas dessa forma de agir são transferidas. Assim, no exemplo, o respeito pela antiguidade pode ser fundamentado como uma forma de se evitarem conflitos que enfraqueçam institucionalmente os tribunais. De acordo com Luhmann, *“Uma tal opinião conceberia o procedimento como um ritual em*

---

<sup>569</sup> Essa é uma prática comum no Poder Judiciário; no Supremo Tribunal Federal, por exemplo, a Presidência da Corte – e, conseqüentemente, de todo o Poder – é ocupada pelo Ministro mais antigo que ainda não a exerceu. Por isso, mesmo em comunicações oficiais, essa escolha é denominada uma *“eleição simbólica”*; cf., por exemplo, STF. Jobim se despede da 2ª Turma para assumir presidência do STF. **Notícias STF**, 4.5.2004, disponível em <[http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?id\\_Conteudo=62661](http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?id_Conteudo=62661)>, acesso em 12.8.2019.

que uma única ação estaria certa em cada caso e as ações estariam de tal forma encadeadas que, excluindo a possibilidade de escolha, uma dependeria da outra”<sup>570</sup>.

Luhmann opõe a esses rituais os procedimentos, que se diferenciam por características que os tornam capazes de *construir e legitimar* uma decisão<sup>571</sup>. O cerne dessa distinção é a diferença entre a *certeza* do ritual e a *incerteza* do procedimento. Rituais são formas de chancela de decisões pré-determinadas, que já estão definidas antes de eles acontecerem. Eles lidam com a certeza, porque a incerteza já foi objeto de uma decisão que lhes é anterior. Procedimentos, de outra forma, só são instrumentos de legitimidade, isto é, de viabilização da aceitação do poder, se têm um resultado necessariamente incerto. Eles precisam estar abertos, no início, a uma *variabilidade de alternativas* e, no final, a uma *decisão não pré-determinável*, porque só assim os seus atores terão incentivos suficientes para tomar parte deles, construindo, efetivamente, o seu resultado final – e, com isso, *aceitando-o* de modo a lhe conferir legitimidade. Procedimentos, por isso, podem ser definidos como *mecanismos que administram, por meio de estruturas, a incerteza da decisão de uma forma regrada*<sup>572</sup>.

---

<sup>570</sup> “Eine solche Auffassung würde das Verfahren als Ritual begreifen, bei dem jeweils nur eine Handlung richtig ist und solche Handlungen so zu einer Kette zusammengeschlossen sind, daß die eine ohne Wahlmöglichkeit sich aus der anderen ergibt”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 38; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 37.

<sup>571</sup> Por oposição não quero dizer que procedimento e ritual são categorias estanques. Nesse sentido, aplicações recentes desse trecho do argumento de Luhmann têm sugerido que ambos são polos entre os quais há categorias intermediárias – isto é, procedimentos com aspectos “*ritualizados*”, o que explicaria problemas como a falta de confiança em processos judiciais em países da modernidade periférica. Cf., nesse sentido, DEWEY, M. Wenn die Entscheidungsakzeptanz scheitert in BIRLE, P., DEWEY, M., MASCAREÑO, A. (org.). **Durch Luhmanns Brille: Herausforderungen an Politik und Recht in Lateinamerika und in der Weltgesellschaft**. Wiesbaden: Springer, 2012, pp. 64/65.

<sup>572</sup> Construo essa definição a partir de um trabalho do Luhmann das décadas de 70 e 80 – porque ele é dedicado a isso –, mas é importante mencionar que essa mesma visão sobre a função do procedimento perpassou trabalhos posteriores em diferentes contextos argumentativos. Em um texto publicado originalmente em 1993 e traduzido em 1995, por exemplo, Luhmann utiliza esse mesmo conceito como base para criticar a Teoria da Ação Comunicativa de Habermas: “A incerteza do futuro é a única variável real da teoria do discurso. Todas as medidas procedimentais servem para oferecer suporte a essas premissas da mesma forma com a que o fazem em processos judiciais ou procedimentos eleitorais de democracias políticas. A decisão precisa ser considerada como aberta até o momento em que ela é feita” (“The uncertainty of the future is the only real invariable of the discourse theory. All procedural measures serve to support these premises much in the same way as they do in court processes or election proceedings of

Uma forma de ver isso com mais clareza é colocando o procedimento em uma perspectiva temporal: ele começa por uma escolha da parte – que, no processo judicial, seleciona um conflito para submeter ao judiciário; no legislativo, elege um problema para ser resolvido pelo parlamento – e é impulsionado por conjunto encadeado de decisões parciais, tomadas por aqueles que nele atuam, alternando posições que o conduzem, passo a passo, para um desfecho. O réu, por exemplo, contesta a petição inicial e decide pedir provas; o parlamentar aquiesce com a proposta, mas decide, antes, pedir uma audiência pública. As partes do procedimento, nesse sentido, atuam alternadamente para influenciar a condução do processo até leva-lo a um termo incerto, mas por elas construído, passo a passo. Cada etapa superada, constituída por uma decisão parcial de uma dessas partes, é um fato que é tomado como pressuposto para a decisão seguinte, construindo uma “*história do processo jurídico*”<sup>573</sup> que molda as expectativas das partes, que se preparam para a decisão final, com a qual concorreram. Isso move as pessoas para uma posição em que fica muito difícil não aceitar o resultado do procedimento, mesmo que ele lhe seja desfavorável: quem decide se opor à decisão judicial tomada em um processo do qual participou carrega o ônus de se contrapor não apenas à decisão, em si, mas a toda a estrutura que a produziu. Não cumprir uma decisão judicial não é uma ofensa à outra parte, mas sim ao sistema judiciário como um todo.

Isso, entretanto, só é possível na medida em que o resultado do procedimento é *incerto*. É essa incerteza que incentiva as pessoas a submeterem o conflito a essa forma procedimental de confina-lo a um âmbito social determinado; é essa incerteza que faz com que as pessoas participem do procedimento, construindo decisões parciais e, assim, tenham contato com a possibilidade da decepção. Por isso, a incerteza marca uma distinção importante entre ritual e procedimento: o segundo só é possível como uma forma de legitimar decisões na medida em que as pessoas decidem participar dele, e as pessoas só participarão efetivamente dele na medida em que não sabem qual será o seu resultado. Não há incentivos para ajuizar uma ação em um tribunal comprado, da mesma forma que não faz sentido concorrer a uma eleição em que o partido único já definiu, de antemão, o vencedor; nem iniciar um processo legislativo de uma proposta que se sabe destinada à

---

*political democracies. The decision has to be regarded as open until it has been made*”). LUHMANN, N. Quod Omnes Tangit: Remarks on Jurgen Habermas’s Legal Theory. **Cardozo Law Review**, v. 17, 1995, p. 886.

<sup>573</sup> “[...] *Verfahrensgeschichte* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 43; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 41.

derrota. É a incerteza, em síntese, que constitui o procedimento como um instrumento de legitimidade:

“[A incerteza] *Dá aos participantes o incentivo de contribuir para o progresso do procedimento por meio das suas próprias tentativas de redução, mantém-lhes vivas as esperanças e conduzi-los através do caminho que, de acordo com as regras do processo jurídico, levará à decisão. [...] Se não existir essa incerteza, então não ocorre um processo jurídico singular, como por exemplo em eleições políticas com listas únicas incontestáveis ou em ‘processos de exibição’; mas antes uma declaração ritualística de valores de um outro sistema, que deve então ser também motivado exteriormente*”<sup>574</sup>.

A incerteza, entretanto, por si só, não é suficiente para caracterizar o procedimento como um instrumento capaz de produzir decisões legítimas. Antes de ser um mecanismo que *confere legitimidade* às decisões, o procedimento precisa ser um instrumento que *produz* decisões. Ele precisa, portanto, ao mesmo tempo em que viabilizar a *incerteza quanto ao conteúdo da decisão*, viabilizar que *uma decisão* será produzida. Isso é feito pelo segundo conjunto de elementos que, acoplados à incerteza, constituem o cerne do procedimento ao qual Luhmann faz referência: as estruturas.

Estruturas são limitações internas, construídas dentro do procedimento, que organizam a forma como ele acontece. Elas se manifestam na forma de “*normas jurídicas gerais, válidas para diversos procedimentos*”, que “*reduzem a tal ponto o número ilimitado de formas possíveis de comportamento, que se torna possível, sem incômodas discussões prévias sobre o sentido e finalidade [...] definir [quanto aos procedimentos] a sua temática e os seus limites e tornar os participantes conscientes disso*”<sup>575</sup>. Elas

---

<sup>574</sup> “*Sie gibt den Beteiligten den Anreiz, mit eigenen Reduktionsversuchen zum Fortgang des Verfahrens beizutragen, hält ihre Hoffnung wach und lotst sie zugleich auf den Weg, der nach den Regeln des Verfahrens zur Entscheidung führt. [...] Fehlt es an solcher Ungewißheit, wie zum Beispiel bei politischen Wahlen mit feststehender Einheitsliste oder bei »Schauprozessen«, dann liegt kein eigenständiges Verfahren vor, sondern eine ritualistische Darstellung der Werte eines anderen Systems, die dann auch extern motiviert werden muß*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 51/52; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 46.

<sup>575</sup> “[...] *allgemeine, für viele Verfahren geltende Rechtsnormen vorgezeichnet [...] Sie reduzieren jedoch die unendliche Vielzahl möglicher Verhaltensweisen so weit, daß es möglich wird, ohne umständliche Vorverhandlungen über Sinn und Zweck*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 42; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 40.

conformam o procedimento a um conjunto de caminhos possíveis, limitados pela necessidade de que um estágio inicial, de *variabilidade de alternativas*, seja conduzido a um estágio final, de *decisão*.

Concretamente, essas estruturas se manifestam, por exemplo, em regras quanto a quem pode iniciar o procedimento, quem pode dele participar, qual é o caminho que ele irá percorrer e em que medida algo que lhe é externo pode tomar-lhe parte. Também fazem parte da estrutura disposições que estabelecem papéis, isto é, funções específicas a determinados atores, que limitam o seu campo de ação como agentes do processo, constringendo-os a agir segundo expectativas pré-estabelecidas. Isso é particularmente claro no processo judicial: um indivíduo pode atuar em diferentes papéis sociais – pai, proprietário, empregado, etc. –, mas, uma vez que se insere no processo judicial, ele tem o escopo de suas ações balizado pela função que exerce ali – de autor, réu, ou mesmo juiz. As estruturas desse mecanismo decisório constringem a sua atuação a um determinado nicho, de modo que ele não pode, por exemplo, trazer as características de seu papel de proprietário para o processo e comprar a decisão: ele é obrigado a influenciá-la por aquilo que a estrutura lhe coloca à disposição – manifestações escritas, pedidos de provas e argumentos orais, por exemplo – e seguir as regras desse jogo processual se quiser vencê-lo.

Um exemplo pode ajudar a auxiliar a entender esse ponto. Suponha um processo judicial em que uma pessoa afirma que é herdeira de outra e que, por isso, tem o direito a receber uma parte de sua herança. Esse é um procedimento que, como um instrumento de legitimidade, só faz sentido se nenhuma das partes sabe quem será o vencedor: caso o autor obtenha a informação de que o juiz do caso foi comprado pela outra parte, ele não terá incentivos para contribuir com o prosseguimento do processo – e, uma vez não engajado, não aceitará a decisão no final. Assegurada a incerteza – ou seja, o desconhecimento do conteúdo da decisão –, é preciso que haja uma decisão. Por isso, o processo avança segundo um conjunto de regras, que define, por exemplo, que o réu deve ser o espólio ou os demais herdeiros do falecido; que, após o suposto herdeiro apresentar seus argumentos, deve ser oportunizado aos demais também falar; que as partes devem ter a possibilidade de pedir que provas sejam realizadas por um perito neutro, como a que reconhece a filiação por um exame de DNA; e que, depois de algum tempo, uma decisão necessariamente deve ser proferida pelo juiz. Além disso, há regras que constituem e delimitam os papéis dos atores processuais: ao autor cabe provar o que argumenta, utilizando argumentos jurídicos; aos réus, defenderem-se; e, ao juiz, decidir de acordo com

o direito. Sob uma perspectiva temporal, jungidas a *incerteza de que uma decisão será proferida* e as *estruturas que garantem que alguma decisão será proferida*, o procedimento caminha garantindo às suas partes a possibilidade de definir o seu destino, fazendo-as participar da construção de uma decisão final que, necessariamente, representará uma decepção para uma delas, algo que, entretanto, é esperado: faz parte do jogo processual ganhar e perder.

Dessa forma, o processo caminha de um momento inicial, de alta complexidade – isto é, de grande variabilidade de opções – para outro de menor complexidade – que culmina com a decisão –, produzindo um resultado transmissível às pessoas como um ponto de partida redutor de complexidade, assim como acontecia com a verdade. Diferente dela, entretanto, aqui são necessárias razões que levem às pessoas a aceitarem essas decisões políticas. Por isso o procedimento é tão importante na estrutura política: ele junte incerteza e estrutura em um modelo que permite às pessoas participarem da construção da decisão, ao ponto em que, mesmo que ela lhes seja desfavorável – gerando um candidato indesejado, uma lei ruim ou uma decisão judicial que lhe prejudica – elas se veem na contingência de aceitá-la.

Na seção seguinte, a relação estabelecida, por hipótese, entre o procedimento como uma incerteza regrada e a produção de decisões legítimas é colocada à prova nos três principais mecanismos em que se manifesta na estrutura política da sociedade moderna, segundo Luhmann: nos procedimentos judiciais, nos procedimentos eleitorais e nos procedimentos legislativos. Como em a *Legitimação pelo Procedimento*, separei uma seção à parte para o primeiro (2.3.1) e condensei os demais no item seguinte (2.3.2).

## 2.3 Três casos: a legitimação pelo procedimento na prática

### 2.3.1 Os processos judiciais

Como essa relação entre legitimidade e processo se manifesta no que Luhmann denomina os “*procedimentos que conseguiram uma importância especial*, [porque] *estruturalmente decisivos nos sistemas políticos atuais: o procedimento de*



*eleição política, o procedimento parlamentar da legislação e o processo judicial*”<sup>576</sup>? Em particular: de que modo esse vínculo que ele estabelece, por hipótese, entre ambos, explica melhor as características desses procedimentos que a relação causal entre, de um lado, processo, e, de outro, verdade? Dedico esta seção e a próxima a mostrar como Luhmann enfrenta essas perguntas, sistematizando suas respostas da mesma forma em que aparecem em a *Legitimação pelo Procedimento*: em duas seções, uma dedicada aos processos judiciais, e, outra, aos processos eleitorais e legislativos. Em cada uma delas, traço, como Luhmann fez, paralelos entre características detectáveis nesses procedimentos e o seu papel legitimador, justificando-as funcionalmente.

No caso dos processos judiciais, Luhmann apresenta uma abordagem que os toma sob uma perspectiva geral. Processos variam<sup>577</sup>: eles podem ser estruturados de maneiras diferentes, que dependem do ordenamento jurídico, de como os seus atores os constroem a partir de suas próprias decisões e mesmo da situação psíquica momentânea das partes ou do juiz. Esses fatores não são irrelevantes, mas eles acontecem dentro de uma estrutura geral legitimadora, que é extraída desse mecanismo pelo recurso à ideia de função – isto é, pela detecção do papel que essa forma de organizar o exercício do poder desempenha na estrutura política da sociedade moderna. É nessa estrutura geral, aplicada ao processo judicial, que Luhmann se concentra.

Essa perspectiva se concretiza em características que Luhmann identifica nos processos judiciais, e que são indicativos de sua função de legitimar o poder. No que segue, exponho os principais elementos apontados nesse sentido, para, em seguida, relacioná-las com esse papel atribuído por Luhmann ao procedimento.

Primeiro: processos constroem fronteiras que criam diferenças entre o que acontece dentro e fora deles. Eles são um mundo dentro do mundo: diferenciam-se de tudo aquilo que os circunda, estabelecendo “*para si próprio uma esfera de sentido, de forma a que os processos seletivos de elaboração de informação do meio ambiente possam ser orientados por regras e decisões próprias do sistema*”. Por isso, “*estruturas e acontecimentos do meio ambiente não sejam automaticamente válidos para o sistema*”:

---

<sup>576</sup> “*Wir wählen dafür Verfahren, die im heutigen politischen System besondere, strukturtragende Bedeutung gewonnen haben: das Verfahren der politischen Wahl, das parlamentarische Gesetzgebungsverfahren und den gerichtlichen Prozeß*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 13; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 18.

<sup>577</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 57/58.

elas precisam, antes, ser parte desse mundo processual à parte para que, então, consigam impactá-lo.

Isso é perceptível em duas características comuns aos processos judiciais. Em primeiro lugar, processos constroem papéis específicos, que separam o indivíduo do ator processual. Assim, é comum que códigos processuais estabeleçam direitos e deveres, vinculados à posição ocupada pelo agente no processo<sup>578</sup>, distinguindo a forma como ele deve se portar naquele contexto específico. Um exemplo disso é o juiz: “*ele tem [...] de ser eximido da consideração dos outros papéis que o podem motivar para decisões definidas como membro de uma determinada família ou camada social, de uma igreja ou seita, de um partido, de um clube, de um bairro*”<sup>579</sup>. Há, nesse sentido, uma construção processual da figura do magistrado, à qual atribuídos direitos e deveres específicos, capazes de estabelecer diferenças entre o indivíduo que conduz o processo, e aquele que é, por exemplo, o proprietário de uma casa, o pai de uma família ou o religioso católico – mesmo que todas essas figuras se concentrem em uma só pessoa. Essa separação de papéis é tão importante, que na medida em que o magistrado impacta na condução do processo não como um juiz, mas por um outro papel que ele também ocupe na sociedade – como o de amigo ou inimigo de uma das partes –, ele imediatamente contamina o procedimento: “*Qualquer outra influência [que não a de magistrado] é desacreditada como corrupção*”<sup>580</sup>. Conduzir processos, por isso, exige rituais de neutralidade – como a posição do assento em que se conduz uma audiência, a prática de receber advogados com a porta aberta e a restrição dos contatos profissionais ao fórum – que demonstrem a aderência da atuação do magistrado aos limites que lhe foram atribuídos por seu papel processual.

---

<sup>578</sup> No Brasil, dois exemplos desses mecanismos são o Livro III, do Código de Processo Civil (“*DOS SUJEITOS DO PROCESSO*”) e o Título VIII, do Código de Processo Penal (“*DO JUIZ, DO MINISTÉRIO PÚBLICO, DO ACUSADO E DEFENSOR, DOS ASSISTENTES E AUXILIARES DA JUSTIÇA*”).

<sup>579</sup> “*Er muß [...] von der Rücksicht auf seine eigenen anderen Rollen entlastet werden, die ihn als Angehörigen einer bestimmten Familie oder Gesellschaftsschicht, einer Kirche oder Sekte, einer Partei, eines Klubs, eines Wohnbezirks usw. zu bestimmten Entscheidungen motivieren könnten*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 64; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 57.

<sup>580</sup> “*Jede andere Einwirkung wird als Korruption diskreditiert*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 64; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 57.

Em segundo lugar, processos judiciais estabelecem filtros, que tanto delimitam aquilo que faz ou não parte deles, quanto fixam critérios que organizam a forma como algo que lhes é externo pode ser absorvido. Fatos externos, nesse sentido, não são capazes de impactar diretamente aquilo que acontece na dinâmica processual: é preciso seguir um ritual próprio, pelo qual critérios processuais atuam sobre eles, inserindo-os no procedimento. Um exemplo disso é a dinâmica probatória: um acontecimento não é, por si só, um elemento de convencimento capaz de gerar uma decisão favorável do juiz. Entre a sua existência externa e a sua inserção no procedimento, existe uma reconstrução, ditada pela dinâmica processual, em que essa circunstância que não faz parte do processo só pode tomar parte dele se seguir critérios específicos, definidos procedimentalmente. Assim, salvo em situações excepcionais, esse fato deve ser trazido por uma das partes – o juiz não pode, de ofício, colhê-lo como algo relevante ao processo –, em um momento específico – há intervalos determinados na dinâmica processual em que esse tipo de argumento deve ser mobilizado – e por meios processualmente definidos – caso esse fato seja, por exemplo, uma questão técnica, ela deve ser aferida por um perito escolhido pelo juiz.

Segundo: o processo judicial é autônomo. Isso significa que ele se desenvolve, em certa medida, por critérios definidos por sua estrutura e atores, e não de acordo com uma lógica de algo que lhe é externo. Como afirma Luhmann, autonomia não é solipsismo: *“que se construa um edifício, que se contratem ou se mantenham juízes, que se fixem ou estabeleçam competências, que se ratifiquem ou promulguem leis, que se conheçam ou participem fatos”*, nada disso contradiz a ideia de autonomia, desde que *“a ação dentro do sistema seja, em parte, orientada segundo critérios próprios de seleção”*. Não faz sentido, entretanto, que o juiz receba um ofício de um deputado dizendo-lhe como o processo deve prosseguir de um ponto em diante; nem que ele tenha a sua remuneração condicionada a conferir um resultado jurídico anteriormente determinado a um procedimento que conduz. Autonomia significa espaço para se conduzir de acordo com os seus próprios critérios e, por isso, aquilo que impulsiona um procedimento até a decisão judicial final são fatos necessariamente processuais

Isso significa que, dentro de certos limites, o processo judicial caminha de acordo com decisões que provêm dele mesmo, tomadas de acordo com os critérios por ele estabelecidos e subscritas por atores construídos naquele próprio ambiente. Isso não se

reduz ao papel do juiz: “o réu que resolve não indicar determinadas testemunhas”<sup>581</sup>, por exemplo, também define o caminho que o processo irá percorrer. Com a petição inicial, oportuniza-se a contestação; saneado o processo, há um momento para a produção de provas; com as alegações finais, o processo avança para a sentença. Mudanças externas ao processo, via de regra, não se traduzem em uma forma diferente de conduzi-lo: se o prédio do fórum vier abaixo, é provável que a base material em que registrados os fatos processuais tenha que se reconstituída, mas são apenas esses fatos – e não um fator externo, por mais catastrófico que seja – que podem conduzir o procedimento ao seu final.

Funcionalmente, diferenciação e autonomia atuam como características processuais que garantem o papel legitimador do procedimento, uma vez que viabilizam a incerteza do resultado do procedimento judicial – e, com isso, incentivam a participação das partes na construção da decisão final, que é o que garante que ela seja aceita. Isso porque, por um lado, diferenciação e autonomia fazem com que certezas externas não se transformem em certezas procedimentais: não é porque um prédio caiu, por exemplo, que o seu zelador será responsabilizado judicialmente. A existência de mecanismos processuais que reconstroem esse fato segundo os critérios jurídicos – que exigem, por exemplo, a demonstração de algum tipo de culpa baseada no laudo de um perito –, tornam incerto o resultado do processo, algo que incentiva tanto autor, quanto réu, a interagirem nesse ambiente para influenciar o seu resultado. Por outro, porque essas características constroem um espaço interno de “*manobra para a atuação dos participantes pleno de alternativas*”, motivando-os “*a tomarem, eles próprios, os riscos de sua ação, a cooperarem, sob controle, na absorção da incerteza e dessa forma contraírem, gradualmente, um compromisso*”<sup>582</sup> – capaz de gerar, ao final, a aceitação mesmo do desfavorável.

Terceiro: processos judiciais constroem uma forma de contato entre as partes. Eles compelem aqueles que estão em conflito a conviver, mesmo que

---

<sup>581</sup> “[...] *der Angeklagte, der sich entschließt, bestimmte Zeugen nicht zu benennen* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 70; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 61.

<sup>582</sup> “[...] *die einen Spielraum für alternativenreiches [...] die Beteiligten motiviert werden, selbst Risiken in ihr Verhalten zu übernehmen, an der Absorption der Ungewißheit durch Eliminierung anderer Möglichkeiten unter Aufsicht mitzuarbeiten und sich dadurch Schritt für Schritt selbst zu binden* [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 73/74; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 64.

temporariamente, em um ambiente controlado por estruturas que conformam esse contato a uma aproximação possível, em que as tensões do problema que as levaram a juízo são atenuadas. Como afirma Luhmann, “*no âmbito de um tal sistema de contato, a decisão isolada converte-se em um momento ao lado dos outros*”<sup>583</sup>. Novamente, isso não significa reconciliação: procedimentos e, por consequência, processos judiciais, não são fábricas de consenso, mas instrumentos que enquadram a divergência em uma estrutura que a torna, dentro do possível, aceitável. Por pior que seja o motivo que traga um autor e um réu a um processo judicial, eles são obrigados a se conformarem a um conjunto de comportamentos que viabiliza a convivência, mesmo que momentânea: o autor deve se demonstrar respeitoso com o réu quando o faz perguntas em uma audiência, porque essa é a única forma de traduzir aquele momento em um convencimento positivo do juiz.

Quarto: processos judiciais compelem as pessoas a adotarem os papéis umas das outras. Como esclareci acima, esse tipo de procedimento é construído a partir da constituição de papéis processuais, que são imputados na forma de um conjunto de direitos e deveres às partes. Juízes, autor, réu, peritos, todos atuam de acordo com esse regime jurídico específico, que estabelecem uma diferenciação entre os papéis que ocupam em sua vida pessoal, e aquele que desempenham no processo.

O ponto é que o contato entre as partes proporcionado por esses procedimentos, e que mencionei acima, faz com que elas vivenciem não apenas o próprio papel, mas também aquele que os outros desempenham. Isso faz parte, inclusive, da estratégia processual para obter uma decisão favorável: é comum que uma parte se pergunte como a outra irá reagir a um determinado movimento no processo – o juiz se pergunta se uma decisão sua irá gerar um recurso; o autor questiona se convém incluir um pedido a mais na sua petição e, com isso, oportunizar que o réu traga mais elementos aos autos; e o réu raciocina que tipo de argumento pertine à sua defesa, para que ele não municie o seu adversário processual com mais informações. Os atores do processo, nesse sentido, veem-se uns nos outros, adotando as posições que lhe são alheias como uma estratégia para vencer.

Funcionalmente, tanto o contato entre as partes, quanto a adoção, pelos atores, de papéis desempenhados uns pelos outros, são identificados por Luhmann como

---

<sup>583</sup> “*Im Rahmen eines solchen Kontaktsystems wird die Einzelentscheidung zu einem Moment neben anderen*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 77; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 66.

instrumentos que tornam mais fácil a aceitação da decisão final, mesmo que ela venha a ser desfavorável. Esse ambiente controlado de estruturas, distinto daquele em que emergiu o conflito, viabiliza, mesmo que temporariamente, a convivência regrada das partes, levando-as a vestir os papéis umas das outras, tornando o sucesso ou o fracasso no processo algo palpável por um calibramento mútuo de expectativas: ao entender a posição do meu opositor, é possível mensurar em que medida é possível ou não vencer o caso. Cria-se um distanciamento da própria posição, que se torna uma dentre as várias possíveis, algo que viabiliza uma análise mais fria – e, sobretudo, técnica – quanto às possibilidades que podem advir do processo. Apuram-se, assim, expectativas, que à medida em que o processo se aproxima de seu fim, tornam o desfecho em uma ou outra direção algo esperado, minimizando a decepção.

Quinto: processos exigem que as partes construam representações coerentes de si mesmas, refletidas pelas decisões parciais que tomam e fazem o procedimento avançar em direção à decisão. Essas posturas são, na verdade, compromissos que as partes assumem consigo mesmo, sob pena de serem tachadas de inconsistentes: quem disse, no início do processo, que era hipossuficiente – e, por isso, não poderia arcar com os seus custos – não pode negá-lo à frente, exceto se apresentar uma justificativa – como a de que era desempregado e obteve um emprego. Isso é particularmente verificável porque os processos têm mecanismos de registro daquilo que aconteceu no seu decorrer – eles têm uma história, contada, passo a passo, por cada decisão que tomam e que permanecem registradas em suporte físico ou digital. Como afirma Luhmann, “*os depoimentos registrados documentam acordos bem-sucedidos – se não quanto ao conteúdo, então pelo menos quanto ao fato de um depoimento definido*”<sup>584</sup>. Eles vinculam as partes àquilo que falaram antes, de modo que, ao final, tenha-se uma representação una daquilo que são.

Funcionalmente, esse pacto de coerência consigo mesmo é uma exigência processual que viabiliza que o processo caminhe da incerteza – inicial e necessária – para a certeza da decisão. Na medida em que o processo avança, decisões são tomadas, seja pelas partes, seja pelo juiz, que o impulsionam a um grau cada vez maior de certeza que, construída pelos próprios elementos do procedimento, culmina com uma decisão. Como disse acima, o procedimento obedece a uma incerteza regrada, porque precisa garantir, ao

---

<sup>584</sup> “*Protokollierte Aussagen dokumentieren gelungene Verständigungen – wenn nicht über den Inhalt, so doch über die Tatsache einer bestimmten Aussage [...]*” LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 93; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 81.

mesmo tempo, tanto que uma decisão será proferida, quanto que o seu conteúdo será incerto. Essa exigência de coerência, nessa direção, viabiliza a decisão, porque lhe confere o apoio que é necessário para a sua elaboração: para que um juiz decida a favor de alguém que foi despejado, ele precisa ter certeza que ele não mora mais lá. Se, durante o processo, isso se torna duvidoso, tomar uma decisão a esse respeito se torna também um problema: por isso o processo estimula as partes a construíram uma narrativa coerente sobre si mesmas, garantindo que o processo terá fim.

Sexto: processos judiciais constroem conflitos permitidos. Como a primeira característica expõe, os procedimentos elaboram um mundo de sentido próprio, em que captam parte da realidade externa a partir de filtros criados por suas estruturas. Isso também vale para os conflitos: eles não são absorvidos em toda a sua complexidade pelo processo, mas passam por mecanismos – como, por exemplo, as limitações de competência de cada juízo – que fazem de problemas morais, econômicos, amorosos, etc., problemas jurídicos específicos. O procedimento, nesse sentido, funciona como uma máquina de reconstrução da conflituosidade que lhe é externa, porque capta as características desses problemas e os traduz em linguagem procedimental: o fim de um casal, por exemplo, transforma-se em uma petição inicial de divórcio. Diferente, entretanto, do conflito inicial – que ocorre em um contexto mais amplo e, por isso, tem uma tendência à generalização –, a versão reconstruída desse problema ocorre dentro do procedimento, em um ambiente controlado por estruturas normativas que limitam a sua expansividade. Como afirma Luhmann, “*os conflitos em si tendem para a generalização, para se estenderem a todas as qualidades, situações, relações e meios dos adversários*”, uma tendência que precisa ser “*detida, e, tanto quanto possível desviada para um processo oposto de especificação dos conflitos através de instituições de efeito contrário*”<sup>585</sup> – algo que é feito pelo processo.

Sétimo: ao construir esses conflitos permitidos, o processo não apenas impede que eles se generalizem socialmente, mas também faz deles específicos, traduzindo a questão a uma divergência jurídica. Isso não significa ignorar um contexto de conflituosidade anterior, mas sim que o processo detém de instrumentos técnicos que extraem desse conflito aquilo que importa para a decisão e, com isso, atenuam a

---

<sup>585</sup> “[...] *weil Konflikte an sich zur Generalisierung tendieren, zur Ausdehnung auf alle Eigenschaften, Lagen, Beziehungen und Mittel der Gegner [...] Sie muß daher gestoppt und, soweit möglich, durch entgegenwirkende Institutionen [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 101/102; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 86.

divergência, porque a transformam em um problema jurídico. Como afirma Luhmann, uma vez que um processo judicial é iniciado, percebe-se que *“já não se trata então, como parece, em primeiro lugar, ao queixoso e seus amigos e talvez também à imprensa, de uma usurpação escandalosa dos direitos e liberdades dos cidadãos, de um ato de violência da burocracia contra o qual a coletividade teria de se defender”*, mas apenas, *“de uma questão problemática da lei de expropriação à sentença contrária do tribunal”*<sup>586</sup>. O descontentamento é especificado, traduzido tecnicamente e submetido a uma decisão jurídica, que o resolve de forma cirúrgica. Concretamente, isso é feito por regras que balizam o que e como algo pode ser apresentado ao juízo: as demandas devem ser construídas por objetos delimitados, juridicamente justificados e segundo uma estrutura própria – a exemplo da divisão adotada em petições, de expor os fatos, o direito e os pedidos.

Funcionalmente, essa construção do conflito permitido, associada à tradução jurídica da divergência, viabiliza moldá-lo a um padrão em que ele possa ser resolvido por uma decisão aceitável pelas partes. Como afirma Luhmann, *“a função do procedimento é, portanto, a especificação do descontentamento e o fracionamento e absorção dos protestos”*<sup>587</sup>: ele constrói os conflitos que pode resolver, por mecanismos capazes de decompô-los até que as suas soluções possam se tornar aceitáveis para as partes. A separação de um casal, por exemplo, é um fato do mundo que congrega conflitos que vão da divisão dos bens à guarda e pensão dos filhos, que frequentemente se entrecruzam na carga sentimental do momento: o pai que se vê privado da guarda dos filhos pode tentar retaliar a mãe, tolhendo-a de uma parte justa do patrimônio que construíram juntos. A técnica procedimental de resolver essas questões pelo processo judicial gera, simultaneamente, conflitos permitidos – seccionando a questão inicial em processos judiciais distintos, que isolam os problemas de guarda, pensão e divisão de bens, uns dos

---

<sup>586</sup> *“Es handelt sich dann nicht mehr, wie es dem Kläger und seinen Freunden und vielleicht auch der Presse zunächst schien, um einen empörenden Übergriff in Rechte und Freiheiten des Bürgers, einen Gewaltakt der Bürokratie, gegen den die Gemeinschaft sich zur Wehr setzen müsste; sondern nur noch um eine schwierige Zweifelsfrage aus dem Enteignungsgesetz, zu der entgegengesetzte Gerichtsurteile, aber noch keine höchstrichterlichen Entscheidungen vorliegen”*. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 116; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 97.

<sup>587</sup> *“Funktion des Verfahrens ist mithin die Spezifizierung der Unzufriedenheit und die Zersplitterung und Absorption von Protesten”*. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 116; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, pp 97/98.



outros – e traduções jurídicas desses problemas – o complexo problema sentimental do fim da intimidade se transforma em uma análise técnica jurídica do regime de bens que vigia à época do casamento. Essa junção de decomposição de conflitos e conversão deles em termos jurídicos, viabiliza, nesse sentido, tanto que uma decisão seja atingida, quanto que ela seja aceita pelas partes: o ex-marido pode continuar entendendo que a separação se deu por culpa da mulher, mas ele se vê premido a reconhecer que, sob uma perspectiva jurídica, a comunhão total de bens com a qual se comprometeu dá à ex-esposa o direito à exata metade dos bens do casal.

Oitavo: procedimentos decidem a partir de mecanismos que garantam a representatividade daqueles que não participam dele. Assim, ao mesmo tempo em que o processo tem instrumentos de filtragem em relação àquilo que lhe é externo, construindo um mundo à parte dedicado aos que estão envolvidos em produzir uma decisão legítima para aquele caso específico, ele também dispõe de formas de contato com os indivíduos que não participam dele. Concretamente, uma manifestação disso é a publicidade: “*o decurso do processo deve poder ser presenciado pelos não participantes*”<sup>588</sup>. Mesmo que as decisões de um processo só impactem, concretamente, naquilo que interessa às suas partes – ou seja, mesmo em procedimentos quanto a fatos estritamente particulares, como os relacionados ao cumprimento de um contrato – há publicidade, de modo que terceiros inteiramente desinteressados possam participar, mesmo que como expectadores, daquilo que se decide ali.

Funcionalmente, a representatividade daqueles que não participam do procedimento é importante para a função exercida por ele, porque garante que a decisão final, mesmo que impacte apenas as partes, seja robustecida por um pressuposto de aceitabilidade geral. À primeira vista, o procedimento destina-se a que aqueles que tomam parte nesse mundo à parte sejam levados a aceitar a decisão final, mesmo que ela seja desfavorável a eles. Essa aceitação, entretanto, não é um ato exclusivamente individual de submissão ao decidido, porque o que a robustece é pressupor que, qualquer outra pessoa, na mesma posição, consentiria da exata mesma forma em ser objeto daquela decisão. Esse não é um consenso explícito: ninguém condiciona a execução de uma sentença a uma pesquisa de opinião, que afere se as pessoas concordam ou não com o que consta nela. O

---

<sup>588</sup> “*Der Ablauf des Verfahrens muß für Unbeteiligte miterlebbar sein*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 123; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 105.

mecanismo funciona de forma diferente: com a publicidade do procedimento – isto é, com a possibilidade de que qualquer um tenha a ele acesso –, pressupõe-se um consenso velado, isto é, implícito, quanto àquilo que ocorreu ali, mesmo por parte dos terceiros que não fazem parte do processo judicial. É como se estivesse implícita, sempre, a seguinte afirmação: tudo aconteceu dentro da normalidade, caso contrário alguém teria falado algo – já que todos podem saber o que aconteceu. Com isso, a legitimidade – isto é, a prontidão generalizada ao cumprimento – que marca a decisão proferida em um processo judicial não advém das partes que o construíram, mas de todos aqueles que, tendo oportunidade, nada falaram.

Nono: processos funcionam de acordo com programas condicionais que restringem a crítica a um escopo técnico, minimizando a divergência. A referência a programas condicionais, nesse contexto, designa a circunstância de que as estruturas que conformam os processos judiciais funcionam para decidir de acordo com uma lógica de que *se* algo acontecer, *então* isso resultará uma consequência processual. Processos judiciais, em síntese, são programados para decidir de acordo com normas jurídicas condicionais. *Se* o contrato tiver uma garantia válida, *então* o réu pode ser desprovido de seu patrimônio; *se* o acusado tiver sonogado uma quantia muito grande, *então* a ele será aplicada uma pena maior; *se* o autor for hipossuficiente, *então* ele terá direito ao benefício. Essa programação prévia é importante, porque direciona a argumentação viável no processo a essa lógica condicional: a única forma de obter sucesso é direcioná-la a um dos lados desse programa condicional. O autor, portanto, dedica-se a demonstrar que a garantia era válida, porque testemunhas subscreveram o contrato; que a quantia sonogada não era tão grande; ou que não dispõe de meios suficientes para viver e, por isso, é hipossuficiente. Essa focalização da argumentação jurídica protege a decisão, justificando-a a partir desse critério condicional e a imunizando “1) contra uma crítica à pessoa; 2) contra uma crítica ao procedimento; 3) contra uma crítica da competência (não-jurídica) especializada e principalmente 4) contra uma crítica de repercussões”<sup>589</sup>. Não faz sentido, portanto, recorrer de uma decisão desfavorável dizendo que o juiz é tímido, que o procedimento não foi o melhor possível, que o advogado é incompetente ou que ela terá, como consequência,

---

<sup>589</sup> “(1) gegen eine Kritik der Person, (2) gegen eine Kritik des Verfahrens, (3) gegen eine Kritik des fachlichen (nichtjuristischen) Sachverständes und vor allem (4) gegen eine Kritik der Auswirkungen (im Unterschied zu den rein juristischen Konsequenzen)”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 132; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 111.

deixar a parte mais pobre. Recursos são fundamentados em uma argumentação jurídica que trabalha a programação condicional que justificou o seu resultado, afirmando, por exemplo, que a decisão deixou de considerar uma das garantias do contrato, que o volume de valores considerado grande pelo juiz na decisão está desatualizado, ou que uma parte da renda do autor está comprometida com remédios e, por isso, mesmo ganhando acima da média do país ele é hipossuficiente.

Funcionalmente, a programação condicional gera uma focalização da crítica, que só pode ocorrer dentro dos estritos limites da normatização que ocasionou a decisão. Isso é importante, porque torna a decisão mais sustentável e, por consequência, passiva de ser aceita: estabelece-se um esteio a partir da qual ela pode ser defendida, mesmo contra um conjunto robusto de argumentos capazes de infirmá-la sob outras perspectivas. Um exemplo disso é o argumento finalístico, que se opõe ao condicional, e que pode ser utilizado para atacar uma decisão que, mesmo assim, consegue se manter em pé – e, portanto, aceitável. Assim, determinar que alguém perca a casa porque não honrou uma dívida é uma decisão atacável sob uma perspectiva finalística: sua consequência direta – deixar alguém sem um lugar para morar – não é desejável, e isso pode ser utilizado como uma forma de resistência. A decisão, entretanto, sempre poderá ser defensável sob uma perspectiva técnica-jurídica, como o resultado de uma programação condicional: como a norma afirma que, se alguém não pagar o financiamento, então o bem que o garantiu deve ser tomado para saldá-lo, a decisão que faz cumprir esse programa é, sob essa perspectiva, correta – mesmo que gere consequências indesejáveis.

### 2.3.2 Os processos eleitorais e legislativos

A mesma pergunta feita no caso dos processos judiciais se estende aos outros “*procedimentos que conseguiram uma importância especial, [porque] estruturalmente decisivos nos sistemas políticos atuais*”<sup>590</sup>: como a relação entre legitimidade e processo se manifesta nos procedimentos eleitorais e legislativos? De que forma a dinâmica do que ocorre nesses processos é esclarecida com mais precisão por

---

<sup>590</sup> “*Wir wählen dafür Verfahren, die im heutigen politischen System besondere, strukturtragende Bedeutung gewonnen haben [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 13; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 18.

Luhmann em relação à perspectiva inicial, que via nesses mecanismos uma forma de construir verdades?

Como aconteceu com os processos judiciais, os processos eleitorais e legislativos têm características que, segundo Luhmann, expõem o seu papel na estrutura política da sociedade moderna. No caso dos procedimentos eleitorais, esses elementos se manifestam em três características, que refletem, na escolha de alguém para um cargo político, a função legitimadora do procedimento.

Primeiro: processos eleitorais são “*especificáveis como funcionais e separáveis das outras relações de papéis*”<sup>591</sup>. Essa característica, que é paralela à primeira dos processos judiciais, manifesta-se concretamente no procedimento eleitoral pela forma como ele apreende a vontade das pessoas, utilizando-a para construir uma decisão. Também o procedimento eleitoral, nesse sentido, cria um mundo dentro do mundo, e, também ele diferencia papéis, como uma forma de distinguir a ação que se espera de alguém que atua no interior desse processo. Luhmann aponta três indícios para isso: a “*universalidade do acesso ao papel de eleitor para toda a população*”, a “*igualdade do peso dos votos*” e o “*segredo da votação*”<sup>592</sup>. O procedimento eleitoral, por esses mecanismos, estabelece, primeiro, um ambiente à parte daquele em que as pessoas vivem, criando ali regras de universalidade, igualdade e sigilo que não existem fora dele. Ele despe o indivíduo de seus papéis fora do procedimento, homogeneizando-os em um modelo abstrato de “*eleitor*”: alguém que, rico ou pobre, católico ou protestante, homem ou mulher, terá sua escolha ponderada da mesma forma que os demais. Como afirma Luhmann, “*todas as diferenças podem ou devem ser ignoradas*” e, por isso, “*o acesso à atuação política no papel de eleitor e os seus efeitos situam-se independentemente de outros papéis sociais e o tipo de decisão como eleitor não precisa ser justificado em outros contextos sociais, pois goza da garantia do segredo*”<sup>593</sup>.

---

<sup>591</sup> “*Sie müssen von anderen Rollenzusammenhängen abtrennbar und funktional spezifizierbar sein [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, pp. 158/159; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 133.

<sup>592</sup> “[...] (i) *durch die Allgemeinheit des Zugangs zur Rolle des Wählers für die gesamte Bevölkerung (mit Ausnahme functional begründbarer Einschränkungen für Unmündige, Entmündigte, Verbrecher)*, (2) *durch die Gleichheit des Stimmgewichts und* (3) *durch die Geheimhaltung der Stimmabgabe*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 159; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 134.

<sup>593</sup> “*Alle Unterschiede dürfen bzw. sollen ignoriert werden außer solchen [...] Der Zugang zum politischen Handeln in der Rolle des Wählers und deren Effekt sind dadurch von anderen gesellschaftlichen Rollen*

Funcionalmente, essa primeira característica tem um papel assemelhado às duas primeiras do processo judicial: nos dois casos, a existência de filtros e papéis que criam um processo eleitoral à parte do mundo são importantes para o seu papel legitimador porque garantem a incerteza. Se as pessoas pudessem trazer características suas para a escolha eleitoral, definindo-a a partir delas, seria possível saber antes o seu resultado, esvaziando o procedimento – uma vez que é precisamente essa incerteza que leva as pessoas a tomarem parte dele e, ao final, aceitarem as suas consequências. Se, por exemplo, o papel do eleitor não fosse desvinculado de sua inserção econômica na sociedade, seria possível – ao menos em sociedades desiguais – definir quem seria o vencedor. É justamente nesse ponto que entram em ação os mecanismos do procedimento eleitoral – como a igualdade e o voto secreto – que fazem dele um mundo à parte: eles viabilizam que diferenças econômicas não sejam traduzidas em diferenças eleitorais, como ocorreria, por exemplo, se o voto das pessoas tivesse pesos diferentes de acordo com a sua capacidade econômica.

Segundo: processos eleitorais se desenvolvem para “*formar alternativas e mantê-las em aberto*”. Há, como afirma Luhmann, a “*criação e conservação de uma complexidade indecisa e contraditória*”<sup>594</sup> no interior de um processo de escolha, construindo-se papéis – como o de candidato – que se mantêm vivos na disputa na medida que conseguem se colocar como alternativas uns aos outros. Essa dinâmica aparece com clareza nos debates eleitorais: diante dos eleitores, os candidatos se apresentam como alternativas uns aos outros que, para sobreviver àquele momento, precisam continuamente argumentar que são diferentes entre si. Eles lutam, portanto, para manterem abertas as alternativas que eles mesmo constituem para a decisão política, dizendo, por exemplo, que irão conduzir políticas públicas diferentes daquelas que o candidato à reeleição executou no decorrer do seu mandato. Eles precisam se postar como diferentes – caso contrário, a sua atuação como candidatos não tem razão de ser.

---

*unabhängig gestellt und die Art der Entscheidung als Wähler braucht in anderen gesellschaftlichen Zusammenhängen nicht verantwortet zu werden, da sie den Schutz des Geheimnisses genießt*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 160; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 134.

<sup>594</sup> “*Daher ist es eine zweite wesentliche Funktion des Wahlverfahrens, Alternativen zu formieren und offenzuhalten*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 161; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 135.

Funcionalmente, formar alternativas e mantê-las abertas é a manifestação procedimental da necessidade de manter o resultado do procedimento incerto para que ele retenha o seu papel legitimador. O procedimento eleitoral, nesse sentido, não viabiliza apenas o voto, mas também possibilita a candidatura, como uma forma de construir, dentro de si, as alternativas que garantam a incerteza de seu resultado. Nesse sentido, não basta apenas a existência desses candidatos: eles precisam ser constantemente colocados, pelo decorrer da corrida eleitoral, uns contra os outros, em um ambiente de divergência controlada, e que permita manter viva, até o último momento, a incerteza do resultado. Esse ponto é chave no procedimento legitimador, porque é essa indefinição que compromete as pessoas com o processo eleitoral, e é esse compromisso das pessoas com o processo eleitoral que as leva a aceitar, no final, a decisão, mesmo que ela seja desfavorável. Não há estímulos para que alguém participe, seja como candidato, seja como eleitor, seja, até mesmo, como um cabo eleitoral, de uma eleição vendida: se é possível saber antes quem será eleito, então não há razão para se envolver com um procedimento que apenas chancela essa decisão externa.

Terceiro: processos eleitorais devem ter uma “*norma reguladora de apoio e controle, que permita que estes problemas sejam decididos*”<sup>595</sup>. Como a característica anterior indicou, eleições viabilizam e estimulam a existência e a permanência de alternativas contraditórias entre si como uma forma de se manter a incerteza latente. Isso é feito, entretanto, apenas até o ponto em que não inviabiliza a formação de uma decisão final: há regras que, assim como no processo judicial, estruturam o processo eleitoral, limitando o conflito a certos parâmetros que o direcionem a um desfecho. Assim, por exemplo, há regras que estabelecem as fases desse processo: há períodos delimitados para que as pessoas se alistem como eleitores; candidatem-se; façam propaganda de suas propostas políticas; e, por fim, realizem a escolha pelo escrutínio propriamente dito. Essas regras também garantem a existência de uma decisão ao conformar o papel dos candidatos e dos eleitores a certos padrões que não prejudiquem o prosseguimento do procedimento, impedindo que o conflito chegue a um nível prejudicial à sua continuidade: um candidato não pode, por exemplo, imputar um fato falso ao seu oponente, porque isso pode elevar a tensão a um nível que leve a métodos de decisão – como as vias de fato – que esvaziariam

---

<sup>595</sup> “[...] und sie müssen ein Regulativ für Unterstützung und Kontrolle enthalten, das diese Probleme entscheidbar werden läßt”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 159; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 133.

o procedimento. Há freios que fazem com que esse procedimento de decisão não se convolve em uma batalha campal, porque isso impede a decisão legítima.

Funcionalmente, essa regulação é importante para o papel legitimador do procedimento porque, como aconteceu no processo judicial, ao mesmo tempo em que é preciso haver *incerteza* sobre a decisão, é preciso que haja uma *decisão*. A técnica de legitimação pelo procedimento envolve, por isso – e como afirmei na seção anterior – uma forma de *incerteza regrada*, que se traduz em mecanismos que assegurem que o desfecho do procedimento irá gerar, necessariamente, uma decisão. Isso é necessário porque o ambiente construído pelo procedimento é um reflexo conflituoso de um conflito que lhe é exterior: a escolha de um candidato não significa selecionar uma pessoa, mas sim um conjunto de soluções para problemas políticos, que são “*compactados*” em um candidato – e, por consequência, em uma só escolha. O procedimento eleitoral, nessa direção, atua como um redutor de complexidade, porque consolida esse conjunto de decisões em uma decisão só – o que não significa que ele ocorre de forma pacífica. Como é preciso manter abertas todas as possibilidades, há um clima de divergência mútua entre os candidatos, porque cada um disputa para se apresentar como uma alternativa aos demais. Isso precisa ser controlado por essas estruturas normativas, caso contrário pode ocasionar não uma decisão, mas um ambiente de belicosidade sem fim. É interessante, nesse sentido, que diferente dos processos judiciais – que, via de regra, não têm dia para acabar – o procedimento eleitoral tem marcos temporais mais estritos, o que atenua essa conflituosidade: sabe-se que, após a última semana de outubro, uma escolha sobre o destino político do país terá sido tomada.

Nos procedimentos legislativos, repete-se o quadro geral dos processos judiciais e eleitorais: Há, primeiro, a “*disposição de possibilidades de procedimento através de preceitos jurídicos e organizações*”<sup>596</sup>, de modo que o processo legislativo legítimo se constrói a partir da escolha de uma dentre várias possibilidades, que necessariamente permanecem em aberto no seu decorrer, por exemplo, por regras que permitem aos parlamentares emendarem a proposta inicial, modificando-a; segundo, “*a individualização dos processos como sistemas especiais de duração limitada, que aspiram*

---

<sup>596</sup> “*Die Einrichtung der Möglichkeit von Verfahren durch Rechtsvorschriften und Organisation, [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 180; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 149.

a uma decisão de compromisso”<sup>597</sup>, isto é, o estabelecimento de regras que balizam o procedimento, fazendo-o caminhar na direção de uma conclusão por mecanismos como a delimitação de etapas – proposta legislativa, tramitação por órgãos colegiados e publicação – e a ideia de que existem oportunidades específicas para que modificações sejam feitas – uma vez realizadas as votações, o projeto não pode mais ser alterado; terceiro, a formação de uma “complexidade própria destes sistemas”<sup>598</sup>, o que significa que, assim como nos processos anteriores, forma-se um mundo procedimental à parte do mundo real, em que filtros e papéis atribuídos aos seus autores viabilizam uma autonomia que garante a incerteza do resultado do processo legislativo; e, quarto, a “criação de uma história própria do sistema para a redução dessa complexidade”<sup>599</sup>, algo perceptível na sistemática de registros que permite testes de coerência quanto àquilo que os atores desse procedimento – os parlamentares, por exemplo – disseram desde o início, e faz a escolha legislativa avançar até um desfecho. Como nos casos anteriores, essas quatro características se resumem à construção de uma incerteza regrada, que garante tanto que haverá uma decisão ao final – ou seja, que uma proposta legislativa vai ser aceita ou não – quanto que ela é incerta – que é a chave para que as pessoas tomem parte nesse procedimento e, conseqüentemente, aceitem o seu resultado, tornando-o legítimo.

Luhmann argumenta, adicionalmente, que há especificidades do procedimento legislativo que não são captadas por esse quadro geral. Elas decorrem de situações específicas que são impostas a ele, mas que não estão presentes nos procedimentos judiciais e eleitorais. Em particular, essas estruturas têm relação com a circunstância de que processos legislativos têm que lidar com um número de possibilidades – isto é, uma complexidade – muito maior que a imposta aos modelos eleitorais e judiciais. Isso acontece porque, por um lado, lida-se com uma infinidade de problemas frequentemente não-programados – isto é, que ainda não foram submetidos a uma

---

<sup>597</sup> “[...] die Individualisierung der Verfahren als besondere Systeme von begrenzter Dauer, die auf eine verbindliche Entscheidung zustreben”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 180; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 149.

<sup>598</sup> “[...] eigene Komplexität dieser Systeme [...]”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 180; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 149.

<sup>599</sup> “[...] Aufbau einer eigenen Systemgeschichte zur Reduktion dieser Komplexität”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 180; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 149.



sistemática legislativa –, que, por outro, podem ser traduzidos em uma infinidade de formatos legislativos diferentes. Procedimentos legislativos, assim, são redutores de complexidade mais premidos que os demais, porque lidam com um grau maior de complexidade tanto fora deles, quando dentro deles próprios. Isso lhes confere características específicas, que consolidei em três linhas abaixo.

Primeiro: processos legislativos têm sistemas informais de contato, isto é, mecanismos que viabilizam que elementos que lhe são externos sejam absorvidos com mais facilidade pelo procedimento, fornecendo indicativos seguros que o orientem em uma direção ou em outra. Como disse acima, esse tipo de processo está sujeito a uma imensa variabilidade de situações, porque praticamente qualquer situação do mundo pode se converter em uma questão legislativa: desde a regulação urbana até o modo como crimes de lavagem de dinheiro são investigados, tudo pode ser “*programável*” pela legislação. Por isso, há formas específicas pelas quais essas influências externas são captadas e inseridas no procedimento, filtrando-as para que esse impacto seja atenuado, mas não eliminado. Uma forma de fazer isso é condicionar a iniciativa legislativa à representatividade: o processo legislativo constrói estruturas que viabilizam o seu início apenas quando há uma provocação de alguém como, por exemplo, um parlamentar, que foi eleito para essa função. Isso comprime as demandas pela resolução de problemas a um número limitado de pessoas, que filtram e apresentam propostas legislativas de forma limitada, o que, ao mesmo tempo que mantém contato com a parte externa do procedimento, reduz a sua complexidade, ao selecionar aquilo que será ou não apresentado para deliberação no parlamento. Há, como afirma Luhmann, uma “*dependência de processos cognitivos alheios*”, pela qual o procedimento se confia nas seleções realizadas por seus atores – como, por exemplo, os parlamentares – para funcionar.

Funcionalmente, essa forma de agir do procedimento é importante porque viabiliza a decisão: caso o processo legislativo não tivesse mecanismos de filtragem, pelos quais essa imensa complexidade ao qual submetido conseguisse ser organizada – por exemplo, na figura do representante – ele não conseguiria chegar a uma decisão. A ideia de se confiar em uma informação já comprida, isto é, que já passou por processos cognitivos alheios, funciona como uma pré-redução da complexidade, que viabiliza a redução final em uma decisão, que converte o não-programado em programado. Isso pode ser observado, por exemplo, no caso de uma cidade que se veja diante de problemas causados pelo surgimento de um novo sistema de transporte por aplicativo. No lugar de o procedimento legislativo ser exposto diretamente à toda a complexidade desse tipo de questão, ele a

compartimenta em uma série de processos que, para serem iniciados, devem contar necessariamente com a iniciativa de um representante parlamentar, que, traduz previamente essa complexidade em uma proposta concreta de lei. Dessa forma, as tensões sociais que decorrem dessa nova forma de se deslocar em um município se veem conformadas, por exemplo, em um projeto legislativo iniciado por um parlamentar que propõe regular a identificação que os veículos de transporte por aplicativo devem utilizar, sob pena de serem multados.

Segundo: procedimentos legislativos funcionam a partir de organizações prévias dos posicionamentos de seus participantes, que permitem esquematizar possíveis resultados de propostas de normatização. A forma mais clara de ver isso concretizado é por meio da divisão entre situação e oposição: parlamentos costumam ser separados entre os que apoiam e os que não apoiam os ocupantes atuais do Poder Executivo. Isso confere uma maior previsibilidade ao procedimento legislativo, porque uma vez iniciado um processo legislativo, desenham-se cenários de probabilidade quanto à sua viabilidade, baseados nessa organização prévia.

Funcionalmente, essa pré-organização é importante, porque ao mesmo tempo em que ela mantém o resultado incerto, ela torna mais provável a decisão, na medida em que consolida os posicionamentos em torno de soluções com o grau máximo de consenso. Nesse sentido, ao invés de se ter um parlamento dividido em tantas posições quantos são os parlamentares, essa pré-organização entre situação e oposição as consolida em um número menor e mais viável de ser trabalhado politicamente. Fica mais fácil, por exemplo, encontrar uma solução de compromisso quanto a uma normativa polêmica, porque é mais fácil achar um ponto intermediário entre duas posições do que entre centenas. Novamente, o que se tem, funcionalmente, são mecanismos que garantem, assim como nos processos anteriores, a existência de uma decisão ao final, tendo que lidar, entretanto, no procedimento legislativo, com um grau mais alto de complexidade, porque tanto os problemas quanto as soluções são muito mais numerosos.

Terceiro: processos legislativos são organizados não de forma, via de regra, pública, como ocorre nos procedimentos judiciais, mas têm um certo espaço para alternar momentos de publicidade com outros de caráter reservado. Assim, as votações e os debates são realizados publicamente: há transmissões na televisão dos discursos que são feitos e é possível acessar, na internet, os dados de cada votação, identificando-se qual foi o posicionamento de cada representante quanto à matéria. Há no processo legislativo, entretanto, toda uma dinâmica de negociações que ocorrem fora desses espaços – a

exemplo das que ocorrem nos gabinetes parlamentares – em que os posicionamentos são frequentemente definidos antes da sessão pública: “*assim, se torna possível trabalhar nas sessões privadas com pressões ou oferecimentos de permuta, que não podem ser indicadas publicamente ou, ao inverso, mostrar maior sensatez e boa vontade do que se podem coadunar com a linha declarada do partido, ou com o seu papel como oposição*”<sup>600</sup>.

Funcionalmente, essa alternância entre momentos públicos e momentos privados é uma engrenagem que facilita a decisão, porque torna possível acordos e decisões que seriam inviáveis publicamente. Se, por exemplo, a aprovação de uma determinada legislação se mostra importante, mas ela é contrária àquilo que defendem publicamente determinada ala dos parlamentares, a possibilidade de confeccionar acordos não públicos a respeito dessa aprovação pode viabilizá-la, ao mesmo tempo em que gera comprometimentos, por aqueles que propõe a inovação, quanto a também se flexibilizarem em uma nova oportunidade. Esse acordo que, se levado ao público, poderia soar ao eleitor como uma quebra de sua confiança, ou uma forma de agir incoerentemente, é viabilizado por essa alternância entre público e restrito e viabiliza a tomada de decisão legislativa.

### 3. CONCLUSÃO

Como afirmei no capítulo anterior, a segunda parte da dissertação se propõe a investigar o funcionamento da moderação de conteúdo para entender por que o procedimento utilizado pelas plataformas para definir aquilo que os seus usuários podem dizer tem se transformado, em particular com a implantação de mecanismos que reduzem a arbitrariedade desse modelo decisório e o aproximam do utilizado pelos Estados. Dividi a estratégia para resolver esse problema em dois passos, sustentados na abordagem funcionalista que justifiquei antes<sup>601</sup>: primeiro, entender qual é o papel que esses procedimentos têm na tomada de decisões nos Estados; segundo, utilizar essa resposta para, com as adaptações necessárias, tentar explicar o que acontece nas plataformas. No

---

<sup>600</sup> “*Dadurch ist es möglich, divergierende Reduktionsstile nacheinander zu verwenden, in den nichtöffentlichen Sitzungen mit Pressionen oder Tauschangeboten zu arbeiten, die öffentlich nicht gezeigt werden könnten, oder umgekehrt mehr Vernunft und Entgegenkommen zu zeigen, als mit der öffentlichen Linie der Partei oder mit ihrer Oppositionsrolle vereinbar sind*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 189; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 155.

<sup>601</sup> Cf. *supra*, Cap. II.3, 3.2.

que segue, concluo esse capítulo sintetizando o primeiro passo – que abordei, acima, reconstruindo a *Legitimação pelo Procedimento* – e expondo as limitações que exigem que essa perspectiva inicial seja aperfeiçoada para que tenha o alcance suficiente para explicar o movimento das plataformas.

Qual é a função que os procedimentos têm na tomada de decisões nos Estados? Essa é uma pergunta importante, porque, ao mesmo tempo em que os procedimentos permeiam a estrutura política da sociedade moderna – atuando, principalmente, nos processos eleitorais, judiciais e legislativos, que Luhmann toma como paradigmas para a sua análise –, as respostas disponíveis quanto ao papel que exercem são questionáveis. Isso porque elas os reduzem a mecanismos que produzem respostas certas – “*verdades*”, tais como o melhor candidato, a decisão mais justa e a forma mais apropriada de se responder a um problema não normatizado –, a partir de um vínculo causal frágil: há eleições que escolhem pessoas inapropriadas, decisões judiciais que aplicam mal o direito e leis que não resolvem problemas da sociedade.

A hipótese de Luhmann a esse respeito é a de que procedimentos têm a função de legitimar as decisões tomadas pelo Estado: segundo ele, a estrutura política da sociedade moderna é marcada por uma *legitimação pelo procedimento*. Há dois termos nessa hipótese que precisam ser esclarecidos para que ela seja compreendida corretamente: “*legitimidade*” e “*procedimento*”. Legitimidade designa o mecanismo capaz de criar nas pessoas uma disposição generalizada para aceitar a decisão de alguém. Inicialmente, os modelos de legitimidade a atribuíam à coação – as pessoas obedecem porque, se não o fizessem, seriam compelidas a obedecer de qualquer forma –; e ao consenso – as pessoas obedecem porque compartilham um conjunto de valores que lhe impõem obrigações com as quais concordam. Assim como nas respostas causais à função de procedimento, Luhmann detecta nessas perspectivas fragilidades que as tornam incompatíveis com a estrutura política da sociedade moderna: não há como o Estado posicionar estruturas de controle que coajam as pessoas em todos os contextos de uma sociedade complexa como a moderna; e também não há como se falar na existência de um consenso em uma estrutura social marcada exatamente pelo oposto – uma variabilidade de possibilidades que se reflete em divergências que surgem, inclusive, nos próprios procedimentos.

Procedimento, por sua vez, é um mecanismo que administra, por meio de estruturas, a incerteza da decisão de uma forma regrada. Uma forma de se entender isso é opondo o *procedimento* ao *ritual*: o ritual é uma sequência fixa de ações, que chancelam uma decisão que lhe é externa. Procedimentos, de outra forma, são mecanismos de decisão

que lidam com a incerteza. Eles se constroem, nesse sentido, ao redor de dois polos: de um lado, a *necessidade de produzir decisões*, que gera *estruturas* capazes de garantir que, no final, alguma definição será atingida. De outro, a *necessidade de que essa decisão seja incerta*, o que assegura que essa é uma definição que surge do procedimento, e não de um fator que lhe é externo. Por isso a ideia de uma *incerteza regrada*: procedimentos lidam com o indefinido – porque a decisão que eles produzem não pode ser a eles anterior –, mas fazem isso a partir de estruturas que progressivamente a conformam no sentido da definição – porque eles precisam produzir uma decisão.

Luhmann argumenta, nesse sentido, que é essa incerteza regrada do procedimento que constrói a legitimidade das decisões na estrutura política da sociedade moderna. Isso acontece porque o procedimento, nos termos em que exposto no parágrafo anterior, começa a partir de uma variabilidade de alternativas, que caminha para uma decisão que não é pré-determinável – mas que é o resultado da atuação das partes. Isso as estimula a participar, atuando de forma alternada e por decisões parciais, que fazem o procedimento caminhar adiante até a decisão – que, nesse ponto, deixa de ser uma surpresa, mas se constitui como o resultado daquilo que fizeram as pessoas. A incerteza, nesse sentido, é o instrumento que incentiva as pessoas a contribuir para o procedimento, trazendo-as para esse formato de decidir. Isso é essencial, porque é justamente essa participação induzida pelo incerto que gera a legitimidade: à medida que as pessoas contribuem para a decisão, tomando progressivamente contato com ela e com a outra parte, elas ajustam as suas expectativas, vislumbram perspectivas de vitória e derrota e, no final, isso torna mais fácil aceitar a decisão, mesmo que ela seja desfavorável. O procedimento, nesse sentido, tanto facilita a aceitação da decisão, quanto dificulta a reação de rejeitá-la – porque o insatisfeito não está mais se opondo à outra parte, mas sim a toda a forma estabelecida de se tomarem decisões.

Luhmann expõe como isso acontece no que considera ser os três principais procedimentos da estrutura moderna: os processos judiciais, eleitorais e legislativos. À parte de especificidades – como, por exemplo, um grau maior de complexidade a que sujeitos os procedimentos legislativos –, todos são marcados por um quadro geral procedimental, isto é, mecanismos internos que garantem, por um lado, a incerteza, e, por outro, a existência de uma decisão ao final. Assim, em todos esses processos há a criação de um mundo procedimental à parte, em que diferenciados papéis e mecanismos que filtram as influências externas – garantindo que a decisão final caberá apenas ao processo; a construção de modos de contato entre as partes, que são estimuladas a adotar os papéis

umas das outras e construir representações coerentes de si próprias – viabilizando que, ao final, a decisão seja aceita mesmo pela parte que não foi favorecida; a elaboração de conflitos permitidos, que traduzem os problemas em divergências jurídicas, eleitorais ou legislativas – viabilizando que decisões sejam tomadas mesmo que o problema de fundo tenha outras implicações; e a viabilização de que pessoas que não tomam parte do conflito sejam nele representados – assegurando que a solução contará com o apoio pressuposto de um espectro mais amplo do social.

A pergunta que fica dessa reconstrução é a seguinte: como isso pode ajudar a resolver o problema do capítulo anterior? De que forma o vínculo que Luhmann estabelece entre procedimentos e legitimidade ajuda a entender as razões por trás do movimento das plataformas da internet, de implantar mecanismos que aproximam o procedimento de moderação de conteúdo ao utilizado pelos Estados para tomar decisões?

Uma resposta direta seria estabelecer um paralelo entre a legitimação pelo procedimento de Luhmann e o movimento das plataformas, e concluir que a primeira explica o segundo. Essa, entretanto, é uma postura arriscada: ela confere pouco valor ao contexto em que construída a relação que Luhmann estabelece entre procedimento e legitimidade, e, principalmente, não percebe os pressupostos que essa teoria traz consigo: antes de trazer a legitimidade pelo procedimento de Luhmann às plataformas, é preciso se perguntar em que medida as conclusões tiradas neste capítulo são aplicáveis a elas.

Essa ressalva é importante, porque o argumento de Luhmann é construído para identificar o papel de um elemento característico da estrutura política da sociedade moderna – o procedimento –, o que, no texto, significa investigar uma característica do Estado. O trecho a seguir, do prefácio de *Legitimação pelo Procedimento*, deixa isso particularmente claro:

*“Tal como a categoria do contrato para o âmbito da ‘sociedade’, assim a categoria do procedimento para o âmbito do ‘estado’ parece apresentar aquela fórmula mágica que combina a mais alta medida de segurança e liberdade que se pode praticar concretamente no dia a dia e que transmite, enquanto instituição, todas as resoluções do futuro”<sup>602</sup>.*

---

<sup>602</sup> “Wie im Bereich der »Gesellschaft« die Kategorie des Vertrags, so scheint im Bereich des »Staates« die Kategorie des Verfahrens jene Zauberformel zu bieten, die ein Höchstmaß an Sicherheit und Freiheit kombiniert, die täglich konkret praktikabel ist und doch als Institution alle Bestimmung der Zukunft überläßt”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. VII; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 7.

É interessante perceber que Luhmann admite a existência de um contrato fora do Estado – na “*sociedade*” –, mas o procedimento só existe no âmbito do “*estado*”. Como outros excertos indicam, parece não estar no horizonte de Luhmann a ideia do procedimento como um instrumento jurídico de legitimação política fora de uma estrutura estatal: a perspectiva teórica inicial com a qual os procedimentos são abordados é a de “*pensadores liberais da teoria do Estado*”<sup>603</sup>; “*o procedimento da eleição política, o procedimento parlamentar da legislação e o processo judicial*”<sup>604</sup> referidos por Luhmann são os instrumentos de decisão do Estado; o paradigma da forma como os procedimentos são entendidos, na prática, é a “*cláusula do ‘due process’ da constituição americana*”<sup>605</sup>; e o direito que é objeto da legitimidade pelo procedimento, que confere aceitação às regras legisladas e às decisões proferidas, é aquele produzido pelo Estado<sup>606</sup>.

Perceber isso é importante, porque o problema do capítulo anterior não acontece no Estado. Este não é um trabalho sobre, por exemplo, a crise nos métodos de legitimidade estatal – um tema que admitiria uma abordagem, como a acima, que toma as conclusões de Luhmann como pontos de partida. A moderação de conteúdo, como insisti no primeiro capítulo<sup>607</sup> e no final do segundo<sup>608</sup>, acontece em um espaço de decisão que, criado pela era digital, surge nas empresas responsáveis pela infraestrutura que viabiliza a Web 2.0. Esse é um dado importante, porque ele condiciona de diferentes maneiras a abordagem que pode ser feita a seu respeito: diferente do que acontece nos Estados, por exemplo, a moderação de conteúdo não está vinculada a protocolos de transparência, o que exigiu, neste trabalho, um esforço de esclarecimento e sistematização quanto àquilo que estava por trás de decisões como a que retirou a foto publicada por Egeland. Inversamente,

---

<sup>603</sup> “*Für liberale Staatsdenker [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 11; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 17.

<sup>604</sup> “[...] *das Verfahren der politischen Wahl, das parrlamentarische Gesetzgebungsverfahren und den gerichtlichen Prozeß*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 13; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 18.

<sup>605</sup> “[...] *die ‘due process’-Klausel der amerikanischen [...]*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 19; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 22.

<sup>606</sup> LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 36.

<sup>607</sup> *Cf. supra*, Cap. I.1.

<sup>608</sup> *Cf. supra*, Cap. I.2, 3.2.2.

a conclusão de Luhmann quanto à relação entre legitimidade e procedimento não pode simplesmente ser aplicada à moderação de conteúdo, porque essa forma de decidir é fruto de um contexto que estava completamente fora do horizonte que sustentou tanto o problema quanto a hipótese que a resultaram. É preciso, no mínimo, perguntar: faz sentido, no funcionalismo de Luhmann, falar em legitimidade do direito em um contexto que não seja o do Estado? Ou, em um nível mais inicial: faz sentido falar em direito em um contexto que não seja o do Estado?

O próximo capítulo se dedica a responder a essa pergunta: nele, acoplo à abordagem de Luhmann perspectivas que lhe são posteriores e que se destinam justamente a torna-la capaz de apreender uma realidade que, posterior a esses primeiros trabalhos, dissocia direito e Estado, permitindo entender fenômenos como a moderação de conteúdo sob uma perspectiva funcionalista da Teoria dos Sistemas. Após esse ajuste, proponho e testo duas hipóteses quanto ao problema que apresentei no capítulo anterior, sustentando que a transformação nas plataformas se explica tanto a partir da lógica da legitimidade, quanto pela perspectiva de uma constitucionalização do mundo digital.



## CAPÍTULO 5 – DA LEGITIMIDADE ÀS PLATAFORMAS

### 1. INTRODUÇÃO

O objetivo deste capítulo é testar as hipóteses apresentadas para resolver o problema proposto na segunda parte da dissertação. Para contextualizar essas respostas com mais clareza, é importante, nesse início, reconstruir o caminho percorrido no trabalho até agora, conectando as conclusões dos capítulos anteriores com a forma pela qual as seções seguintes estão organizadas.

Em síntese, este é um trabalho sobre como a era digital transformou a regulação da liberdade de expressão. No Capítulo 1, esse tema foi particularizado a partir do argumento de que as modificações causadas por mudanças tecnológicas nos últimos anos, em particular as relacionadas com as telecomunicações, não significaram apenas o surgimento de novos *espaços de manifestação* na internet – em que as pessoas passaram a expressar aquilo que pensam – mas também se traduziram em novos *espaços de decisão* – constituídos nas empresas por trás das plataformas que viabilizam essas manifestações, nos quais aquilo que as pessoas podem ou não dizer nesses ambientes é definido. Como essas decisões são tomadas de forma relativamente independente e a partir de políticas próprias, construídas por essas companhias, entender a regulação da liberdade de expressão apenas a partir do que os Estados fazem deixou de ser o suficiente: é preciso compreender também como funcionam esses mecanismos regulatórios empresariais, que definem, em extensões por vezes maiores que a do Estado, aquilo que as pessoas podem dizer.

Essa proposta, entretanto, encontra um obstáculo no primeiro problema que este trabalho se propôs a resolver: a opacidade. Como a moderação de conteúdo é um mecanismo decisório localizado no interior de empresas, ele não está submetido aos mesmos protocolos de transparência que a atividade regulatória do Estado, e, por isso, não há fontes consolidadas de informações que permitam compreender – e problematizar – a forma como essas decisões são tomadas. Para superar esse obstáculo, o Capítulo 2 se dedicou, primeiro, a esquematizar o modo como esse mecanismo funciona – propondo tanto conceitos quanto sistematizações que permitissem superar a opacidade característica da moderação de conteúdo – e, segundo, a expor as perspectivas disponíveis para interpretá-lo.

Enfrentar a opacidade da moderação de conteúdo na primeira parte da dissertação permitiu, no Capítulo 3, entender os movimentos que a tem transformado e, a

partir disso, avançar para problematizações mais complexas. Em particular, argumentei que uma análise da forma como essas decisões são tomadas pelas plataformas e do conjunto de exigências feitas por seus usuários quanto ao procedimento adotado na moderação de conteúdo indica que essas empresas têm, progressivamente, deixado de lado um formato arbitrário de definir aquilo que as pessoas podem dizer; adotando, em seu lugar, mecanismos procedimentais de transparência, participação e recorribilidade, que propõem uma forma de decidir cada vez mais próxima à do Estado – em particular, semelhante à do devido processo legal utilizado em processos judiciais.

O problema é que essa transformação, que tem moldado a forma como empresas regulam a liberdade de expressão, não tem explicação nas interpretações disponíveis da moderação de conteúdo. Como moderar conteúdo de uma forma opaca e unilateral assegura mais poder e discricionariedade às plataformas, faz mais sentido manter esse modelo do que abrir mão dele – em particular porque não há o que as obrigue a decidir dessa maneira. Esse movimento se torna ainda mais incompreensível porque, como a primeira parte da dissertação indica, a moderação de conteúdo não é construída por acasos, mas sim por decisões cuidadosamente tomadas por essas empresas, que as calcularam para construir o modelo que lhes fosse mais conveniente – tornando essa transformação, no mínimo, implausível. Por isso, faz sentido perguntar: por que as plataformas decidiram mudar o seu modelo decisório, guinando-o a um formato sustentado no que obriga os Estados? Que razões existem para que elas, depois de desenvolverem um sistema próprio de moderação de conteúdo – construído para que detivessem total controle sobre os seus resultados – agora se empenhem em transformá-lo em algo que as faz perder o domínio que têm sobre o ambiente que administram? Por que, afinal de contas, essas empresas estão abrindo mão do seu poder?

Construí uma estratégia para resolver esse problema em dois passos: primeiro, utilizar a Teoria dos Sistemas para identificar a função que os procedimentos desempenham na estrutura política dos Estados – em relação à qual as plataformas têm se aproximado. Segundo, adaptar essas ferramentas dos Estados às plataformas, testando em que medida elas – e a hipótese construída a partir desses instrumentos – são aplicáveis à transformação identificada no procedimento utilizado para moderar conteúdo.

O primeiro passo dessa estratégia foi dado no Capítulo 4. Ele é dedicado a responder à seguinte pergunta: qual é o papel que os procedimentos desempenham na estrutura política dos Estados? Com reconstruí no capítulo anterior, a resposta do funcionalismo estrutural para essa pergunta passa por estabelecer uma relação entre

procedimentos e legitimidade: para Luhmann, procedimentos como o judicial têm a função de legitimar as decisões tomadas pelo Estado, porque organizam mecanismos capazes de criar uma prontidão generalizada para que elas sejam aceitas pelas pessoas. Essa prontidão é construída por estruturas que trazem a conflituosidade para o ambiente controlado do procedimento, submetendo-a a uma incerteza regrada que assegura tanto que uma decisão será proferida, quanto que o seu resultado é incerto – ou seja, que ele é determinado pelo próprio procedimento. Isso induz as pessoas a participarem do processo – isto é, a tentarem influenciá-lo a seu favor –, assumindo posições e se vendo constringidas por regras que o levam, passo a passo, a uma decisão – que, por ser construída pelas partes, torna-se esperada e, mais importante, aceitável, mesmo que seja provavelmente desfavorável a uma delas. O procedimento, nesse sentido, é uma forma de reger a incerteza decisória que induz a participação e confere, com isso, aceitação e legitimidade às decisões do Estado na estrutura política da sociedade moderna.

O segundo passo dessa estratégia é o objeto deste capítulo. Como mencionei no final do anterior, a relação entre procedimento e legitimidade construída por Luhmann em a *Legitimação pelo Procedimento* é, no limite, algo que fundamenta uma hipótese a respeito daquilo que acontece nas plataformas: assim como os Estados estabelecem procedimentos porque, mesmo abrindo mão de parte de seu poder, isso garante aceitabilidade às suas decisões, as plataformas têm construído mecanismos de transparência, participação e recorribilidade porque eles são capazes de fazer das suas decisões algo legítimo. Para que essa hipótese seja testada como uma possível resposta à pergunta do Capítulo 3, entretanto, duas etapas precisam ser percorridas: primeiro, é preciso apurar as ferramentas teóricas que Luhmann utilizou para estabelecer esse diagnóstico, porque elas foram construídas para analisar uma estrutura política que não existe no problema que propus – o Estado; segundo, uma vez concebida essa versão mais apurada da Teoria dos Sistemas, é preciso aplicá-la à construção feita no decorrer da dissertação ao caso da moderação de conteúdo, testando se essa explicação tem respaldo nos dados que refletem a transformação que tem ocorrido nas plataformas.

O capítulo que segue está organizado de acordo com essas duas etapas. Assim, a primeira seção (2) se dedica a refinar as ferramentas da Teoria dos Sistemas que Luhmann utilizou em seu diagnóstico, empreendendo uma tentativa de tornar o funcionalismo estrutural um instrumento útil para analisar a transformação que aconteceu nas plataformas. A pergunta, nesse caso, é a seguinte: a Teoria dos Sistemas é capaz de enfrentar o problema de definir a função de procedimentos, mesmo quando ele se

apresenta fora da “*estrutura política da sociedade moderna*” – isto é, do Estado? Em particular: em que medida a relação entre procedimento e legitimidade que Luhmann sustenta existir nessa estrutura vale, por hipótese, para as plataformas? Argumentarei, em síntese, que a resposta para essa pergunta é afirmativa, mas que ela exige um esforço teórico que dissocie direito e Estado, como na versão original da Teoria dos Sistemas, adaptando-a à realidade sociológica do pluralismo jurídico – algo feito, em específico, por Gunther Teubner. Mais: sustentarei que esse movimento expõe uma segunda função dos procedimentos, associado a um processo de constitucionalização, em que eles atuam autonomizando o direito em relação àquilo que ele regula – algo que é particularmente útil em contextos transnacionais como o das plataformas.

A segunda seção (3) se dedica a tomar essa versão mais apurada da Teoria dos Sistemas – e as relações que ela estabelece entre procedimento e legitimidade e procedimento e autonomização do direito – e aplica-las àquilo que reuni nesta dissertação para esclarecer como funciona a moderação de conteúdo. Faz sentido sustentar que as plataformas adotaram mecanismos que as aproximam da forma como os Estados lidam, por exemplo, com processos judiciais, primeiro, para que a moderação de conteúdo produza decisões legítimas e, segundo, para que ela, produza o direito de uma forma autônoma? Sustentarei que a resposta é sim: em síntese, argumentarei, primeiro, que características da regulação da liberdade de expressão feita pelas plataformas – em particular, a multiplicidade de valores envolvidos e a ausência de outros meios legitimadores, como eleições e processos legislativos – associadas às dificuldades que têm sido enfrentadas por elas na atividade de moderação de conteúdo evidenciam que há, nesse contexto, um problema de legitimidade – que tenta ser, em alguma medida, resolvido pela adoção de procedimentos mais próximos ao do Estado. Segundo, argumento que a adoção desses procedimentos também é um sinal de um movimento mais amplo, de constitucionalização do direito digital, que se concretiza no âmbito das plataformas pela criação de procedimentos que conferem autonomia ao direito das plataformas – isto é, à moderação de conteúdo –, transformando a sua relação tanto com aquilo que ele regula, quanto com as influências que vêm de contextos externos aos da plataforma.

No final, uma conclusão (4) reúne os resultados do capítulo, abrindo caminho para o encerramento da dissertação.

## 2. POR UMA TEORIA DOS SISTEMAS DIFERENTE

## 2.1 Introdução

O objetivo desta seção é oferecer uma versão apurada do funcionalismo estrutural – e, em particular, da Teoria dos Sistemas –, capaz de enfrentar o problema proposto na segunda parte da dissertação. Como o capítulo anterior expôs, o funcionalismo de Luhmann é capaz de explicar o papel desempenhados por determinadas estruturas sociais, construindo relações entre problemas e soluções que existem na sociedade moderna. Essa, entretanto, é uma abordagem que se mostrou limitada: ela fornece uma resposta plausível quanto à razão pela qual existem procedimentos na sociedade, posicionando-os como uma solução ao problema da falta de legitimidade nas decisões tomadas pelo Estado, mas não dispõe, nesse formato inicial, das ferramentas necessárias para estender esse diagnóstico às plataformas<sup>609</sup>. Como a Teoria dos Sistemas de Luhmann associa, ao menos a princípio, o direito ao Estado, a relação entre procedimento e legitimidade é, no limite, uma hipótese, que, para ser testada, exige instrumentos capazes de fazer com que o funcionalismo dê um passo adiante, tornando-se capaz de lidar com uma realidade que não estava claramente posta quando a *Legitimação pelo Procedimento* e *O Direito da Sociedade* foram escritos: aquela em que fenômenos normativos semelhantes ao direito, como a moderação de conteúdo, passam a existir fora do âmbito local e desvinculados do Estado.

Esse passo adiante no funcionalismo estrutural é dado, nesta seção, a partir da reconstrução da releitura que Gunther Teubner<sup>610</sup> – um dos mais importantes intérpretes

---

<sup>609</sup> Essa necessidade de desdobrar a construção original da Teoria dos Sistemas em ferramentas capazes de enfrentar problemas que surgiram com essa nova realidade tecnológica decorre também de uma questão cronológica: mesmo que Luhmann tivesse abordado, em termos gerais, a interseção entre a Teoria dos Sistemas e o progresso tecnológico (*cf.*, por exemplo, os capítulos dedicados a temas correlatos em LUHMANN, N. **Soziologie des Risikos**. Berlin: Walter de Gruyter, 1991, pp. 93/110 e LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 302/312), seus últimos textos datam de 1998, momento em que os impactos decorrentes da internet não estavam claros ainda – a transição para a Web 2.0, por exemplo, começa em 2004; *cf. supra*, Cap. I.2, 2.2.

<sup>610</sup> Teubner não é o único que realiza uma interpretação crítica da Teoria dos Sistemas, e, por isso, a adaptação do funcionalismo estrutural para uma realidade em que direito e Estado são desvinculados não é a única que decorre do trabalho de Luhmann – mas é a que se mostra capaz, como explico nesta seção, de abordar o problema colocado neste trabalho, o que justifica a escolha. Para outras abordagens, *cf.*, em caráter geral, FISCHER-LESCANO, A. Critical Systems Theory. **Philosophy and Social Criticism**, pp. 1/21, 16.11.2011, disponível em <<http://psc.sagepub.com/content/early/2011/11/15/0191453711421600>>, acesso

e repercutores críticos de Niklas Luhmann, em particular no âmbito privado – faz da Teoria dos Sistemas. Em particular, as seções seguintes se dedicam ao modo como Teubner mobiliza a Teoria dos Sistemas para que ela seja capaz de abordar o pluralismo jurídico, isto é, o surgimento de ordens jurídicas além do Estado, como é o caso da *lex mercatoria* e dos códigos corporativos. Com isso, minha intenção é traçar o caminho percorrido por ele quanto a esses dois casos para, a partir das ferramentas que ele aperfeiçoa para conseguir tanto, refazer esse percurso com a moderação de conteúdo das plataformas da internet, testando as hipóteses que aventei no Capítulo 3.

Isso é feito, a seguir, em uma seção única (2.2), em que reproduzo esse trajeto quanto aos códigos corporativos e à *lex mercatoria*. Com isso, meu objetivo é apreender os aperfeiçoamentos realizados por Teubner na Teoria dos Sistemas para, em seguida, aplica-los ao procedimento, chegando a uma conclusão quanto ao seu papel no direito que é produzido de forma desvinculada ao Estado, em particular, nas plataformas. Esse percurso é dividido em três etapas: na primeira (2.2.1), reconstruo o problema que originou essa abordagem; na segunda, a solução proposta por Teubner (2.2.2), e, na terceira (2.2.3), os benefícios que essa resposta tem para resolver o problema do Capítulo 3.

## 2.2 Por um direito fora do Estado

### 2.2.1 O problema de Teubner

Mesmo que situações que ocorrem na internet tenham sido objeto da atenção do funcionalismo estrutural – e, em particular, de Gunther Teubner<sup>611</sup> –, as

---

em 4.1.2020; e, especificamente, por exemplo, o diálogo que tem se desenvolvido entre a teoria crítica e a Teoria dos Sistemas, gerando uma “*Teoria Crítica dos Sistemas*”, inicialmente em BACHUR, J. P. **Às portas do labirinto: para uma recepção crítica da teoria social de Niklas Luhmann**. Rio de Janeiro: Azougue, 2010, 300 pp. e, mais recentemente, em AMATO, L. F., BARROS, M. A.. (org.) **Teoria Crítica dos Sistemas? Crítica, teoria social e direito**. Porto Alegre: Fi, 2018, 430 pp.. O próprio Teubner tem outras interpretações críticas a respeito da Teoria dos Sistemas que deixo de mencionar neste trabalho em razão de sua restrição temática; *cf.*, por exemplo, TEUBNER, G. **Le droit: Un Système Autopoïétique**. Paris: Presses Universitaires de France, 1993, 296 pp.

<sup>611</sup> *Cf.*, a esse respeito, TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, pp. 1/28,

tentativas de adaptá-lo a uma realidade em que Estado e direito não estejam necessariamente relacionados partiram de um contexto anterior. Elas surgem para entender os impactos da globalização no direito, em particular na forma como ela se manifesta, primeiro, na criação, por multinacionais, de códigos de conduta que regulam a forma como os seus empregados devem se comportar, mesmo que localizados em diferentes países; e, segundo, no ressurgimento da *lex mercatoria*, um instrumento de regulação e resolução de conflitos comerciais transnacionais.

Assim, o primeiro contexto de onde esse problema é extraído é o das normatizações internas de empresas multinacionais, que passaram a consolidar em códigos corporativos as normas que regulam a atuação de seus funcionários. Assim, com a ideia de “*códigos corporativos de multinacionais*”<sup>612</sup>, Teubner refere-se aos complexos normativos estabelecidos por companhias que, por atuarem em mais de uma jurisdição, atravessam horizontalmente as regulamentações dos Estados, uniformizando essas condutas pelo mundo. Concretamente, essa forma de normatização toma a forma em códigos, estruturados em três níveis: no primeiro, há “*compromissos específicos da empresa e diretrizes*”<sup>613</sup> que orientam em um grau alto de generalidade o caminho a ser trilhado pela companhia; no segundo, estabelecem-se as “*atividades internas regulatórias e o trabalho dos órgãos executivos*”<sup>614</sup>; no terceiro, subsistem orientações pormenorizadas, como “*regras técnicas e organizacionais [...] que compreendem deveres e tarefas*”<sup>615</sup>.

---

2003, tratado *infra*, Cap. II.5, 3.3. As traduções dos textos de Teubner foram feitas a partir das versões em alemão e cotejadas, quando possível, com as existentes em inglês e/ou francês.

<sup>612</sup> “*Corporate Codes multinationalaler Unternehmen [...]*”. TEUBNER, G. Codes of Conduct multinationalaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. *et al* (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>613</sup> “*1.auf der oberen Ebene existieren firmeneigene Selbstverpflichtungen und Leitlinien; unter den Bedingungen zunehmender rechtlicher Porösität sind sie ein wichtiges Mittel der Erzeugung von Vertrauen und der Legitimitätsbeschaffung*”. TEUBNER, G. Codes of Conduct multinationalaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. *et al* (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>614</sup> “*2.Auf der mittleren Ebene rangieren die Aktivitäten der internen Kontroll- und Implementationsorgane, die teilweise als Berater auftreten, teils als Ermittler und Durchsetzungsorgane fungieren und häufig auch die Funktion übernehmen, die geltenden Normen weiterzuentwickeln*”. TEUBNER, G. Codes of Conduct multinationalaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher

Como tem acontecido na moderação de conteúdo das plataformas da internet, esses códigos corporativos têm semelhanças com a perspectiva tradicional do direito – que o associa à produção normativa dos Estados: eles estabelecem estruturas decisórias, imputando-lhes competências e procedimentos específicos a serem obedecidas, sob pena de nulidade – como é o caso dos conselhos de administração e de fiscalização interna; estabelecem princípios que orientam a conduta das pessoas – a partir de ideias como as de “*missão da empresa*”; e fixam direitos e deveres individuais, como os associados àquilo que os seus empregados devem fazer em situações específicas, ou a como eles são remunerados. A questão é que, apesar dessas semelhanças tanto na existência de uma estrutura normativa, quanto no papel que ela desempenha, esses códigos corporativos são, essencialmente, o resultado de uma atividade desvinculada do Estado, porque surgem no interior de sociedades empresárias, que os estabelecem a partir de critérios que lhe são próprios. Isso fica particularmente claro pela relativa independência que eles têm quanto às fronteiras da jurisdição dos Estados: como os códigos são unificados por estruturas empresariais comuns – e não pelo pertencimento a um país –, as suas regulamentações valem mesmo que em jurisdições diferentes. Por isso, as diretrizes de comportamento, a forma de organização decisória e a conduta ética que se espera de um empregado da Microsoft que exerce suas funções nos Estados Unidos são essencialmente as mesmas que são aplicáveis a um empregado brasileiro. Assim, e, a despeito dessas semelhanças, esses códigos corporativos não são entendidos como direito, porque não são produzidos pelo Estado.

O segundo contexto de onde esse problema é extraído é o da *lex mercatoria*. Ela é, em síntese, “*o direito transnacional das transações econômicas*”<sup>616</sup>, ou seja, uma

---

Mitbestimmung, in Holand, A. et al (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>615</sup> “3. Auf der unteren Ebene dominieren konkrete technische und organisatorische Normen und Regeln, die gleichsam den operativen Kern der Regulative ausmachen; sie umfassen Aufgaben und Pflichten, die man den Leitlinien in dieser Konkretheit nicht hätte entnehmen können und die streckenweise im Format impliziter Basisregeln und Normalitätskriterien vorliegen”. TEUBNER, G. Codes of Conduct multinationalen Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. et al (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>616</sup> “[...] die transnationale Rechtsordnung der Weltmärkte, als der bisher erfolgreichste Fall eines “Weltrechts” jenseits der internationalen politischen Ordnung erwiesen”. TEUBNER, G. Globale Bukowina:



forma de normatização que perpassa os atores econômicos privados que atuam em escala global e que regulam o seu comportamento como partes comerciais umas com as outras. Inicialmente, essa prática normativa foi identificada como um “*‘direito costumeiro’ do comércio internacional*”<sup>617</sup>, construído tanto a partir de práticas entendidas como obrigatórias pelas partes de uma avença comercial, quanto, sob a perspectiva contenciosa, pela “*arbitragem comercial internacional*”<sup>618</sup>, responsável, através dos “*grandes organismos da arbitragem*”<sup>619</sup> – isto é, as cortes arbitrais – por fixar, via precedentes, as regras aplicáveis ao comércio internacional. Teubner se contrapõe a essa visão da *lex mercatoria* como um direito costumeiro, sem, entretanto, contradizer essa base fática que o constitui: é como se houvesse um “*contrato autorregulatório, que vai muito além de uma transação comercial particular e estabelece toda uma ordem jurídica que reclama para si validade global*”, construída tanto por “*regras substantivas*” quanto por “*cláusulas que referem conflitos para uma ‘corte’ de arbitragem que é idêntica à instituição privada que foi responsável por ‘legislar’ o contrato-modelo*”<sup>620</sup>.

Como no caso dos códigos corporativos, também na *lex mercatoria* há características que a aproximam do fenômeno jurídico, mesmo que ela também não resulte da atividade do Estado. Assim, a *lex mercatoria* estabelece, por um lado, mecanismos legislativo-contratuais de criação do direito, isto é, contratos, que constituem a fonte imediata de normas – por disporem sobre direitos e obrigações das partes – e definem a forma como essas regras devem ser “*identificadas, interpretadas e os procedimentos a*

---

Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>617</sup> “[...] *dans un ‘droit coutumier’ du commerce international*”. GOLDMAN, B. Les frontières du droit et ‘lex mercatoria’. **Archives de Philosophie du Droit**, n. 9, 1964, p. 183.

<sup>618</sup> “[...] *c’est l’arbitrage commercial international [...]*”. GOLDMAN, B. Les frontières du droit et ‘lex mercatoria’. **Archives de Philosophie du Droit**, n. 9, 1964, p. 183.

<sup>619</sup> “[...] *le cadre des grands organismes d’arbitrage [...]*”. GOLDMAN, B. Les frontières du droit et ‘lex mercatoria’. **Archives de Philosophie du Droit**, n. 9, 1964, p. 183.

<sup>620</sup> “*Hierbei handelt es sich um einen selbstregulierenden Vertrag, der über eine einfache ökonomische Austauschbeziehung insofern hinausgeht, als er eine eigenständige private Rechtsordnung mit dem Anspruch auf universelle Geltung schafft. Neben substantiellen Regeln enthalten diese Verträge auch Vorschriften, die Konflikte an ein Schiedsgericht verweisen, das gleichzeitig mit dem Gremium identisch ist, das den Modellvertrag ursprünglich geschaffen hat*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

serem adotados para resolver conflitos”<sup>621</sup>. Por outro, surgem instituições adjudicatórias, que também fazem parte da *lex mercatoria* e definem, tal como os tribunais dos Estados, o correto sob a perspectiva “jurídica” da questão: são as cortes arbitrais, por vezes reunidas em estruturas institucionais mais complexas – as câmaras de arbitragem internacional –, que são responsáveis por determinar, a partir de conflitos que lhe são submetidos, quem cumpriu ou não com o que dispunha o contrato.

Novamente, há aqui uma estrutura essencialmente independente dos Estados – tanto os contratos internacionais, nos quais vêm à luz as regras da *lex mercatoria*, quanto as cortes arbitrais, em que eles são julgados, emergem de atores privados –, mas que muito se assemelha com o direito estatal. Nesse sentido, e assim como nos códigos corporativos, a despeito das semelhanças, há uma resistência em se entender essas regras como um fenômeno jurídico: como relata Teubner, a *lex mercatoria* expõe, a partir das tentativas de classifica-la, a existência de um preconceito em se entender algo que não advém do Estado como direito, mesmo que isso implique em “argumentos teóricos cuja pobreza só é contrabalançada pela estreiteza conceitual de seus opostos”<sup>622</sup>.

Assim, há, primeiro, quem entenda que a *lex mercatoria* seria uma forma secundária do fenômeno jurídico, uma espécie de “direito costumeiro”<sup>623</sup> – uma perspectiva incompatível com o nível de complexidade e organização dessa estrutura jurídica; segundo, quem a veja como o direito de uma “*societas mercatorum*” – ou seja, o direito de uma “*comunidade mundial de comerciantes*”<sup>624</sup>, um paralelo moderno da *lex mercatoria* medieval – algo que também dificilmente capta a complexidade dessa estrutura normativa –; e, terceiro, quem a reduza a uma estrutura contratual feita no vácuo, ou seja,

---

<sup>621</sup> “[...] enthält ebenfalls Sekundärregeln, mit denen sie das Verfahren der Identifizierung von Primärregeln sicherstellt und deren Interpretation und Konfliktlösungsprozeduren steuert”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>622</sup> “Diese Fürsprecher der *lex mercatoria* haben allerdings theoretische Argumente entwickelt, deren gedankliche Armut nur von der konzeptuellen Dürftigkeit ihrer Widersacher erreicht wird”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>623</sup> “[...] *Gewohnheitsrechts*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>624</sup> “[...] *eine societas mercatorum - fast als formale Organisation*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

afastada da estrutura normativa do Estado moderno e que, por isso, não faria parte do fenômeno jurídico – algo que ignora a existência de estruturas além das contratuais, como as de arbitragem.

Assim, tanto no caso dos códigos corporativos, quanto no da *lex mercatoria*, Teubner detecta situações em que práticas associadas ao fenômeno jurídico – a produção normativa que estabelece direitos e deveres, a existência de regras que separam o lícito do ilícito, o surgimento de estruturas responsáveis por determinar se a conduta de alguém foi a correta ou não, a emergência de procedimentos regulados por essas regras, a disposição de competências específicas a certas instituições, etc. – não são consideradas como parte do direito porque não decorrem do Estado. A questão é que essa é uma postura que não gera consequências apenas taxonômicas: a questão aqui não é saber como posicionar essas práticas em uma tabela que coloque, de um lado, o que é direito e, do outro, o que não é direito. O problema é mais grave, porque, ao entendê-las como algo à parte do fenômeno jurídico, utilizando um critério que o distingue em razão de sua origem estatal, as ferramentas teóricas desenvolvidas para analisar o direito – inclusive as da Teoria dos Sistemas – tornam-se inaplicáveis à *lex mercatoria* e aos códigos corporativos. Práticas como essas – e, como argumento a seguir, a própria moderação de conteúdo das plataformas de internet – tornam-se aquilo que Teubner chama de “‘ONI’ – *objetos normativos não identificados*”<sup>625</sup>: estruturas sociais que em tudo se parecem com o direito, mas que por razões classificatórias não podem ser interpretadas como direito.

Na Teoria dos Sistemas, essa situação adquire uma configuração particular. Isso porque, por um lado, o funcionalismo de Luhmann pareceu aderir, em *Legitimação pelo Procedimento*, a essa perspectiva que associa o direito ao Estado – e, por isso, não seria, à primeira vista, uma alternativa para entender melhor o funcionamento das práticas normativas que mencionei acima. Por outro, essa associação não é inerente à forma como Luhmann conceitua o direito, mas antes um modo contingente que ele teve de entendê-lo: a Teoria dos Sistemas não define o fenômeno jurídico a partir da estrutura política do Estado, mas sim de acordo com uma prática comunicativa que se autoconstitui de um modo independente aos outros âmbitos do social. Isso abre um flanco de oportunidade, para que se questione em que medida a associação que aparece em Luhmann não possa ser

---

<sup>625</sup> “‘UNOs’ (*unidentified normative objects*)”. TEUBNER, G. *Transnationales Recht: Legitimation durch horizontale Grundrechtswirkung in* VIELLECHNER, L. **Transnationalisierung des Rechts**. Velbrück: Weilerswist, 2013, prefácio (n. p.).

atenuada, ao ponto em que a leitura que faz a Teoria dos Sistemas do fenômeno jurídico permaneça coerente, ao mesmo tempo em que útil para analisar essas novas práticas normativas. Esse é, no fundo, o projeto de Teubner.

Para entender melhor esse movimento – ao qual dedico a próxima seção –, é preciso compreender o modo como Luhmann conceitua o direito e a relação que ele estabelece entre o fenômeno jurídico e o Estado. Só assim a viabilidade dessa adaptação se tornará clara o suficiente para que ela não seja confundida com uma distorção do funcionalismo construído por Luhmann – o que não traria, para a análise dessas manifestações normativas, a capacidade de análise da Teoria dos Sistemas que justifica esse movimento.

O conceito de direito na Teoria dos Sistemas vive dois momentos: o primeiro, que remonta a uma das primeiras obras sociológicas de Luhmann, a *Sociologia do Direito*<sup>626</sup>; e o segundo, que é uma perspectiva de maturidade, inserida em sua teoria social e que vem à luz em sua versão mais consolidada em *O Direito da Sociedade*<sup>627</sup>.

No primeiro momento, Luhmann concebe um conceito de direito de forma paralela às ideias de procedimento e legitimidade, reconstruídas no capítulo anterior. Em todos esses casos, o ponto de partida é um diagnóstico comum: a sociedade moderna é moderna porque é *complexa*. Por complexidade, Luhmann refere-se à circunstância de que, nesse tipo de sociedade, há uma multiplicidade de experiências e de cursos de ações possíveis, que tornam indispensável ao seu participante exercer algum tipo de *seleção*. Ele vive, como disse antes, uma “*situação existencial*”<sup>628</sup> que exige *escolhas*. Mais do que isso, essas são escolhas *difíceis*: como as pessoas têm uma capacidade limitada de assimilar informações, escolher um curso de ação não é uma tarefa fácil, porque é impossível apreender tudo aquilo que é necessário para saber o que é melhor fazer.

Como o capítulo anterior expôs, emergiram, na sociedade moderna, estruturas que aliviam essa pressão seletiva, tornando essas escolhas possíveis. Um mecanismo de ação identificado por Luhmann tanto na verdade científica, quanto no manejo do poder, é o da *redução da complexidade*: seja utilizando seleções já realizadas anteriormente por uma metodologia científica, seja tendo as suas opções constringidas por uma autoridade, o indivíduo se vê, com essas estruturas, diante de um cenário em que há menos

---

<sup>626</sup> LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, 393 pp.

<sup>627</sup> LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, 598 pp.

<sup>628</sup> Cf. *supra*, Cap. II.4, 2.2.2.

opções dentre as quais escolher, o que facilita a tomada de decisão. Nos exemplos do capítulo anterior, se antes as opções de cursos somavam centenas para o vestibulando, agora elas foram reduzidas às três que melhor remuneram seus profissionais; se as possibilidades de escolhas incluíam cursos menos tradicionais, agora elas se restringem à medicina, engenharia e direito, por uma imposição familiar.

Existem, entretanto, outros mecanismos de orientação decisória do indivíduo, que o auxiliam a navegar na complexidade que caracteriza a sociedade moderna. O direito é um deles: ele funciona como um organizador de estruturas sociais capazes de fazer frente à multiplicidade de possibilidades da sociedade moderna, dispondo-as em cursos de ação que enfrentam a imprevisibilidade nas diferentes vertentes da complexidade conferindo-lhes *congruência*. Para entender melhor isso, é importante compreender *quais* estruturas sociais o direito organiza e *como* ele o faz.

O ponto de partida para isso é o conceito de *expectativa*. Expectativas são estruturas sociais pelas quais as pessoas orientam a decisão em uma sociedade complexa pelo que esperam que as outras pessoas façam: alguém dirige mais rápido pela esquerda porque espera que os motoristas mais lentos permaneçam na faixa da direita; uma pessoa que compra algo na padaria entrega uma nota de maior valor porque espera que o caixa lhe devolva a diferença, como troco; nós andamos na calçada porque temos a expectativa de que os motoristas farão com que os carros transitem apenas pela rua. Essas estruturas – que são profundamente inseridas no viver moderno, como os exemplos acima indicam – têm um imenso potencial de auxiliar as pessoas a tomarem decisões, mas trazem consigo um problema: a possibilidade de desapontamento. Um motorista que não liga para as regras de trânsito pode atravancar a faixa da esquerda; um caixa desonesto pode reter parte do troco; um terrorista pode invadir a calçada com um carro.

Estruturalmente, para além da tragédia particular desses casos, a decepção dessas expectativas lança dúvidas sobre esse mecanismo como um meio eficaz de reduzir a complexidade: afinal, devemos ou não esperar que as pessoas respeitem a faixa mais rápida, devolvam o troco ou mantenham o carro na rua? Luhmann argumenta, nesse sentido, que a sociedade moderna construiu *estruturas que protegem o mecanismo de expectativa contra a decepção*, mantendo-o funcionando. Essas estruturas trabalham em três dimensões diferentes: temporal, social e prática.

Na *dimensão temporal*, estabelecem-se, na sociedade moderna, *regras*, que protegem determinadas expectativas contra a decepção futura<sup>629</sup>. Concretamente, isso é perceptível, segundo Luhmann, pela existência de dois tipos de expectativas: as cognitivas e as normativas, que se diferenciam de acordo com o tipo de reação que elas têm uma vez que são desapontadas. Expectativas cognitivas são aquelas que se adaptam à decepção: elas a absorvem como uma forma de aprendizado, que as molda de um modo mais preciso com a realidade que as circunda. Assim, se alguém vai para a praia esperando sol e descobre que o dia está nublado, essa decepção é incorporada como uma parte do aprendizado da pessoa: há dias em que o tempo não corresponde ao da expectativa. Expectativas normativas, de outra forma, são aquelas que se mantêm híidas diante da decepção: elas “*caracterizam-se pela determinação em não assimilar os desapontamentos*”<sup>630</sup>. Não é porque, por exemplo, alguém invade a contramão e provoca uma batida no trânsito, que as pessoas passam a esperar isso como parte da normalidade: antes, aquele que avançou na faixa será considerado culpado pelo acidente e sancionado porque violou uma regra. As expectativas normativas, nesse sentido, contam com mecanismos específicos – como é o caso da sanção – que as reafirmam diante da decepção, tornando claro que aquilo que elas propõem continua sendo o esperado e, mais importante, que elas continuam sendo um mecanismo seguro de orientação das pessoas: elas poderão, mesmo depois de um caso como o relatado, trafegar nas rodovias, porque as outras pessoas continuarão a transitar em suas próprias faixas. Luhmann identifica, nessas expectativas normativas, o que o discurso jurídico chama de *normas*: “*expectativas de comportamento estabilizadas em termos contrafáticos*”<sup>631</sup> – isto é, mantidas híidas independentemente do que acontece.

Na *dimensão social*, estabelecem-se, na sociedade moderna, *instituições*, que protegem determinadas expectativas, na medida em que, na amplitude da sociedade, e dos inúmeros participantes que a compõem, elas podem vir a ser decepcionadas<sup>632</sup>. Em

---

<sup>629</sup> LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, pp. 40 e ss.

<sup>630</sup> “[...] *normative Erwartungen dagegen durch die Entschlossenheit, aus Enttäuschungen nicht zu lernen*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 43. Para a absorção dessa perspectiva na teoria do direito do Brasil, cf. FERRAZ JR., T. S. **Introdução ao estudo do direito: técnica, decisão, dominação**. São Paulo: Atlas, 2018, 4.1.2 (“*Concepção dos fenômenos sociais como situações normadas, expectativas cognitivas e normativas*”), n. p..

<sup>631</sup> “*Normen sind demnach kontrafaktisch stabilisierte Verhaltenserwartungen*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 43.

<sup>632</sup> LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, pp. 64 e ss.

síntese, o esquema de expectativas protegidas na dimensão temporal é sujeito a decepções, porque a complexidade da sociedade moderna pressupõe situações que envolvem mais do que duas pessoas: mesmo em arranjos contratuais, em que duas partes interagem por meio de um acordo, há, para além da expectativa de que elas irão cumpri-lo, a expectativa de que as pessoas que não fazem parte desse arranjo não irão intervir nele. As instituições jurídicas, nesse sentido, são consolidações sociais que protegem essas expectativas quanto a terceiros, ao pressupor o consenso deles nesse tipo de relação. No caso do contrato, por exemplo, ninguém consulta todos os terceiros antes de assinar um instrumento de compra e venda de um carro; mas essas partes tomam como pressuposto que as outras pessoas concordam com que isso seja feito dessa forma, e que irão respeitar esse arranjo. Isso é feito precisamente a partir da instituição contrato: ela confere a chancela de validade do direito àquele arranjo, protegendo-o não só dos próprios contraentes, mas dos terceiros, que implicitamente são considerando como “*concordes*” em relação ao que aconteceu ali.

Na *dimensão prática*, estabelecem-se, na sociedade moderna, *complexos de expectativas*, que protegem as expectativas contra decepções comportamentais que podem se mostrar plausíveis, porque justificáveis<sup>633</sup>. O problema identificado por Luhmann aqui é a possibilidade de que o mecanismo de expectativas seja minado por decepções que se apresentem como razoáveis: é o caso, por exemplo, de quem decide não pagar tributos porque diz que o governo faz mau uso do dinheiro público; ou de quem afirma que continuará poluindo um rio porque não vê sentido na proteção do meio ambiente. Nesses casos, a estabilização das expectativas normativas – isto é, a sua imunização contra a crítica que não apenas a destrói no caso concreto, mas que a enfraquece enquanto um mecanismo que orienta condutas – exige que essas expectativas sejam vinculadas a um complexo mais amplo, que lhes dê *sentido*. Assim, na dimensão prática, surgem na sociedade complexos, por exemplo, de valores, que são posicionados como aceitos consensualmente e que conferem suporte às normas, justificando-as enquanto o resultado de algo que lhes é superior. Como afirma Luhmann, “*a solidez dessa relação baseia-se em um esforço de abstração*”<sup>634</sup>: o sonegador é punido mesmo que ache que o governo emprega mal o dinheiro, porque é dever de todos contribuir para a continuidade do Estado, algo que beneficia a todos nós; o poluidor, por sua vez, precisa ser impedido porque o

---

<sup>633</sup> LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, pp. 80 e ss.

<sup>634</sup> “*Die Festigkeit des Zusammenhangs beruht auf einer Abstraktionsleistung*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 82.

meio ambiente é algo precioso para toda a sociedade e deve, por isso, ser protegido. Em ambos os casos, essas expectativas particulares são ancoradas, na prática, em valores que lhe são superiores e em relação aos quais se supõem haver consenso, como uma forma de manter estáveis as expectativas que foram produzidas a partir deles e, assim, viabilizar que elas continuem a exercer seu papel ante a complexidade.

Essa estabilização dos mecanismos de expectativa nas dimensões *temporal*, *social* e *prática*, que garante que eles continuem cumprindo suas funções, tem um problema: ela nem sempre ocorre de forma congruente. Como afirma Luhmann, “*através da generalização são superadas as discontinuidades típicas a cada dimensão, eliminando os perigos que lhe são típicos*”, mas, como “*os mecanismos de generalização temporal, social e prática são muito heterogêneos*”, “*não se pode supor que eles se processem continuamente de modo idêntico, que eles estejam aferidos por uma espécie de verdade natural do ser e sempre generalizem as mesmas expectativas*”<sup>635</sup>. Dito de outra forma: se é possível garantir a consistência *dentro* de cada uma dessas dimensões – não faz sentido haver expectativas normativas temporalmente contraditórias, por exemplo –, não é possível fazê-lo *entre* cada uma dessas dimensões – a expectativa normativa temporalmente assegurada pode ser uma, enquanto que aquela que corresponde ao valor fixado na dimensão prática é outra. É diante desse problema de incompatibilidade de dimensões “*que o direito constitui sua função social*”<sup>636</sup>: o direito é o mecanismo de alívio de expectativas que garante a sua generalização congruente nas dimensões temporal, social e prática, estabilizando-as de forma socialmente coerente. Quando dizemos, por exemplo, que o direito de um país admite a pena de morte, estamos refletindo o fato sociológico de que as expectativas normativas estão congruentemente alinhadas em todas as suas três dimensões, e é de se esperar não só a aplicação desse tipo de sanção em alguns casos, quanto que essa

---

<sup>635</sup> “*Das Gemeinsame besteht darin, daß durch Generalisierung die jeweils dimensionstypischen Diskontinuitäten überbrückt, die jeweils dimensionstypischen Gefahren ausgeschaltet werden. [...] Die Mechanismen zeitlicher, sozialer und sachlicher Generalisierung, die wir im vorstehenden analysiert haben, sind sehr heterogener Art. Man kann daher nicht unterstellen, daß sie von vornherein gleichgeschaltet laufen; daß sie durch eine Art natürliche Wahrheit des Seins darauf geeicht seien, stets dieselben Erwartungen zu generalisieren*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, pp. 94/95.

<sup>636</sup> “[...] *im Hinblick auf dieses Problem hat das Recht seine gesellschaftliche Funktion*”. LUHMANN, N. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, p. 95.



expectativa não irá mudar, e, mesmo que seja em algum caso desapontada, continuará sendo uma baliza confiável para a conduta das pessoas.

No segundo momento<sup>637</sup>, Luhmann concebe o direito integrando-o a uma perspectiva que abrange o fenômeno social como um todo – transitando de uma sociologia jurídica para uma sociologia da sociedade. Esse movimento insere o direito em um quadro geral de interpretação dos fenômenos da sociedade, agrupando-os em subsistemas sociais que constroem diferenças capazes de delimitá-los em relação ao todo social, particularizando-se não só como subsistema direito, mas também em subsistemas como economia, ciência, política e religião, por exemplo<sup>638</sup>.

A base desse movimento é a identificação da comunicação como o fenômeno social por excelência. Em síntese, o Luhmann sociólogo da sociedade – e não apenas o do direito – vê-se diante do desafio de fundar sua teoria em um conceito de sociedade que não se veja premido pelas restrições dos anteriores. Em particular, ele argumenta que a existência de quatro “*obstáculos epistemológicos*”<sup>639</sup>, característicos das formas anteriores de se ver o social, impedem a sociologia de avançar como um instrumento analítico da modernidade. Eles se configuram, em sequência, pela prática de, *primeiro*, conceituar a sociedade como algo constituído por indivíduos ou por relações entre indivíduos; *segundo*, estabelecer o consenso como um elemento que estabelece ou integra o social; *terceiro*, identificar a multiplicidade de Estados com a multiplicidade de sociedades; e *quarto*, ter a necessidade de abordar epistemologicamente a relação entre

---

<sup>637</sup> A divisão entre um primeiro momento marcado por “*Sociologia do Direito*” e um segundo, associado a “*O Direito da Sociedade*” reflete o primeiro e o início do terceiro círculos identificados por Jean Clam no desenvolvimento da obra de Luhmann; *cf.*, a esse respeito, CLAM, J. **Droit et Société dans la sociologie de Niklas Luhmann: Fondés en contingence**. Paris: Presses Universitaires de France, 1997, p. 5. Essa divisão reflete mais uma necessidade explicativa do que a existência de momentos estanques na Teoria dos Sistemas; é possível identificar textos intermediários, isto é, anteriores à publicação de “*O Direito da Sociedade*” em que as suas principais ideias já eram abordadas. *Cf.*, por exemplo, LUHMANN, N. Operational Closure and Structural Coupling: The Differentiation of the legal system. **Cardozo Law Review**, v. 13, pp. 1.419/1.441, 1991.

<sup>638</sup> Resumi o fenômeno da diferenciação funcional dos subsistemas sociais ao necessário para construir o meu argumento; para uma abordagem mais aprofundada a respeito, *cf.* LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 2, pp. 595 e ss, LUHMANN, N. Differentiation of Society. **The Canadian Journal of Sociology**, v. 2, n. 1, 1977, pp. 36 e ss e LUHMANN, N. **Soziologische Aufklärung 1**, Opladen: Westdeutscher Verlag, 1970, pp. 148 e ss.

<sup>639</sup> LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 24/25.

sociólogo e sociedade a partir da lógica de, respectivamente, um sujeito observador e um objeto observado.

O argumento de Luhmann prossegue, nesse sentido, em dois passos: no *primeiro passo*, ele expõe a fragilidade das concepções que sustentam os conceitos pretéritos de sociedade. Elas são imprecisas conceitualmente – dizer que a sociedade é constituída por indivíduos diz muito pouco sobre aquilo que ela é, porque deixa aberta a pergunta do que exatamente isso significa – como diz Luhmann, “*ela [a sociedade] é composta de braços, pernas, pensamentos e enzimas?*”<sup>640</sup>; não refletem as características da sociedade moderna – mais marcada pelo dissenso do que pelo consenso, como argumentei acima<sup>641</sup>; adotam critérios arbitrários – como o território de um país, algo injustificável em uma sociedade moderna em que há aspectos sociais, como as conclusões científicas, que não dependem dessas divisões; e supõem, sem demonstrar, que o sociólogo é capaz de se colocar à parte da sociedade para descrevê-la.

No *segundo passo*, Luhmann propõe uma saída conceitual capaz de superar essas limitações. Sua resposta, em síntese, acopla dois conceitos – de *comunicação* e de *sistema* – para sustentar que *a sociedade é um sistema social que reproduz comunicações*. Sua proposta é de que o elemento constitutivo da sociedade não seja mais a *ação social*, mas sim a *comunicação*: ele a toma como a atividade social por excelência, porque ela só pode acontecer *dentro* da sociedade<sup>642</sup> – e, por isso, reflete melhor o fenômeno social do que a ação, que não depende desse contexto específico para ocorrer. Essa centralidade das comunicações como os elementos que constituem a sociedade é associada à ideia de sistema, que oferece um recurso intelectual capaz de explicar como esses componentes se

---

<sup>640</sup> LUHMANN, N. **Der Begriff der Gesellschaft**. Disponível em <[http://www.hyperkommunikation.ch/literatur/luhmann\\_begriff\\_gesellschaft.htm](http://www.hyperkommunikation.ch/literatur/luhmann_begriff_gesellschaft.htm)>, acesso em 4.1.2020, n. p.; LUHMANN, N. O Conceito de Sociedade, in NEVES, C. E. B.; SAMIOS, E. M. B. (Org.). **Niklas Luhmann: a nova teoria dos sistemas**. Porto Alegre: Editora da UFRGS; Goethe-Institut, 1997, p. 76; no mesmo sentido, LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 26/27 e LUHMANN, N. **Einführung in die Theorie der Gesellschaft**. Heidelberg: Carl Auer Verlag, 2005, pp. 32 e ss.

<sup>641</sup> Cf. *supra*, Cap. II.4, 2.2.2.

<sup>642</sup> Para uma abordagem geral sobre esse argumento, cf., LUHMANN, N. ‘Was ist der Fall’ und ‘Was steckt dahinter?’ Die zwei Soziologie und Gesellschaftstheorie. **Zeitschrift für Soziologie**, v. 22, n. 4, 1993, pp. 252/253; para uma perspectiva específica quanto a esse mecanismo social de observação *na* sociedade e *da* sociedade, cf. LUHMANN, N. Deconstruction as Second-Order Observing. **New Literary History**, v. 24, n. 4, 1993, pp. 775/777 e LUHMANN, N. **Die Realität der Massenmedien**. Opladen: Westdeutscher Verlag, 1996, pp. 169 e ss.

articulam entre si, construindo o fenômeno social. As comunicações, nesse sentido, são tomadas como resultados umas das outras, constituindo um ciclo que se autoreproduz de acordo com uma lógica que as diferencia de todo o resto. Cada comunicação produzida, nesse sentido, é uma diferença que é estabelecida entre aquilo que faz parte da sociedade – e é comunicação – e aquilo que não faz parte da sociedade – e não é comunicação:

*“[...] podemos, a partir de agora, determinar o conceito de sociedade. A sociedade é o sistema abrangente de todas as comunicações, que se reproduz autopoieticamente, na medida em que produz, na rede de conexão recursiva de comunicações, sempre novas (e sempre outras) comunicações. A emergência de um tal sistema inclui comunicações, pois elas só são passíveis de conexão internamente, excluindo todo o resto.”<sup>643</sup>*

Luhmann sustenta, nesse sentido, que o conceito de sociedade como um sistema social de comunicações é melhor que os que o antecederam, porque é mais preciso – ele não depende de se definir o que exatamente do indivíduo constitui a sociedade; dispensa o consenso; não estabelece critérios arbitrários como o da divisão territorial; e, como a comunicação sobre a sociedade ocorre dentro dela própria, torna desnecessária uma postura epistemológica que pressupõe a capacidade de o sociólogo se colocar externamente ao seu objeto de estudo para analisá-lo.

Lançadas as bases de sua teoria social, Luhmann acomoda a existência de diferentes âmbitos da sociedade – como o direito, a ciência, a economia, a política e a religião – distinguindo diferentes tipos de comunicações. Assim, se *a sociedade abrange todas as comunicações, âmbitos específicos da sociedade abrangem tipos específicos de comunicações*<sup>644</sup>. Luhmann argumenta, nessa direção, que da mesma forma que a comunicação, em geral, é o resultado de um sistema que reproduz uma diferença entre aquilo que *é* comunicação e aquilo que *não é* comunicação; a comunicação jurídica é o resultado de um subsistema, localizado dentro da sociedade, que reproduz a diferença entre

---

<sup>643</sup> LUHMANN, N. **Der Begriff der Gesellschaft**. Disponível em <[http://www.hyperkommunikation.ch/literatur/luhmann\\_begriff\\_gesellschaft.htm](http://www.hyperkommunikation.ch/literatur/luhmann_begriff_gesellschaft.htm)>, acesso em 4.1.2020, n. p.; LUHMANN, N. O Conceito de Sociedade, in NEVES, C. E. B.; SAMIOS, E. M. B. (Org.). **Niklas Luhmann: a nova teoria dos sistemas**. Porto Alegre: Editora da UFRGS; Goethe-Institut, 1997, p. 83.

<sup>644</sup> LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 2, pp. 745 e ss; *cf.*, no mesmo sentido, LUHMANN, N. **Einführung in die Theorie der Gesellschaft**. Heidelberg: Carl-Auer Verlag, 2005, pp. 235 e ss.

aquilo que *é* comunicação jurídica e aquilo que *não é* comunicação jurídica – algo que também vale para cada um dos outros âmbitos sociais.

Nesse contexto, o direito tem a sua especificidade – e o seu conceito –, na forma de um subsistema social que reproduz comunicações jurídicas que se distinguem das demais por uma *função* que lhe é específica: a de garantir, congruentemente, as expectativas normativas. É nesse ponto que o raciocínio do primeiro e do segundo Luhmann se encontram: a generalização das expectativas normativas nas dimensões temporal, espacial e prática é acoplada ao contexto sociológico geral da Teoria dos Sistemas pelo primado da *função*, que imputa a esse subsistema social o papel de assegurar a higidez dessa forma de orientar socialmente o indivíduo na sociedade complexa. Concretamente, esse papel é exercido por um *discrímen*, identificável na comunicação jurídica, que é a sua capacidade de atribuir a cada hipótese a característica de ser lícita ou ilícita: dizemos, por exemplo, que o direito do Brasil proíbe o homicídio, o que significa atribuir a esse tipo de situação a marca do “*ilícito*” – ao mesmo tempo em que exceções, como a legítima defesa, atribuem a outras circunstâncias a marca do “*lícito*”. Como acontece na sociedade em geral, o direito é um constituído por um ciclo comunicativo que se reproduz a partir das suas próprias comunicações, tomando a si mesmo como uma referência para aquilo que ele é. Em síntese, repercutindo-se o primeiro Luhmann no segundo:

*“Concretamente, o direito lida com a função de estabilizar expectativas normativas regulando como elas são generalizadas quanto às suas dimensões temporal, prática e social. O direito permite saber quais expectativas atingirão a aprovação social e quais não irão [...] [e] isso significa que podemos viver em uma sociedade mais complexa, em que os mecanismos e de interação para assegurar a confiança não são mais suficientes”*<sup>645</sup>.

Como o excerto acima – retirado de *O Direito da Sociedade* – demonstra, o primeiro e o segundo Luhmann não construíram conceitos essencialmente diferentes do direito, mas o abordaram por caminhos diversos: primeiro, enfocando a sua função como generalizador de expectativas normativas; segundo, contextualizando-o em uma teoria

---

<sup>645</sup> “*Konkret geht es um die Funktion der Stabilisierung normativer Erwartungen durch Regulierung ihrer zeitlichen, sachlichen und sozialen Generalisierung. Das Recht ermöglicht es, wissen zu können, mit welchen Erwartungen man sozialen Rückhalt findet, und mit welchen nicht. [...] Man kann sich in höherem Maße riskantes Vertrauen oder auch Mißtrauen leisten, wenn man dem Recht vertrauen kann*”. LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, pp. 131/132.

social mais ampla, que o localiza como um subsistema social, com função e códigos específicos que o acoplam à percepção de a *Sociologia do Direito*.

Seja em seu acesso inicial ou posterior ao direito, a característica que transpassa esses momentos é o fato de que *a Teoria dos Sistemas não define o direito em função do Estado*. Isto é: *o Estado não é um elemento essencial ao conceito de direito*. O direito é definido a partir de uma *observação*, que revela a dinâmica do fenômeno jurídico a partir da sua *função* – e nenhuma das duas está associada ao Estado. Isso fica claro quando cada um dos momentos que delineei acima é analisado sob essa perspectiva. Assim, no primeiro momento, Luhmann concebe o direito a partir de uma característica da sociedade moderna – a necessidade de seleção imposta pela complexidade. Aqui, ele é uma estrutura de generalização congruente das expectativas normativas, um papel que exerce mantendo-as coerentes nas dimensões temporal, espacial e prática, *para isso não dependendo necessariamente do Estado*. Nesse sentido, é evidente que, por exemplo, em a *Sociologia do Direito*, Luhmann menciona aspectos da atuação dessa estrutura política no direito – como é o caso da sanção estatal, que tem o papel de reforçar o caráter normativo das expectativas. Não há, entretanto, uma exigência teórica na definição de direito que particularize essa atuação como algo que deve necessariamente advir do Estado. Nada impede, por exemplo, que essa reafirmação do caráter normativo das expectativas seja realizada por uma sanção que não seja proveniente do Estado, ou mesmo por outros mecanismos que não sejam necessariamente sancionatórios.

No segundo momento, em que esse direito generalizador de expectativas aparece como um subsistema inserido em uma sociedade constituída por comunicações, a sua própria formação – isto é, a sua diferenciação em relação a todas as outras comunicações – pressupõe um isolamento operativo, isto é, uma capacidade de se reproduzir independentemente daquilo que lhe é externo, que torna o seu conceito incompatível com a presença necessária do Estado como parte de sua formação. Se o direito existe como um âmbito social específico porque ele reproduz uma diferença entre aquilo que é direito e aquilo que não é direito, ele não pode, ao mesmo tempo, depender da ideia política de Estado para existir. Novamente, isso não significa que conceitos sociológicos que estabeleçam vínculos entre um e outro não possam ser construídos: a ideia de constituição, por exemplo, que discuto com mais detalhes na seção seguinte<sup>646</sup>, é um exemplo de como isso pode acontecer. Esse, entretanto, é um vínculo contingente, e

---

<sup>646</sup> Cf. *infra*, Cap. II.5, 3.3.

que não precisa existir para que se possa falar em direito: na Teoria dos Sistemas, o fenômeno jurídico não depende do fenômeno estatal.

Assim, por um lado, o direito não resulta do Estado<sup>647</sup> – mas sim do que caracteriza a sociedade moderna; por outro, ele não tem como função resguardá-lo, nem depende dessa estrutura política para existir – mas é concebido justamente a partir da sua capacidade de se constituir autonomamente. Isso afasta a Teoria dos Sistemas de perspectivas que vinculam um ao outro – conceituando o direito, por exemplo, como um conjunto de normas promulgadas por um soberano – e, por isso, mobilizar as suas ferramentas para entender o fenômeno jurídico não implica trazer, junto, a existência do Estado como um pressuposto necessário. O recurso, primeiro, à comunicação como constitutiva da sociedade e do direito, e, segundo, ao sistema como estrutura que constrói e articula a comunicação jurídica, cria uma forma de se apreender o fenômeno jurídico que não depende de considerá-lo estatal. Como afirma Teubner, ao aventar possibilidades de abordagens para a *lex mercatoria*:

*“Como a teoria sociológica contemporânea do direito lida com a lex mercatoria e outras formas de direito global sem o Estado? [...] Há, entretanto, um tipo de teoria do direito que explicitamente se faz dependente das práticas jurídicas. Ela observa o direito como um processo de auto-organização que define, autonomamente, as suas fronteiras. [...] A teoria não tenta delinear o que está dentro ou fora do direito, mas produz, em seu lugar, instrumentos de observação. Ela observa as observações da prática jurídica”*<sup>648</sup>.

---

<sup>647</sup> Um elemento que corrobora essa constatação é a forma como a relação entre política – âmbito ao qual Luhmann vincula o Estado (cf. LUHMANN, N. **Essays on Self-Reference**. New York: Columbia University Press, 1990, pp. 166 e ss) – e direito é trabalhada na Teoria dos Sistemas. Assim como ocorre, por exemplo, com a economia, a política tem uma função à parte do direito, o que se reflete em uma espécie de “*localização teórica*” também distinta: a ela é dedicada uma obra à parte, em que delineado o seu papel como um instrumento de tomada de decisão coletivamente vinculante (cf. LUHMANN, N. **Die Politik der Gesellschaft**. Frankfurt: Suhrkamp, 2002, pp. 69 e ss), ao qual associados capítulos, nas obras de Luhmann dedicadas ao direito, em que a intersecção entre ambos é trabalhada (cf., por exemplo, LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, pp. 407/439). No mesmo sentido, e quanto à ausência de identidade entre o Estado e o fenômeno social, cf. LUHMANN, N. **Soziologische Aufklärung 4: Beiträge zur funktionalen Differenzierung der Gesellschaft**. Opladen: Westdeutscher Verlag, 1987, pp. 67 e ss.

<sup>648</sup> “*Wie würde die gegenwärtige gesellschaftstheoretische Diskussion des Rechts mit der lex mercatoria und andere Formen globalen Rechts jenseits des Staats umgehen? [...] Es gibt jedoch Spielarten der Rechtstheorie, die explizit Selbstbeschreibungen des Rechts zu ihrem Ausgangspunkt machen und sich damit ihrerseits an die Abgrenzungen der Rechtspraxis selbst binden. Sie beobachten das Recht als einen selbstorganisierenden Prozeß, der autonom seine eigenen Grenzen festlegt [...] Sie beobachten, wie*

Essa constatação é importante, porque expõe que a associação entre direito e Estado, que baseia diagnósticos como o de a *Legitimação pelo Procedimento* – e que é exposta, por exemplo, no recorte que analisou o procedimento a partir dos processos judicial, eleitoral e legislativo<sup>649</sup> – é contingente. Isso significa que, para mobilizar os instrumentos da Teoria dos Sistemas em uma investigação sociológica um fenômeno jurídico, não é necessário “*comprar*” também a postura de Luhmann quanto à relação entre direito e Estado: ela não é constitutiva do conceito de direito, mas, antes, reflete uma forma de ver essa relação que pode ser questionada sem que a capacidade analítica da Teoria dos Sistemas seja, a princípio, afetada. Isso abre, ao menos, um flanco de atuação para que essas ferramentas sejam utilizadas para analisar fenômenos como os códigos corporativos e a *lex mercatoria* – e, como faço na seção seguinte, a própria moderação de conteúdo – em que características do fenômeno jurídico-estatal surgem em contextos não estatais. É essa a solução proposta por Teubner para a ausência de instrumentos capazes de interpretar esse novo fenômeno, que reconstruo a seguir para, mobilizando-a, tentar compreender o que tem acontecido na moderação de conteúdo.

### 2.2.2 A solução de Teubner

Como essa proposta – viabilizada no flanco aberto pela Teoria dos Sistemas e seu conceito de direito desvinculado do Estado – concretiza-se em uma perspectiva capaz de fazer frente a esse conjunto de fenômenos jurídicos não estatais? Para responder a essa pergunta, reconstruo, a seguir, a forma como Teubner solucionou esse problema, em particular quanto à *lex mercatoria* e os códigos corporativos – os primeiros que ocasionaram uma visão diferente do direito. Minha estratégia, a esse respeito, é entender o que foi feito nesses casos para, em seguida, aplicar essa estratégia na moderação e conteúdo das plataformas da internet – que, como argumento a seguir, faz parte de um contexto semelhante.

Antes de particularizar essa proposta nos casos acima, é preciso entender o quadro geral da adaptação proposta por Teubner. Ela se desdobra em dois passos: no

---

*Rechtsoperationen ihrerseits die Welt beobachten*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>649</sup> Cf. *supra*, Cap. II.4, 2.3.1 e 2.3.2.

primeiro, ele unifica esses eventos particulares em um fenômeno conjunto – os impactos da globalização para o direito. No segundo, ele propõe que a Teoria dos Sistemas apreenda esse fenômeno a partir de uma reinterpretação do pluralismo jurídico, algo viável porque, como a seção anterior mostrou, a forma como Luhmann conceitua o direito permite que ele seja concebido fora do Estado, mesmo que essa possibilidade não tenha sido explorada em toda a sua potencialidade nas versões originais da teoria.

O primeiro passo, portanto, compreende entender as linhas gerais desse movimento que produz “*ONI – objetos normativos não identificados*”<sup>650</sup>, isto é, estruturas que, como os códigos corporativos e a *lex mercatoria*, têm características que os aproximam do direito – normas, processos, atribuições de competência, instâncias decisórias vinculantes, etc. –, mas que não são associados a um Estado.

O surgimento desses fenômenos é observado, por Teubner, a partir da percepção de que a sociedade moderna passa por um processo de *globalização*. Assim como procedimento e legitimidade, esse é um termo com significado específico na Teoria dos Sistemas: “*globalização, como nós a vivemos hoje, significa uma mudança de proeminência no princípio primário da diferenciação: uma mudança da diferenciação territorial para a funcional em um nível mundial*”<sup>651</sup>. Ele capta, em síntese, a circunstância de que os âmbitos do social identificados por Luhmann como formas específicas de se comunicar – a cultura, a economia e a ciência, por exemplo – passaram a romper as fronteiras territoriais dos sistemas políticos para se constituírem como sistemas globais próprios e independentes do Estado: um Picasso é uma obra de arte no Brasil e na Itália; os impactos da recessão de 2008 não cessaram quando chegaram à fronteira dos Estados Unidos com outros países; uma conclusão científica que leve à cura da AIDS vale tanto para um paciente europeu quanto um asiático<sup>652</sup>. Assim, “*enquanto o processo político atingiu apenas uma proto-globalidade nas relações internacionais [...] os outros*

---

<sup>650</sup> ‘UNOs’ (*unidentified normative objects*)”. TEUBNER, G. *Transnationales Recht: Legitimation durch horizontale Grundrechtswirkung in* VIELLECHNER, L. **Transnationalisierung des Rechts**. Velbrück: Weilerswist, 2013, prefácio (n. p.).

<sup>651</sup> “*“Globalisierung” bedeutet in ihrer heutigen Form einen Wandel des Grundprinzips der Differenzierung: einen Wandel von territorialer zu funktionaler Differenzierung auf der Welt-Ebene [...]*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290, (n. p., nota de rodapé).

<sup>652</sup> Cf., para uma abordagem quanto à característica global da comunicação científica, LUHMANN, N., BEHNKE, K. *The Modernity of Science*. **New German Critique**, v. 61, 1994, pp. 21 e ss.



*subsistemas sociais já começaram a formar uma autêntica sociedade global, ou melhor, uma multitude fragmentária de diversas sociedades globais*<sup>653</sup>. Globalização, nesse sentido, significa construção de sistemas comunicativos globais correspondentes a âmbitos – e funções – específicas da sociedade, que, constituindo sociedades globais temáticas, paulatinamente ignoram as restrições políticas locais – que, de certa forma, ficaram para trás nessa transformação – em nome de uma forma universal de tratar o social – e que é manifesta, em particular, em âmbitos como o direito, a economia e a ciência.

A pergunta que conecta esse fenômeno com o surgimento da *lex mercatoria*, dos códigos corporativos e, como argumento a seguir, da moderação de conteúdo, é a seguinte: como a globalização muda o direito? Ou: “*o que esse cenário de múltiplos movimentos implica para o direito?*”<sup>654</sup> De que forma o direito se transforma em uma realidade em que os discurso sociais não estão mais restritos pelas fronteiras políticas? Como ele se globaliza? Diferente do que acontecia com a economia, a ciência e a arte – que não estavam, em âmbito local, vinculados necessariamente à política – direito e Estado têm uma ligação umbilical, que se reflete, por exemplo, na própria abordagem do Luhmann de a *Legitimação pelo Procedimento*, que vincula ambos, mesmo que de maneira contingente. Isso leva a uma situação particular, porque não há uma estrutura política global – ao menos não no sentido de um Estado centralizado – que possa replicar a relação que existe entre política e direito na globalização. *A globalização, por isso, não apenas traz o direito do âmbito local para o global, mas faz com que ele rompa as suas amarras com o Estado*. Ela o transforma, na medida em que *um direito global não pode mais ser vinculado a uma estrutura política, porque uma estrutura política nesses moldes*

---

<sup>653</sup> “*Während die Politik in Gestalt der "inter - nationalen Beziehungen" erst einen Zustand der Proto-Globalität [...] erreicht hat, haben die anderen sozialen Subsysteme bereits begonnen, eine echte Weltgesellschaft zu formen - oder, besser gesagt: eine fragmentierte Anzahl unterschiedlicher Weltsysteme*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.). No mesmo sentido, cf. LUHMANN, N. **Soziologische Aufklärung 2: Aufsätze zur Theorie der Gesellschaft**. Wiesbaden: Springer, 1975, pp. 51 e ss, LUHMANN, N. *Globalization or World Society: How to Conceive of Modern Society?* **International Review of Sociology**, v. 7, n. 1, pp. 67/79, 1997, LUHMANN, N. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, pp. 145 e ss e LUHMANN, N. **Die Politik der Gesellschaft**. Frankfurt: Suhrkamp, 2002, pp. 220 e ss.

<sup>654</sup> “*Welche Bedeutung hat dieses Szenario der 'Globalisierung mit unterschiedlichen Geschwindigkeiten' nun für das Recht?*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

*simplesmente não existe*. Por isso, a emergência de um discurso jurídico com dimensões globais pressupõe, mesmo que por hipótese, que esse é um direito sem uma correspondência política direta, ao menos de alguma forma paralela à do Estado local – algo que, como exponho adiante, impacta o próprio direito, porque, diferente da ciência, que nunca teve esse tipo de amarra, o fenômeno jurídico foi essencialmente construído a partir de uma amálgama sua com o Estado. Retirá-lo dessa posição em que constituído, portanto, tem seu preço.

Concretamente, a globalização do direito se manifesta em uma “*multitude de ‘discursos fragmentados’ que são hermeticamente fechados entre si e podem ser identificados nos numerosos tipos de direito informal que são gerados de forma independente ao Estado e que operam em vários níveis de formalização*”<sup>655</sup>. Surgem, portanto, códigos corporativos, ou seja, discursos jurídicos internos às empresas, que regulam o seu funcionamento perpassando, de maneira horizontal, jurisdições estatais e construindo um fenômeno jurídico global; a *lex mercatoria*, como uma forma de regular condutas e tomar decisões a respeito de relações comerciais que se localizam além dos limites da jurisdição; e, por fim, como discuto melhor na próxima seção, a moderação de conteúdo, um instrumento de regulação da liberdade de expressão que define e aplica regras quanto àquilo que as pessoas podem dizer em um nível global.

Seja, assim, na regra que fixa os limites éticos do empregado de uma multinacional, no laudo de uma corte arbitral internacional, ou na decisão que retira um conteúdo publicado em uma plataforma da internet, simplesmente não faz sentido procurar correspondências políticas semelhantes à do Estado para esse discurso semelhante ao jurídico, porque, a nível global, isso não existe. Isso, mais do que uma questão taxonômica – que distingue um direito local, vinculado ao Estado<sup>656</sup>, de um “*direito*” global,

---

<sup>655</sup> “*In ihrer hermetischen Geschlossenheit lassen sich die vielfältigen ‘fragmentierten Diskurse’ an Hand zahlreicher informeller Regeln identifizieren, die staatsunabhängig entstehen und auf verschiedenen informellen Ebenen operieren*”. TEUBNER, G. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne in SCHMIDT, E., WEYERS, H. (ed.). **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, 1995, p. 191.

<sup>656</sup> Nem todo direito local é vinculado ao Estado; o fenômeno do pluralismo jurídico, nesse sentido, pode ocorrer em nichos localizados dentro do território de um Estado. *Cf.*, a esse respeito, SANTOS, B. The Law of the Oppressed: The Construction and Reproduction of Legality in Pasargada. **Law & Society Review**, v. 12, n. 1, pp. 5/126, 1977, quanto às normatizações que surgiram nas comunidades brasileiras à margem do Estado; e ELLICKSON, R. **Order without Law: How Neighbors Settle Disputes**. Cambridge: Harvard University Press, 1991, 316 pp., quanto às normatizações que surgiram em comunidades americanas quanto à

desvinculado do Estado – é algo importante de ser percebido, porque traz questionamentos quanto a *como o direito pode ser abordado pela sociologia*. Nas palavras de Teubner: “*Como a teoria sociológica contemporânea lida com a lex mercatoria e outras formas de direito global sem o Estado?*”<sup>657</sup>.

A resposta de Teubner a essa pergunta é o segundo passo de seu argumento: sua proposta é de que a Teoria dos Sistemas apreenda esse novo fenômeno jurídico dando corpo funcionalista ao pluralismo jurídico. De um lado, ele justifica mobilizar a Teoria dos Sistemas porque, como expus na conclusão da seção anterior, ela compreende o direito não a partir do seu *conteúdo* – normas jurídicas produzidas pelo Estado, por exemplo –, mas sim como uma *prática social de generalização de expectativas operativamente isolada* e, por isso, que não depende de um vínculo com o Estado para existir. Como citei antes, “*essa Teoria [dos Sistemas] não tenta delinear o que está dentro ou fora do direito, mas produz, no lugar, instrumentos que o observam*”<sup>658</sup>. Por isso, utilizá-la para entender o direito não depende da adoção de uma posição que o vincule ao fenômeno estatal, mesmo que – de forma contingente – essa relação tenha aparecido no Luhmann de *Legitimação pelo Procedimento*.

De outro, o pluralismo jurídico é o conceito referido por Teubner ao analisar práticas sociais a exemplo da *lex mercatoria*, dos códigos corporativos – e, por outros autores, mesmo fenômenos locais, como as normatizações construídas em comunidades periféricas do Brasil<sup>659</sup> –, como uma forma de melhor compreendê-los por uma visão

---

responsabilidade pela fuga de gado. Para uma abordagem sistêmica do pluralismo jurídico, mas diferente da de Teubner, cf. NEVES, M. Do pluralismo Jurídico à Miscelânea Social: o problema da falta de identidade da(s) esfera (s) de juridicidade da modernidade periférica e suas implicações na América Latina. **Revista Direito em Debate**, v. 4, n. 5, 2013; e CAMPILONGO, C. F. **O Direito na Sociedade Complexa**. São Paulo: Saraiva, 2011, pp. 61 e ss (“*Crise do Estado, Mudança Social e Transformação do Direito no Brasil*”).

<sup>657</sup> “*Wie würde die gegenwärtige gesellschaftstheoretische Diskussion des Rechts mit der lex mercatoria und andere Formen globalen Rechts jenseits des Staats umgehen?*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>658</sup> “*Eine solche Theorie legt nicht von sich aus, "analytisch" wie es heißt, fest, was innerhalb der Grenzen des Rechts liegt und was nicht*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>659</sup> SANTOS, B. The Law of the Oppressed: The Construction and Reproduction of Legality in Pasargada. **Law & Society Review**, v. 12, n. 1, pp. 5/126, 1977; não há uma cisão entre as análises sustentadas pelo pluralismo jurídico quanto a fenômenos locais e globais, mas antes uma linha de continuidade que permite aplicar as conclusões das análises de um em outro; para um exemplo disso, cf. a análise feita em FARIA, J. E.

diferente quanto àquilo que é o direito. Essa transformação, que é a principal marca desse movimento e, por isso, fundamenta o seu diálogo com a Teoria do Sistemas, centra-se em uma contraposição à perspectiva que define o direito a partir de sua origem estatal. O pluralismo jurídico opõe-se, nesse sentido – no caso específico das práticas comerciais transnacionais – à posição que “*evoca a soberania do estado nação para atacar a lex mercatoria como uma ‘ficção legal’, como um ‘fantasma’ conjurado por alguns professores especulativos da Sorbonne*”, cujos “*argumentos são baseados na noção do século XIX de unidade entre o estado e o direito: o assim chamado ‘direito anacional’ é impensável!*”<sup>660</sup>: segundo essa perspectiva, a *lex mercatoria* nunca vai se desenvolver como uma ordem legal autêntica “*porque para isso é necessário um vínculo mínimo com o direito nacional*”<sup>661</sup>.

Esse tipo de leitura posiciona, no centro do fenômeno jurídico, uma forma de direito “*por excelência*”, “*autêntico*”, que é criado e aplicado pelo Estado, relegando à periferia desse esquema fenômenos como a *lex mercatoria* e os códigos corporativos, que seriam, no limite, um “*quase-direito*”<sup>662</sup>, incapaz de exercer o poder de vincular a conduta das pessoas e que, por isso, jamais poderiam ser entendidos como parte do fenômeno jurídico propriamente dito. O que o pluralismo jurídico faz, nesse sentido, é precisamente *pluralizar* a ideia de direito, argumentando que a origem no Estado não é um critério adequado para definir aquilo que é o jurídico, porque, em uma sociedade moderna, adotar essa restrição significaria deixar de lado fenômenos normativos que, por mais que impactem na vida das pessoas e tenham características idênticas às do direito estatal, são posicionadas fora desse espectro simplesmente porque não se originaram em uma estrutura política específica. *Lex mercatoria*, códigos corporativos – e, como argumento a seguir, a

---

A Globalização Econômica e sua Arquitetura Jurídica (10 tendências do Direito Contemporâneo). **Revista da Academia Judicial**, v. 1, pp. 41/59, 2010.

<sup>660</sup> “[...] *die sich auf die Souveränität der Nationalstaaten berufen, um die lex mercatoria als ‘Phantom spekulativer Sorbonne-Professoren’ zu attackieren. Ihre Argumente beruhen im Prinzip auf der historisch und rechtsvergleichend reichlich beschränkten Vorstellung der notwendigen Einheit von Recht und Staat. Ein sogenanntes “a-nationales” Recht sei undenkbar*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>661</sup> “[...] *zumindest müsse es eine minimale Verbindung zum nationalen Recht aufweisen*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>662</sup> “[...] *Quasi-Recht* [...]”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

moderação de conteúdo – não seriam fenômenos marginais ao direito, mas indicativos de que aquilo que é o direito mudou, deixando para trás a perspectiva que marginalizava essas construções não-estatais como “*quase-direito*” para uma que reconhece neles expressões de fenômenos tão jurídicos quanto aqueles que surgem do Estado.

O caminho da solução proposta por Teubner, portanto, para o problema de se definir os impactos da globalização para o direito, é acoplar o funcionalismo da Teoria dos Sistemas ao pluralismo jurídico. A estratégia, assim, é construir um *pluralismo funcionalista*<sup>663</sup>, que utiliza as ferramentas da Teoria dos Sistemas para apurar o alcance do pluralismo jurídico, em particular para compreender o funcionamento de estruturas normativas como a *lex mercatoria*, os códigos corporativos – e, a seguir, a moderação de conteúdo.

Isso toma forma em uma perspectiva que organiza o pluralismo jurídico – em particular a visão de que o direito é “*um conjunto de normas sociais conflitantes em um determinado campo*” –, passando a definir o direito como “*uma multiplicidade de processos comunicativos que observam a ação social sob o código binário do lícito / ilícito*”<sup>664</sup>. A Teoria dos Sistemas, nesse sentido, dá corpo teórico ao pluralismo jurídico, apreendendo esse fenômeno jurídico em uma estrutura que define o direito como um âmbito social que produz comunicações capazes de atribuir àquilo que acontece na sociedade a marca do lícito / ilícito. Como nem o direito, nem a aposição desse código são atribuições exclusivas do Estado, essa definição é capaz de abranger tanto aquilo que os Tribunais decidem, quanto o que a *lex mercatoria* produz, sem a necessidade de colocar um à margem do outro: rejeita-se “*categoricamente qualquer posição hierarquicamente superior do direito oficial do Estado*”. As “*múltiplas ordens do pluralismo jurídico sempre produzem expectativas normativas no sentido sociológico*”<sup>665</sup> e, dessa forma, generalizam

---

<sup>663</sup> O termo não é utilizado pelo próprio Teubner; construí como um neologismo para facilitar uma referência a essa amálgama, elaborada pelo autor, entre pluralismo jurídico e funcionalismo-sistêmico estrutural.

<sup>664</sup> “*Danach wird der Rechtspluralismus nicht mehr al seine Menge konfligierender gesellschaftlicher Normen in einem gegebenen Feld definiert, sondern als eine Vielfalt kommunikativer Prozesse, die gesellschaftliches Handeln über den Code Recht/Unrecht beobachten*”. TEUBNER, G. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne in SCHMIDT, E., WEYERS, H. (ed.). **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, 1995, p. 201.

<sup>665</sup> “*Die hier gewählte Ansatz lehnt kategorisch jede hierarchische Überordnung des offiziellen staatlichen Rechtes ab [...]. Die vielfältigen Rechtsordnungen des Rechtspluralismus erzeugen ständig rechtscodierte normative Erwartungen in soziologischen Sinn*”. TEUBNER, G. Die zwei Gesichter des Janus:

congruente essas expectativas pelo mecanismo direito – esteja ele concretizado em consolidações normativas estatais, nos códigos corporativos, na *lex mercatoria* ou, por fim, na moderação de conteúdo.

Essa perspectiva inicial do pluralismo funcionalista é aperfeiçoada em textos posteriores de Teubner, em que ele acopla ao uso do código lícito / ilícito outro elemento que constitui – e diferencia – essas ordens não-estatais como direito. Segundo ele “*as normas promulgadas por essas ‘comunidades transnacionais’ não têm a qualidade jurídica simplesmente porque elas aderem ao código binário lícito/ilícito no julgamento de comportamentos*”<sup>666</sup>; é decisivo, para o direito, que seja identificável um segundo fenômeno: o de que as regras estabelecidas esteja, elas próprias, sujeitas ao próprio escrutínio, de modo que a estrutura normativa seja o resultado dela própria. Um fenômeno normativo não pode ser considerado direito enquanto ele não for reflexamente normativo, isto é, enquanto *as normas criadas não estiverem, elas mesmas, sujeitas às próprias normas*. Uma forma de constatar isso é pela existência de procedimentos institucionalizados de alteração das expectativas normativas: assim como códigos corporativos dispõem quanto a órgãos e procedimentos específicos para que as suas regras sejam alteradas; a *lex mercatoria* detém uma lógica de precedentes, em que mudanças jurisprudenciais devem ser fundamentadas, por exemplo, na mudança das condições presentes à época em que proferida uma decisão vinculante anterior, via, por exemplo, o mecanismo de “*overruling*”. Como afirma Teubner:

“[...] *para que a ordem privada se qualifique como direito genuíno, não é suficiente que as regras pertinentes de comportamento estejam ligadas à noção de lícito ou ilícito. [...] Esse processo reflexivo requer algumas precauções institucionais, em particular, o desenvolvimento de atores ou instâncias, que são responsáveis pelo estabelecimento, modificação, interpretação e implementação da formação de normas primárias*”<sup>667</sup>.

---

Rechtspluralismus in der Spätmoderne in SCHMIDT, E., WEYERS, H. (ed.). **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, 1995, pp. 201/202.

<sup>666</sup> “[...] *besitzen die von diesen „transnationalen Gemeinschaften“ gesetzten Normen nicht schon dann Rechtsqualität, wenn sie nach dem binären Code Recht/Unrecht Verhalten beurteilen*”. TEUBNER, G., KORTH, P. Zwei Arten des Rechtspluralismus: Normkollisionen in der doppelten Fragmentierung der Weltgesellschaft, in KÖTTER, M., SCHUPPERT, G. (Ed.), **Normative Pluralität ordnen**. Baden-Baden: Nomos, 2009 (n. p.).

<sup>667</sup> “[...] *der binäre Code Recht/Unrecht nicht nur auf Verhalten angewendet wird, sondern zugleich auf die Normierungen selbst. Dieser Prozess erfordert gewisse institutionelle Vorkehrungen, besonders die*

Em resumo: Teubner se propõe a resolver o problema de explicar quais são os impactos da globalização no direito. Ele o detecta em dois contextos – o da *lex mercatoria* e dos códigos corporativos – e argumenta que a Teoria dos Sistemas é um bom instrumento para enfrentá-lo, porque tanto o Luhmann de *Legitimação pelo Procedimento*, quanto o de *O Direito da Sociedade*, conceituam o direito como um fenômeno social que não é necessariamente vinculado ao Estado – mesmo que essa conexão apareça, de forma contingente, em algumas análises funcionalistas. Essa proposta toma forma em uma solução em que a Teoria dos Sistemas dá corpo teórico ao pluralismo jurídico, construindo um *pluralismo funcionalista* que conceitua o direito não de acordo com o seu vínculo com o Estado, mas sim como um mecanismo que (i) é capaz de atribuir, a qualquer acontecimento social, a marca do lícito / ilícito; e (ii) tem suas normas reflexivamente sujeitas a elas próprias.

### 2.2.3 Os benefícios da solução de Teubner

Uma vez construído o pluralismo funcionalista como um novo ponto de partida para analisar o que acontece nas empresas – quando instituem códigos corporativos – ou entre os atores comerciais internacionais – que construíram a *lex mercatoria* –, o passo seguinte é utilizar isso em favor da solução do problema da segunda parte desta dissertação. Ou seja: mobilizar essa nova estrutura teórica para explicar o que tem acontecido na moderação de conteúdo das plataformas da internet.

No capítulo anterior, uma primeira aproximação para essa solução foi feita reconstruindo o argumento de Luhmann em a *Legitimação pelo Procedimento*. No texto, a Teoria dos Sistemas é utilizada para explicar o papel que os procedimentos têm na estrutura política moderna – ou seja, no Estado. O argumento de Luhmann, em síntese, é de que, na sua formação, os Estados passaram a tomar decisões por procedimentos que garantem uma incerteza regrada, porque, com isso, as suas decisões se tornam legítimas –

---

*Herausbildung von Verfahrensrollen, die für die Setzung, Modifizierung, Interpretation und Implementation der primären Normierungen zuständig sind*". TEUBNER, G. Codes of Conduct multinationaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. et al (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

ou sejam, tornam-se aceitáveis mesmo que sejam desfavoráveis para as pessoas que são impactadas por elas.

Esse diagnóstico, entretanto, não pode ser simplesmente transportado para as plataformas, porque ele é sustentado em uma relação entre direito e Estado que não está presente no caso da moderação de conteúdo: como a primeira parte deste trabalho demonstrou, não faz sentido entender as decisões tomadas nas plataformas como consequência da atuação dos Estados<sup>668</sup>. Por isso, foi necessário apurar as ferramentas da Teoria dos Sistemas, trazendo-as para a realidade das plataformas, reconstruindo, para isso, o pluralismo funcionalista de Teubner. Agora, de posse das ferramentas adequadas, é possível enfrentar a pergunta seguinte: qual é o papel que os procedimentos têm nessa nova realidade, diferente da posta na *Legitimação pelo Procedimento*, em que Estado e direito estão desvinculados? Como a relação entre procedimento e legitimidade é tensionada – e transformada – pela realidade globalizada do direito? Em particular: que papel têm os procedimentos na realidade das plataformas?

Nesta seção, meu objetivo é responder a essas perguntas. Em específico, exponho a seguir que interpretar situações que antes eram considerados periféricas ao direito propriamente dito – a exemplo da formação da *lex mercatoria* e dos códigos corporativos – como manifestações autênticas do fenômeno jurídico contribui para compreendê-los melhor. Esse ganho analítico acontece, em síntese, porque, ao considerá-los como parte do direito, é possível trazer conceitos antes restritos àquilo que era produzido pelo Estado para essa análise, manejando-os para esclarecer os movimentos pelos quais passam essas novas esferas de manifestação do jurídico. Minha estratégia, nesse sentido, é apresentar esses ganhos analíticos nos exemplos apresentados por Teubner para, a seguir, testá-los na moderação de conteúdo das plataformas da internet, utilizando-os para demonstrar minhas hipóteses a respeito do problema apresentado no capítulo anterior. Quero, assim, expor o diagnóstico de Teubner quanto aos seus exemplos para, em seguida, percorrer esse caminho com as plataformas da internet.

Um primeiro passo para responder a essas perguntas é perceber uma diferença entre o contexto analisado em *Legitimação pelo Procedimento* e o que circunda a *lex mercatoria* e os códigos corporativos. Diferente do Estado, que é uma estrutura política consolidada por um conjunto de fatos históricos que o levaram ao que ele é hoje, esses são fenômenos relativamente recentes, que, assim como a globalização, datam de décadas – e

---

<sup>668</sup> Cf. *supra*, Cap. I.1, 4.



não de séculos – atrás. Isso é algo importante de se perceber, porque tanto a *lex mercatoria* quanto os códigos corporativos não são trazidos por Teubner como o produto de um *fiat* jurídico – isto é, como o resultado pronto de um ato único que faz do não-direito, direito – mas sim como fenômenos em transformação, que têm se aproximado, cada vez mais, desse modelo de direito estruturado pelo pluralismo funcionalista. É justamente na medida em que esse processo se desenvolve, fazendo emergir nessas esferas transnacionais algo que cada vez mais se aproxima do fenômeno jurídico, que surgem as consequências da desvinculação entre direito e Estado, dentre as quais, como argumento a seguir, o tensionamento de relações que existiam na estrutura política tradicional entre procedimento e legitimidade, ressignificando-as.

Assim, Teubner relata, por exemplo, quanto à *lex mercatoria*<sup>669</sup>, que ela satisfaz com facilidade a exigência do pluralismo funcionalista de ser capaz de atribuir, a acontecimentos sociais, um dos lados do código lícito / ilícito. Como o ponto de partida para a instituição desse regime jurídico próprio das relações comerciais internacionais é o contrato, isso não é um problema, porque é inerente a esse mecanismo dispor sobre os direitos e os deveres daqueles que estão envolvidos em arranjos comerciais, estabelecendo, por exemplo, quais são as alternativas que um comprador tem para receber um produto; ou quais são os deveres que o vendedor tem de conservá-lo até ser entregue.

A questão que a *lex mercatoria* verdadeiramente enfrenta diz respeito à segunda característica atribuída por Teubner direito transnacional: a reflexividade. Se o mecanismo contratual operado dentro e fora dos Estados é essencialmente idêntico – no sentido de que, em ambos os casos, ele diz às partes o que é ou não lícito fazer –, o mesmo não pode ser dito quanto àquilo que faz com que “*as regras [...] [sejam] subjugadas a um processo, em que elas são julgadas de acordo com o código jurídico*”, notadamente o “*desenvolvimento de atores ou instâncias, que são responsáveis pelo estabelecimento, modificação, interpretação e implementação da formação de normas primárias*”<sup>670</sup>. É

---

<sup>669</sup> “*Was wir hier beobachten, ist ein sich selbst reproduzierender Rechtsdiskurs globalen Ausmaßes, der seine Grenzen durch Benutzung des binären Codes Recht/Unrecht schließt und sich selbst durch Prozessieren eines Symbols globaler (nicht: nationaler) Geltung reproduziert*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>670</sup> “[...] *institutionelle Vorkehrungen, besonders die Herausbildung von Verfahrensrollen, die für die Setzung, Modifizierung, Interpretation und Implementation der primären Normierungen zuständig sind*”. TEUBNER, G. Codes of Conduct multinationalaler Unternehmen: Unternehmensverfassung jenseits von

nessa segunda característica que a transição do contexto estatal para o não estatal revela o seu verdadeiro desafio, porque esse tipo de questão simplesmente não se colocava antes: um contrato de compra e venda no Brasil tinha as suas normas julgadas de acordo com as legislações produzidas pelo Estado brasileiro, conforme as instituições e processos previamente estabelecidos no país. A reflexividade do direito produzido por um Estado é um dado incontroverso – mencionado, inclusive, como algo que lhe é característico por Luhmann<sup>671</sup>.

Na realidade de um fenômeno jurídico sem Estado, entretanto, como é o caso da *lex mercatoria*, esse pano de fundo jurídico-político simplesmente não existe – a menos a princípio. Por isso, Teubner interpreta o movimento da *lex mercatoria* no sentido de não apenas fazer contratos, mas também instituições, como parte de uma transformação que constrói um fenômeno jurídico extraestatal, que tanto imputa direitos e deveres por expectativas normativas, quanto as avalia de acordo com os seus próprios critérios a partir de um pano de fundo institucional que o próprio direito precisa criar. Esse é o ganho fundamental do pluralismo funcionalista: ele torna a Teoria dos Sistemas capaz de interpretar esse movimento – que, na *lex mercatoria*, surge, por exemplo, com a criação de câmaras arbitrais internacionais, que interpretam, aplicam e julgam a validade das normas<sup>672</sup> – como o nascimento e o desenvolvimento do direito, entendido aqui como um fenômeno comunicativo, e não como o resultado da atuação de uma estrutura política. Como afirma Teubner:

*“[...] mais importante que isso, o contrato auto-referencial [...] externaliza a sua auto-validação [...] referindo as condições de validade e os conflitos futuros para instituições externas ‘não-contratuais’ que são, nada obstante, ‘contratuais’, uma vez que eles são o produto interno do próprio contrato. A mais proeminente dessas instituições externas auto-criadas é a arbitragem que tem que julgar a validade dos contratos, embora a sua própria validade seja baseada na validade que ela deve julgar”<sup>673</sup>.*

---

Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. et al (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>671</sup> LUHMANN, N. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, p. 145.

<sup>672</sup> Para um detalhes quanto a esse funcionamento, cf. BERGER, K. **The Creeping Codification of the New Lex Mercatoria**. Alphen aan den Rijn: Kluwer Law International, 2010, pp. 43 e ss.

<sup>673</sup> “[...] und dies ist der wichtigste Punkt, bedient sich der selbstreferentielle Vertrag der Entparadoxierungstechnik der Externalisierung. Er externalisiert die unvermeidliche Selbst-Validierung des

O mesmo movimento é observado nos códigos corporativos. Novamente, imputar direitos e deveres não é um problema: faz parte da própria lógica de se fazer um código estabelecer aquilo que as pessoas podem ou não fazer – ou seja, dizer o que é lícito ou não. Assim, códigos corporativos têm seções dedicadas a expor os deveres éticos de seus empregados, fixando, por exemplo, que é ilícito receber presentes de outras empresas acima de um determinado valor. Também aqui há um movimento de construção institucional, que, sob a perspectiva do pluralismo funcionalista, é iluminado como a formação de estruturas que submetem essas regras a elas mesmas, notadamente por normas de competência, processos de revisão e criação de órgãos específicos responsáveis por verificar a legalidade e a aplicabilidade dessas regras a determinados casos. Também por trás dos códigos corporativos, portanto, estabelecem-se progressivamente instituições internas, capazes de interpretá-los, aplica-los e revisá-los:

*“É apenas quando existirem arranjos institucionais que sistematicamente subordinam as observações de primeira ordem às observações de segunda ordem que podemos falar de um direito autônomo com ou sem o envolvimento do Estado. Não se deve assumir que há um ‘direito global sem Estado’ apenas baseado sobre instituições não estatais que julgam o comportamento de acordo com um código normativo; essa qualidade só pode ser reconhecida quando processos que observam essas funções judiciais sobre o código binário jurídico forem institucionalizados. Apenas nesse momento os códigos corporativos satisfazem esses pré-requisitos estruturais do direito transnacional fora do direito do Estado”<sup>674</sup>.*

---

*Vertrags, indem er die Beurteilung der Geltungsbedingungen und die Lösung zukünftiger Konflikte externen, nicht-vertraglichen Institutionen zuweist, die aber dennoch "vertraglich" sind insofern, als sie bloße interne Produkte des Vertrags selbst sind".* TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. Rechtshistorisches Journal*, v. 15, 1996, pp. 255/290 (n. p.).

<sup>674</sup> *“Only when institutional arrangements which systematically subordinate the first order observations to the second order observation of a legal code exist, can we speak of autonomous law (with or without the involvement of the state). A “global law without a state” should not yet be assumed upon the basis that non-state institutions judge behaviour pursuant to the normative code, but, rather, that it may be acknowledged only when processes which observe these judicial functions under the binary legal code have been institutionalised. Only then do corporate codes satisfy the structural pre-requisites of a transnational law outside of state law”.* TEUBNER, G. *The Corporate Codes of Multinationals: Company Constitutions Beyond Corporate Governance and Co-Determination*, in NICKEL, R. (ed.). **Conflict of Laws and Laws of Conflict in Europe and Beyond: Patterns of Supranational and Transnational Juridification**. Oxford: Intersentia, 2009, n. p. (cito a versão em inglês e alemão do texto porque elas diferem em alguns trechos).

O ponto é que, no decorrer desse processo de formação de um fenômeno jurídico em nível transnacional, questões antes acomodadas pelo vínculo anteriormente existente entre direito e Estado são tensionadas pela ausência do segundo, tanto retendo características da realidade anterior – em que direito e política estavam associados – quanto os resignificando em determinados aspectos. Esse é justamente o caso das ideias de legitimidade e procedimento, como argumento a seguir.

Sob a perspectiva da legitimidade, ela mantém, por um lado, nesse novo contexto, características do diagnóstico feito por Luhmann em *Legitimação pelo Procedimento*. A sua questão-base ainda se coloca: tanto em contextos estatais, quanto em contextos não estatais, permanece o problema de se explicar porque as pessoas obedeceriam a um fenômeno normativo, seja ele uma regra que consta em uma lei aprovada pelo parlamento, uma cláusula de um contrato internacional ou um item de um código corporativo. Também não importa se no Estado, na *lex mercatoria* ou nos códigos corporativos, o consenso e a coação não são alternativas viáveis para explicar esse fenômeno: assim como não faz sentido pressupor que os países são capazes de posicionar controles que forcem as pessoas a fazerem o que o direito manda a cada momento, também não há cabimento em supor que todos os empregados de uma empresa concordam com o seu código corporativo. Nesse sentido, a legitimidade, como algo que “*é capaz de produzir uma prontidão generalizada para aceitação de suas decisões*”<sup>675</sup>, é uma necessidade que se apresenta tanto em um caso, quanto em outro.

Por outro, a desvinculação entre o jurídico e político, que é característica desse direito globalizado, cria diferenças entre o problema de legitimidade que se tinha antes e o que se apresenta nesse novo contexto. Isso acontece, essencialmente, porque uma parte dos mecanismos que legitimam o direito produzido nos Estados depende da existência dessa estrutura política local. Há várias formas de se perceber isso<sup>676</sup>, mas a que melhor se contextualiza com a abordagem deste trabalho é a que resgata a reconstrução de a *Legitimação pelo Procedimento*, feita no capítulo anterior. Como fica claro pela própria

---

<sup>675</sup> “[...] *eine generalisierte Bereitschaft, inhaltlich noch unbestimmte Entscheidungen innerhalb gewisser Toleranzgrenzen hinzunehmen*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 28; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 30.

<sup>676</sup> Cf., por exemplo, a abordagem que Teubner faz da questão a partir do debate sobre a democratização do direito transnacional em TEUBNER, G. *Quod omnes tangit: Transnationale Verfassungen ohne Demokratie?* **Der Staat**, v. 57, n. 2, 2018, p. 171.

forma como o texto é organizado<sup>677</sup>, a legitimação pelo procedimento das decisões do Estado é, na verdade, uma legitimação pelos *procedimentos*: na estrutura política moderna, coexistem ao menos três âmbitos diferentes de construção de decisões legítimas, localizados nos procedimentos judiciais, legislativos e eleitorais, que se reforçam mutuamente para produzir, ao final, a prontidão generalizada à aceitação de que fala Luhmann.

Isso fica mais evidente quando os resultados desses procedimentos se encontram em situações nas quais alguém é obrigado a aceitar uma decisão que lhe é desfavorável. Assim, uma sentença que condena uma pessoa a pagar um valor a outra não é legítima apenas porque resulta de um procedimento judicial, mas porque aplica uma regra que foi o resultado de um procedimento legislativo que, por sua vez, foi construído por parlamentares escolhidos por um procedimento eleitoral. Há uma sobreposição recíproca de procedimentos legitimadores, que se reforçam mutuamente, retroalimentando-se e, com isso, blindando a decisão em um nível cada vez mais elevado, aumentando o ônus de quem a critica a patamares cada vez maiores. Assim, o devedor irresignado precisa se voltar não apenas à ação institucional do juiz, que lhe mandou pagar, mas também à dos legisladores que promulgaram a regra, e à dos eleitores que os escolheram. Ele precisa enfrentar, em síntese, um bloco de procedimentos legitimadores que, entrecruzados entre si, resulta da amálgama que existe entre direito e política no contexto local.

No direito transnacional, entretanto, uma das partes dessa amálgama não existe. Não há correspondente político do Estado para o direito global e, por isso parte dos mecanismos legitimadores<sup>678</sup> – em particular, os procedimentos legislativos e eleitorais – não se colocam como alternativas para solucionar esse problema. É difícil pensar em uma eleição entre os empregados de uma empresa para decidir quem será o seu presidente, assim como, na *lex mercatoria*, as decisões que definem os direitos e deveres que irão

---

<sup>677</sup> Cf. *supra*, Cap. II.4, 2.3.1 e 2.3.2, que replicam a existência de seções à parte em *Legitimação pelo Procedimento* para os procedimentos judiciais, eleitorais e legislativos.

<sup>678</sup> É importante ressaltar que, desde as suas elaborações mais iniciais, a ideia de legitimidade não esteve necessariamente vinculada à de Estado – o que permite transpô-la, como faço aqui, para um contexto não-político, como o do direito global. Nessa direção, Luhmann fala em uma “*politização do tema ‘legitimação’*” (“*Politisierung des Themas ‘Legitimation’*”), referindo-se à circunstância de que, na origem, a ideia de legitimação não é inerentemente política, mas assim se torna quando tomada no contexto estatal. LUHMANN, N. **Soziologische Aufklärung 4: Beiträge zur funktionalen Differenzierung der Gesellschaft**. Opladen: Westdeutscher Verlag, 1987, p. 361.

constar nos contratos não decorrem de uma escolha legislativa, mas sim da composição dos interesses econômicos das partes. Isso não significa que alternativas procedimentais não sejam viáveis, aliviando, em certa medida, a pressão que recai sobre os mecanismos que restam: é possível, por exemplo, que o código corporativo seja elaborado de forma colaborativa com os empregados, conferindo-lhes voz para propor dispositivos, ou mesmo que eles elejam um representante para fazer isso. Trata-se, entretanto, mais de uma solução episódica do que um método de estruturar o poder político, como ocorre nos Estados. Por isso, o problema da legitimidade se apresenta de uma forma diferente – e mais delicada – nesse direito global, na medida em que os instrumentos legitimadores procedimentais são menos densos – isto é, têm um reforço mútuo atenuado pela ausência do Estado e de seus mecanismos de escolha eleitoral e legislativa. É por isso que Teubner fala em uma “*situação de auto-legitimidade*”<sup>679</sup>: sem um referencial político em que o problema da legitimidade possa ser compartilhado pelo direito, resta ao procedimento judicial – que se faz presente, em formatos paralelos ao do Estado, tanto na *lex mercatoria*, via arbitragem, quanto nos códigos corporativos, por processos de apuração das condutas de um empregado, por exemplo – o papel de protagonista na criação dessa prontidão generalizada à aceitação.

Esse déficit de mecanismos de legitimação – motivado pela ausência de uma estrutura política correspondente à do Estado no direito global – traduz-se, no outro lado da relação estabelecida por Luhmann, em uma ressignificação da função do procedimento. De um lado, e como no caso da legitimidade, o diagnóstico de a *Legitimação pelo Procedimento* permanece válido: procedimentos “*judiciais*” viabilizam a tomada de decisão por uma estrutura de incerteza regrada, que legitima o seu resultado, seja ele a decisão de um juiz de direito ou a de um tribunal arbitral. Assim como a definição do que é direito, na Teoria dos Sistemas, não depende dessa amálgama entre direito e Estado, a relação entre legitimidade e procedimento também não a pressupõe – mesmo que, contingentemente, ela tenha sido aplicada à estrutura política que Luhmann tinha à mão à época, ou seja, o próprio Estado. Da mesma forma como as partes de um processo judicial são levadas à aceitar a decisão final, mesmo que uma delas acabe condenada, a arbitragem

---

<sup>679</sup> “*Wir erleben hier eine ‘selbst-legitimierende’ Situation [...]*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

internacional tem regras procedimentais – semelhantes às do Estado<sup>680</sup> – que trazem as partes do contrato a um ambiente controlado, em que a controvérsia é trabalhada a um nível que, no final, uma delas se veja na contingência de aceitar o laudo arbitral, mesmo que ele lhe seja desfavorável.

O mecanismo, nesse sentido, é essencialmente idêntico, mesmo que alguns detalhes desse novo contexto impactem na forma como ele funciona: no âmbito transnacional, por exemplo, há um potencial de exposição do procedimento a contextos em que existe uma variabilidade de valores muito maior que a do Estado. Tratarei da moderação de conteúdo com mais detalhes na seção que segue, mas ela é um bom exemplo disso: as decisões quanto àquilo que as pessoas podem ou não dizer na internet exigem desse procedimento a capacidade de lidar com os valores que perpassam virtualmente qualquer comunidade localizada no mundo – e não mais aquela abrangida por um território<sup>681</sup>. Assim, se em contextos estatais – localmente delimitadas, que reúnem pessoas que frequentemente compartilham alguns valores em comum – o consenso já não era um caminho viável para se produzir aceitação, nos contextos transnacionais ele sequer se apresenta como uma possibilidade no horizonte. Isso escancara e torna ainda mais agudo o diagnóstico feito por Luhmann em *Legitimação pelo Procedimento*, no sentido de que, no limite, o que é possível se obter na sociedade moderna é, nas palavras de Teubner, um “*dissenso organizado*”, isto é, não a “*esperança inocente pela única e melhor solução*” – tal como supõem aqueles que associavam o consenso e a verdade ao procedimento, como seus resultados inevitáveis –, mas, no máximo, “*a confrontação de visões de mundo alternativas*”, permitida por “*procedimentos decisórios que funcionem bem e criem espaço*”

---

<sup>680</sup> Cf., como referência, as ICC Rules of Arbitration, base sobre a qual os julgamentos da Câmara Internacional do Comércio são realizados, em particular os artigos 11 a 30, que estabelecem uma forma de julgamento semelhante à do devido processo legal, viabilizando mecanismos que garantem a neutralidade do julgador (art. 11, 1) e a participação das partes no julgamento (art. 26). ICC. **ICC Rules of Arbitration**. Disponível em <<https://iccwbo.org/dispute-resolution-services/arbitration/rules-of-arbitration/>>, acesso em 8.9.2019.

<sup>681</sup> Cf., nesse sentido, a dificuldade mencionada por Teubner em se definir o “*Omnes*” (“*todos*”) na aplicação atual da frase “*Quod omnes tangit, ab omnibus tractari et approbari debet*” (“*Aquilo que a todos toca, deve por todos ser aprovado*”) em TEUBNER, G. Quod omnes tangit: Transnationale Verfassungen ohne Demokratie? *Der Staat*, v. 57, n. 2, 2018, pp. 179 e ss.

para a ‘possibilidade do dissenso como uma pré-condição de uma seletividade independente distribuída na sociedade’<sup>682</sup>.

De outro lado, entretanto, a diferença que existe entre o contexto local – em que há Estado – e o transnacional – em que não há Estado – não apenas torna mais aguda a função do procedimento na sociedade moderna, mas a ressignifica, conferindo-lhe um novo papel. Isso porque a ausência de uma estrutura política em um nível transnacional correspondente ao do direito não significa que esse novo fenômeno jurídico ocorre no vácuo. Pelo contrário: ele age em outros âmbitos, como o econômico e o científico, regulamentando-os na medida em que também eles não mais se restringem aos limites do território político. Isso faz renascer um problema que estava de certa forma composto pela evolução da relação entre direito e política: o da autonomia do direito em relação àquilo que ele regulamenta. Enquanto restrito ao âmbito local, o direito tinha à mão, no que diz respeito à sua relação com a política, mecanismos que lhe permitiam desenvolver a sua própria lógica decisória, impedindo, por exemplo, que o modo como dividido o poder em órgão de representação legislativa interferisse na decisão de um juiz. Uma vez que o fenômeno jurídico se desvencilha dessas amarras locais, desvinculando-se do político – para se vincular a outros âmbitos –, ele revisita esse problema que parecia inicialmente solucionado, vendo-se na necessidade de adaptar as suas ferramentas para que continue exercendo a sua função com autonomia.

Esse contexto confere uma nova importância ao procedimento – e, em particular, ao procedimento “*judicial*”. Para entender como isso acontece, é preciso compreender como Teubner traz mais um conceito do direito “*tradicional*” para o direito transnacional: o de constituição. Como aconteceu com o conceito de direito, o termo “*constituição*” é identificado como um fenômeno essencialmente político-jurídico e, por isso, local: ele é utilizado para designar um arranjo de normas superior às demais, que organiza e limita o poder do Estado ao qual vinculado – fala-se, por exemplo, na Constituição Federal do Brasil. A aposta de Teubner, nesse sentido, é semelhante à feita com o pluralismo funcionalista: se o direito não é, necessariamente, um fenômeno local, então também faria sentido pensar a constituição – uma ideia central ao direito – não como um acontecimento exclusivo dos Estados, mas também como um conceito que, se

---

<sup>682</sup> “*Im transnationalen Raum ist also Demokratie darauf auszurichten, dass effektiv funktionierende Entscheidungsprozeduren Raum schaffen für die ‘Möglichkeit von Dissens als Voraussetzung eigenständiger in der Gesellschaft verteilter Selektivität’*”. TEUBNER, G. Quod omnes tangit: Transnationale Verfassungen ohne Demokratie? **Der Staat**, v. 57, n. 2, 2018, p. 181.



adequadamente ampliado, poderia ajudar a explicar as transformações pelas quais passa o direito transnacional. Ele questiona: “*é possível fazer frutificar a tradição do constitucionalismo do Estado-nação, para que ele consiga abranger os fenômenos da privatização e globalização?*”<sup>683</sup>; ou, dito de outra forma, em que medida esse “*nacionalismo metodológico*”<sup>684</sup> – que restringe tanto direito, quanto constituição, ao fenômeno do Estado-nação – sustenta-se.

Teubner aventa, nesse sentido, a possibilidade de uma transição, em que a ideia de *constituição* – associada a arranjos normativos particulares, detectáveis junto ao Estado – fosse enfocada a partir do *constitucionalismo* – um mecanismo social não necessariamente vinculado a essa instância política local. Sua ideia, portanto, é propor uma sociologia constitucional, baseada na afirmação de que “*é simplesmente anacrônico restringir o alcance do fenômeno que nós vemos como essencialmente constitucional às constituições dos Estado-nação*”<sup>685</sup>; sua tese, nesse sentido, é de que “*nós estamos testemunhando a emergência de uma multiplicidade de constituições civis além do Estado-nação*”<sup>686</sup>. Para ele, aquilo que, na percepção anterior, era o conjunto de elementos aos quais se resumia o fenômeno – as constituições dos países – torna-se, na sociologia constitucional, apenas um dos casos possíveis pelos quais o constitucionalismo se manifesta, abrindo caminho para que ele também fosse identificado em ambientes transnacionais, tais como os que deram a luz aos códigos corporativos e à *lex mercatoria*.

---

<sup>683</sup> “*In other words, can we make the tradition of nation-state constitutionalism fruitful and redesign it in order to cope with the phenomena of privatisation and globalisation?*”. TEUBNER, G. *Fragmented Foundations: Societal Constitutionalism Beyond the Nation State*, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, 2010, p. 328.

<sup>684</sup> “[...] *methodologischen Nationalismus* [...]”. TEUBNER, G. *Transnationaler Verfassungspluralismus: Neun Variationen über ein Thema von David Sciulli*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 76, 2016, p. 3.

<sup>685</sup> “[...] *it is simply anachronistic to restrict the range of phenomena which we see as essentially constitutional to the constitutions of nation states, as widely remains the case amongst mainstream proponents of constitutional research*”. TEUBNER, G. *Societal Constitutionalism: Nine Variations on a Theme by David Sciulli*, in BLOKKER, P., THORNHILL, C. **Sociological Constitutionalism**. Cambridge: Cambridge University Press, 2017, pp. 313/340 (n. p.; o trecho é exclusivo da versão em inglês do texto).

<sup>686</sup> “[...] *we are witnessing the emergence of a multiplicity of civil constitutions beyond the nation state*”. TEUBNER, G. *Fragmented Foundations: Societal Constitutionalism Beyond the Nation State*, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, 2010, p. 329.

Mas o que particulariza isso? O que é, afinal de contas, esse constitucionalismo capaz de abranger tanto o fenômeno local, quanto o que se manifesta no nível do direito transnacional? A chave para responder a essas perguntas é aplicar a metodologia funcionalista à ideia de constituição. Isso significa deixar de entendê-la como um “*fenômeno normativo monista de regras jurídicas mais importantes*”<sup>687</sup> e passar a pensar em *qual é a função que uma constituição exerce* – seja em um âmbito local, típico dos Estados, seja no âmbito do direito global, arena em que fenômenos como a *lex mercatoria* acontecem.

A resposta de Teubner, em síntese, é de que o constitucionalismo é um *mecanismo de autonomização de âmbitos comunicativos*, que molda a relação que o direito tem com outros nichos da sociedade – sejam eles políticos, como é o caso dos Estados, sejam eles econômicos, como as empresas. Sua proposta é que, à medida que o direito se constrói em âmbitos locais e transnacionais – tanto atribuindo a acontecimentos a marca do lícito e ilícito, quanto estabelecendo mecanismos de reflexividade, em que essa normatividade se sujeita a si mesma –, ele dá um passo além, e *cria instrumentos de autonomização própria*, pelos quais ele se torna independente daquilo que regula, imunizando-o<sup>688</sup>. Ele deixa de ser um anexo, por exemplo, das relações comerciais da *lex mercatoria*, ou da atividade econômica da empresa que resolve ter um código corporativo, e passa a ter um campo de atuação própria. Em particular, “*a situação só se torna constitucionalizada quando normas das normas – isto é, normas secundárias – prescrevem como a identificação, formação, emenda e regulação de competências para a expedição e delegação de normas primárias devem ocorrer*”<sup>689</sup>. O direito se constitucionaliza, nesse

---

<sup>687</sup> “[...] *as a monistic normative phenomenon of higher legal rules* [...]”. TEUBNER, G. Constitutional Drift: On the Spontaneous Co-evolution of Social ‘Ideas’ and Legal ‘Structures’, in DOWDLE, M., WILKINSON, M. (Ed.) **Constitutionalism beyond Liberalism**. Cambridge: Cambridge University Press, pp. 79/95, 2017 (n. p.).

<sup>688</sup> Sobre esse caráter de imunização que é inerente ao constitucionalismo e o papel que o procedimento desempenha nesse contexto, cf. GUERRA FILHO, W. Luhmann and Derrida: Immunology and Autopoiesis in LA COUR, A., PHILIPPOPOULOS-MIHALOPOULOS, A. **Luhmann Observed**. Houndmills: Palgrave Macmillan, 2013, pp. 233 e ss.

<sup>689</sup> “*The situation only really becomes constitutionalized when norms of norms – that is, secondary norms – prescribe how the identification, setting, amendment and regulation of competences for the issuing and delegating of primary norms are to occur*”. TEUBNER, G. Constitutional Drift: On the Spontaneous Co-evolution of Social ‘Ideas’ and Legal ‘Structures’, in DOWDLE, M., WILKINSON, M. (Ed.) **Constitutionalism beyond Liberalism**. Cambridge: Cambridge University Press, pp. 79/95, 2017 (n. p.). Cf., no mesmo

sentido, ao se acoplar a outros âmbitos do social, de modo a tanto *regulá-los* quanto a *se manter independente em relação a eles*, garantindo uma dupla autonomia: tanto de si próprio, quanto daquilo que é regulado por ele.

Nesse sentido, Teubner:

*“É sempre um fenômeno dual: um vínculo entre esses processos intrinsecamente autônomos. De uma perspectiva jurídica, a constitucionalização envolve a produção de normas secundárias que são peculiarmente entrelaçadas com os entendimentos do sistema social. De uma perspectiva social, a constitucionalização envolve a produção de normas epistêmicas secundárias que são, elas mesmas, informadas pelas normas jurídicas reconhecidas pelo sistema. Apenas quando ambas as condições operarem juntos faz sentido, em termos tanto da sociologia quanto da doutrina jurídica, falar em ‘elementos’ de uma constituição política, de uma constituição econômica, de uma constituição dos sistemas da educação e ciência ou de uma constituição digital da internet”<sup>690</sup>.*

Um exemplo pode ajudar a esclarecer esse ponto. Códigos corporativos surgem, como disse acima, em estruturas econômicas: eles são fenômenos que emergem dentro de empresas que atuam transnacionalmente e que constroem, no seu interior, complexos normativos que estabelecem princípios, condutas permitidas ou vedadas e órgãos de controle e deliberação. Para que essas estruturas sejam consideradas parte de um

---

sentido, TEUBNER, G. The Corporate Codes of Multinationals: Company Constitutions Beyond Corporate Governance and Co-Determination, in NICKEL, R. (ed.). **Conflict of Laws and Laws of Conflict in Europe and Beyond: Patterns of Supranational and Transnational Juridification**. Oxford: Intersentia, 2009 (n. p.); “Therefore, an autonomous, non-state, non-political, civic constitutionalisation of multinationals takes place if reflexive social processes, which concern the relationship of the multinational in its various environs, are interwoven with reflexive legal processes”); TEUBNER, G. *Unitas Multiplex: Das Konzernrecht in der neuen Dezentralität der Unternehmensgruppen*. **Zeitschrift für Unternehmens- und Gesellschaftsrecht**, v. 20, n. 2, 1991, pp. 211 e ss.

<sup>690</sup> “It is always a dual phenomenon: a linking of these two innately autonomous processes. From the legal perspective, constitutionalization involves the production of secondary legal norms that are peculiarly interwoven with the fundamental understandings of the social system. From the social perspective, constitutionalization involves the production of secondary epistemic norms that are themselves critically informed by the legal norms recognized by the system. Only when both these conditions operate together does it make sense, in terms of both legal sociology and legal doctrine, to speak of the ‘elements’ of a political constitution, of an economic constitution, of a constitution of the education and science system, or of the digital constitution of the Internet”. TEUBNER, G. *Constitutional Drift: On the Spontaneous Co-evolution of Social ‘Ideas’ and Legal ‘Structures’*, in DOWDLE, M., WILKINSON, M. (Ed.) **Constitutionalism beyond Liberalism**. Cambridge: Cambridge University Press, pp. 79/95, 2017 (n. p.).

fenômeno jurídico transnacional, o pluralismo funcionalista exige delas duas características: primeiro, a capacidade de atribuir a acontecimentos sociais a marca do lícito e do ilícito; e, segundo, a reflexividade, isto é, o fato de que as suas normas estejam sujeitas a elas mesmas, no sentido de que são elas que determinam, por exemplo, qual é o procedimento para que uma regra seja modificada. Esses códigos, entretanto, estão acoplados a uma estrutura econômica – a empresa –, o que significa que eles convivem constantemente com a possibilidade de que essa lógica seja sobreposta à de sua construção jurídica: por mais que faça parte da missão da empresa incluir pessoas com deficiência, por exemplo, é possível imaginar um cenário em que se decide demitir todos os empregados que tenham algum tipo de impedimento a longo prazo, sob o argumento de que produzem menos que os demais. Nesse contexto, a constitucionalização do direito é uma fase adiante do pluralismo funcionalista, em que o direito não é apenas construído em um ambiente transnacional, mas também detém mecanismos que o autonomizam em relação às lógicas não-jurídicas, fazendo, por exemplo que a demissão desses empregados possa ser reavaliada de acordo com os critérios desse direito interno – que pode chegar à conclusão, por exemplo, que cumprir essa determinação contraria as regras que regulam a admissão de pessoas na empresa.

É possível aplicar o mesmo raciocínio na *lex mercatoria* – em que a constitucionalização do direito parece mais desenvolvida: também ela é atrelada a um fenômeno essencialmente econômico – as trocas comerciais transnacionais –, e também aqui se coloca a questão da autonomização do direito, em particular porque as instituições construídas para tomar decisões jurídicas são mantidas por esses atores. É interessante perceber, nesse sentido, que existem barreiras institucionais – como, por exemplo, o procedimento adotado para se julgar um caso em uma câmara de arbitragem, a divisão igual dos custos, a distância mantida entre árbitros e partes, o dever que os julgadores têm de revelar se tiveram alguma relação recente com as empresas participantes<sup>691</sup> – que garantem que a decisão final seja o resultado de uma lógica jurídica – ou seja, da aplicação do código lícito / ilícito ao modo como as partes se comportaram – e não econômica. Não ganha quem paga mais; ganha quem tinha razão – no sentido jurídico da palavra –, porque

---

<sup>691</sup> Cf., a esse respeito, o art. 11, 2, das Regras da ICC, que estabelece um “*dever de disclosure*” do árbitro apontado para compor um painel arbitral. ICC. ICC Rules of Arbitration, disponível em <<https://iccwbo.org/dispute-resolution-services/arbitration/rules-of-arbitration/>>, acesso em 8.9.2019.

a esfera jurídica é compreendida e mantida como algo autônomo em relação àquilo que ela regula. Como afirma Teubner quanto à *lex mercatoria*:

*“É um direito que cresce e muda de acordo com as exigências das transações e organizações econômicas globais (Braeckmans, 1986). Isso a faz extremamente vulnerável aos interesses e pressões dos processos econômicos. Como não há isolamento institucional de sua quase-legislação de sua quase-jurisdição, a relativa autonomia e independência que historicamente as ordens jurídicas nacionais foram capazes de atingir provavelmente irá permanecer como algo desconhecido”*<sup>692</sup>.

É dentro desse passo além, tomado, em algumas circunstâncias, pela constitucionalização do direito em sua relação com aquilo que ele regula, que o procedimento é ressignificado, tomando para si um outro papel. Como disse acima, o movimento de constitucionalização pode ser entendido no âmbito transnacional como uma forma de o direito tentar resguardar a sua autonomia: se, localmente, essa questão era equacionada pelas constituições político-jurídicas, na medida em que o fenômeno jurídico se desvincula do Estado, ele se acopla em novas realidades – como as econômicas – e, nesse sentido, se vê vulnerável não mais ao despota que se recusa a reconhecer os direitos dos governados, mas sim pelo presidente da empresa, que tenta impor, sobre essa lógica jurídica transnacional, a lógica econômica; ou pela parte da arbitragem internacional, que tenta subornar um dos seus juízes. Como afirma Teubner, *“na relação entre o direito e a política, a tradição do Estado nação estabeleceu um modelo que ainda precisa ser desenvolvido na relação entre o direito e outros setores sociais”*<sup>693</sup>. Para fazer frente a

---

<sup>692</sup> *“Es ist ein Rechtsgebiet, das gemäß den Bedürfnissen globaler ökonomischer Transaktionen wächst und sich wandelt (Braeckmans, 1986). Gegenüber der Interessen- und Machtausübung von ökonomischen Akteuren ist es daher extrem anfällig. Aufgrund der nur geringen Isolierung seiner Quasi-Gesetzgebung und Quasi-Rechtsprechung werden die relative Autonomie und Unabhängigkeit, die nationale Rechtsordnungen in der Vergangenheit erreicht haben, hier wohl nicht erreicht werden”*. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>693</sup> *“Im Verhältnis des Rechts zur Politik hat die nationalstaatliche Tradition das Modell vorgelegt, das im Verhältnis des Rechts zu anderen Teilsystemen in analoger Manier erst noch zu entwickeln ist”*. TEUBNER, G. Privatregimes: Neo-Spontanes Recht und duale Sozialverfassungen in der Weltgesellschaft, in SIMON, D., WEISS, M. (Ed.) **Zur Autonomie des Individuums: Liber Amicorum Spiros Simitis**. Baden-Baden: Nomos, 2000, pp. 437/453 (n. p.).

isso, o direito adapta o que Teubner chama de “*reserva de instituições constitucionais*”<sup>694</sup>: mecanismos que, desenvolvidos no âmbito local, “*tiveram sucesso em efetivamente limitar a totalização do poder político*”<sup>695</sup> – e, por isso, são transportados para a esfera transnacional, sob a possibilidade de que possam atuar nessa revisitação que o direito faz ao problema da sua autonomia, para, à semelhança da constituição política – relativa ao vínculo direito/Estado –, viabilizar a criação de instituições econômicas e científicas, por exemplo.

Esses mecanismos são organizados, em síntese, em dois elementos considerados típicos da constitucionalização: “*disposições normativas sobre o estabelecimento e exercício da tomada de decisão (regras procedimentais e organizacionais), de um lado; e definições sobre liberdades individuais e autonomias sociais (direitos fundamentais), de outro*”<sup>696</sup>. A constitucionalização, nesse sentido, traduz o problema da autonomia em dois tipos de soluções: uma procedimental, relacionada ao estabelecimento de normas quanto a como decisões são tomadas, imunizando-as de influências externas ao direito; e outra, material, relativa a direitos fundamentais.

A primeira – que é a que pertine a esse momento da dissertação e que Teubner menciona como tão ou mais relevante que a outra, porque “*a proteção horizontal dos direitos constitucionais deve ser consistentemente transmutada em organização e processo*”<sup>697</sup> –, expõe um papel dos procedimentos à parte do detectado em *Legitimação pelo Procedimento*, e que é trazido pelo direito transnacional à esfera da *lex mercatoria* e

---

<sup>694</sup> “[...] *Reservoir an verfassungsrechtlichen Institutionen*”. TEUBNER, G. Transnationaler Verfassungsppluralismus: Neun Variationen über ein Thema von David Sciulli. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 76, 2016, p. 8.

<sup>695</sup> “*Vorbild sind sie besonders deshalb, weil ihnen die paradoxe Leistung geglückt ist, in Machtprozessen immanent angelegte Totalisierung der politischen Macht gerade mithilfe politischer Macht wirksam zum begrenzen*”. TEUBNER, G. Transnationaler Verfassungsppluralismus: Neun Variationen über ein Thema von David Sciulli. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 76, 2016, p. 8.

<sup>696</sup> “*Here, in such diverse contexts, we find typical elements of a constitution: provisions on the establishment and exercise of decision making (organisational and procedural rules) on the one hand and definitions of individual freedoms and societal autonomies (fundamental rights) on the other*”. TEUBNER, G. *Fragmented Foundations: Societal Constitutionalism Beyond the Nation State*, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, 2010, p. 333.

<sup>697</sup> “[...] *the horizontal protection of constitutional rights must be consistently transmuted into organization and process*”. TEUBNER, G. Horizontal Effects of Constitutional Rights in the Internet: A Legal Case on the Digital Constitution. **Italian Law Journal**, v. 3, n. 1, 2017, p. 201.

dos códigos corporativos. Em síntese, procedimentos não são apenas instrumentos de produção de legitimidade, mas também são *mecanismos que atuam autonomizando o direito, porque constituem esferas de atuação exclusiva da lógica jurídica, imunizando-a em relação àquilo que ela regula*. Eles têm, nesse sentido, um papel importante no processo de constitucionalização do direito transnacional, porque estabelecem meios pelos quais ele possa tomar decisões por ele mesmo – isto é, de forma independente em relação, por exemplo, à economia. Procedimentos, nesse sentido, *garantem ao fenômeno jurídico transnacional que ele tome as suas próprias decisões – antes que alguém as tome por ele*.

Essa nova função atribuída pelo pluralismo funcionalista – em sua versão mais apurada, que reinterpreta a constitucionalização como um fenômeno também desvinculado do Estado – aos procedimentos faz sentido tanto sob a perspectiva dos problemas que emergem para o direito, na medida em que ele se desvincula do âmbito local e, conseqüentemente, da política; quanto de como os procedimentos funcionam no direito, algo destrinchado por Luhmann em a *Legitimação pelo Procedimento*. Como reconstruí acima, não há dúvidas de que esse direito transnacional enfrenta um problema de autonomia: na medida em que ele se desacopla da política e passa a regular novos âmbitos da sociedade em nível transnacional, ele deixa para trás os mecanismos de autonomização que constituiu com o Estado. Como argumentei acima, entretanto, isso não significa abandonar qualquer pretensão de autonomia da esfera do jurídico, mas sim reconquistá-la de outra forma: há uma espécie de memória institucional quanto a esses mecanismos de autonomização que é resgatada pelo direito nessa nova realidade – mas que precisa ser nela novamente construído, na forma de um procedimento de constitucionalização.

Nessa direção, os mecanismos que o procedimento coloca à disposição do direito apresentam-se como uma resposta possível para esse problema, porque, como a *Legitimação pelo Procedimento* demonstrou, eles imunizam a decisão jurídica de influências não jurídicas criando, no seu decorrer, um mundo procedimental à parte. Em particular, as duas primeiras características que mencionei acima quanto aos procedimentos judiciais – e que são refletidas nos demais – têm exatamente como escopo construir esse espaço procedimental à parte, que resguarda a atuação do direito de influências que lhe sejam externas: refiro-me, especificamente, primeiro, à capacidade do processo de “*construir fronteiras que criam diferenças entre o que acontece dentro e fora*

*deles*”<sup>698</sup> – isto é, diferenciar-se pela elaboração de papéis específicos e filtros que delimitam aquilo que faz ou não parte do procedimento, construindo “*para si próprio uma esfera de sentido, de forma a que os processos seletivos de elaboração de informação do meio ambiente possam ser orientados por regras e decisões próprias do sistema*”<sup>699</sup>; e segundo, à sua “*autonomia*”<sup>700</sup> – que o faz se desenvolver por critérios definidos por sua estrutura e atores, “*e não de acordo com uma lógica de algo que lhe é externo*”, de modo que “*a ação dentro do sistema seja, em parte, orientada segundo critérios próprios de seleção*”<sup>701</sup>. Se, em a *Legitimação pelo Procedimento*, essas características são funcionalmente relacionadas a um papel garantidor da incerteza do resultado do procedimento – porque impedem que ele seja externamente determinado –, algo importante porque faz com que as partes sejam estimuladas a participarem da formação da decisão, levando-as a aceitarem no final; no pluralismo funcionalista de Teubner essas características desdobram um papel diferente dos procedimentos, que se soma ao anterior, no sentido de garantir a autonomia da decisão jurídica em relação àquilo que lhe externo, notadamente as influências dos âmbitos regulados pelo direito transnacional. É essa a constitucionalização procedimental à qual ele se refere – em que processos decisórios são convertidos em instrumentos de defesa da autonomia do direito.

É importante mencionar que essa função não nasce com o direito transnacional, mas é exposta por ele, na medida em que, no âmbito local, a independência entre direito e política estava consolidada no formato constitucional clássico, o que deixava esse papel menos latente. Com a desvinculação de ambos – e a abertura de um flanco de intervenção, no direito, pelos âmbitos que atuam *transnacionalmente*, ela volta à tona com

---

<sup>698</sup> Cf. *supra*, Cap. II.4, 2.3.1.

<sup>699</sup> “[...] *eine Sinnsphäre für sich zu konstruieren, so daß selektive Prozesse der Verarbeitung von Umweltinformationen durch systemeigene Regeln und Entscheidungen gesteuert werden können, daß also Strukturen und Ereignisse der Umwelt nicht automatisch auch im System gelten, sondern erst nach Filterung von Informationen anerkannt werden*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 59; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 53.

<sup>700</sup> Cf. *supra*, Cap. II.4, 2.3.1.

<sup>701</sup> “[...] *daß ein Gebäude beschafft wird, Richter angestellt und besoldet, Kompetenzen festgelegt und durchgesetzt, Gesetze erlassen und bekanntgegeben, Tatsachen gewußt und mitgeteilt werden, schließt nicht aus, daß im System das Verhalten zum Teil an eigenen selektiven Kriterien orientiert wird*”. LUHMANN, N. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, p. 69; LUHMANN, N. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, p. 61.



toda a sua importância, que deve ser considerada, portanto, de modo paralelo ao papel legitimador construída por Luhmann em a *Legitimação pelo Procedimento*.

Novamente, um exemplo pode ajudar a compreender esse argumento. No caso da *lex mercatoria*, não há dúvida de que o procedimento “*judicial*”, utilizado por câmaras arbitrais à semelhança dos tribunais de justiça produz legitimidade. Na exata medida em que autor e réu são conduzidos pelo processo judicial de uma pluralidade inicial de possibilidades para uma decisão a ser aceita no final, as partes de um contrato se submetem à arbitragem e, após participarem desse procedimento decisório, se veem na contingência de não recusar o seu resultado, mesmo que ele lhe seja desfavorável. A *lex mercatoria*, entretanto, tem um problema adicional: ela não é vinculada a uma estrutura política, mas sim a um conjunto de estruturas econômicas, que ela regula. Isso gera o seguinte problema: como garantir que a decisão produzida no final da arbitragem não foi influenciada pela posição de uma das partes nessa estrutura econômica? Por exemplo, pode ser interessante para um árbitro adotar uma posição favorável à empresa economicamente mais poderosa, para que, em casos futuros, ele seja novamente chamado a atuar. O que garante, afinal, que ele agirá com autonomia, decidindo apenas de acordo com os próprios critérios jurídicos? É aqui que entra em cena o procedimento: ele produz uma decisão legítima ao final, mas também garante que essa decisão é o resultado do próprio procedimento. Ele cria um mundo – procedimental – à parte do mundo, que restringe aquilo que pode impactar na decisão final, porque posiciona filtros ao que pode fazer ou não parte do procedimento. Por mais que uma das empresas seja muito mais poderosa que a outra, isso só poderá impactar no procedimento na medida em que ela traduza essa condição em argumentos que tenham impacto jurídico – isto é, sejam capazes de convencer o juiz –, por exemplo, contratando um escritório de advogados de qualidade.

Uma mesma situação exemplificativa poderia ser pensada quanto aos códigos corporativos. Uma vez que uma empresa decida estabelecer um código corporativo, suponha que ela fixe normas quanto àquilo que é lícito fazer, mas deixe de criar procedimentos – no sentido funcionalista da palavra – para decidir, por exemplo, se alguém violou ou não esses padrões. Toda vez que há uma suspeita que isso aconteceu, quem define se o empregado permanece ou não na empresa é um de seus diretores, que adota, para isso, os seus próprios critérios – como produtividade, capacidade de gerar lucros, etc.

O que acontece, nessa situação, quando são estabelecidos procedimentos decisórios que viabilizam ao interessado, por exemplo, participar dessa decisão,

posicionando a sua defesa, obtendo transparência a respeito do que está acontecendo e ter a possibilidade de recorrer? De um lado, isso torna a decisão mais legítima, no sentido de que o resultado desse novo procedimento é algo mais aceitável. De outro, o procedimento atua blindando os critérios jurídicos que regulam o código corporativo, dos códigos do regulado, isto é, dos critérios econômicos utilizados monocraticamente pelo diretor. Ele age, nesse sentido, como o construtor de um regime à parte daquele que é regulado, que faz com que, uma vez desenvolvido em um certo nível esse processo de autonomização, não se fala mais de um direito subjugado – ou seja, de um código corporativo submetido aos imperativos econômicos –, mas na *colisão entre regimes*<sup>702</sup>, sejam eles, por exemplo, o jurídico e o político, o jurídico e o econômico e o jurídico e o científico. Confere-se corpo ao direito, ao nível em que ele pode fazer frente àquilo que ele regula, por um processo de “*auto-organização de mútua constituição dos atos jurídicos e das estruturas jurídicas*”<sup>703</sup>.

Com isso, o procedimento emerge desse movimento de apuração das ferramentas da Teoria dos Sistemas – algo, repito, necessário para que elas se tornem aplicáveis ao fenômeno jurídico transnacional – com uma dupla função. Por um lado – e como já se via na *Legitimação pelo Procedimento* – procedimentos produzem legitimidade, no sentido de que constroem ambientes controlados por estruturas capazes de gerar uma incerteza regrada e, assim, uma prontidão generalizada à aceitação de suas decisões. Por outro – e como o diagnóstico de Teubner expõe – procedimentos são parte dessa memória institucional que o direito traz ao âmbito transnacional consigo, servindo, nesse âmbito, como um mecanismo de constitucionalização procedimental do fenômeno jurídico, garantindo-lhe autonomia em relação àquilo que ele passa a regular.

### 3. POR UMA NOVA PERSPECTIVA QUANTO À MODERAÇÃO DE CONTEÚDO

#### 3.1 Introdução

---

<sup>702</sup> Refiro-me à ideia de colisão de regimes desenvolvida por Teubner em FISCHER-LESCANO, A., TEUBNER, G. Regime-Collisions: The Vain Search for Legal Unity in the Fragmentation of Global Law. **Michigan Journal of International Law**, v. 25, n. 4, pp. 999/1.046, 2004 (em particular, o trecho “*Autonomous ‘Private’ Legal Regimes*”, pp. 1.009/1.012).

<sup>703</sup> “[...] *vielmehr geht es um einen selbstorganisierten Prozeß wechselseitiger Kontituierung von Rechtsakten und Rechtsstrukturen*”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

Este último capítulo – dedicado a oferecer uma resposta ao problema da segunda parte da dissertação – é dividido em duas partes. Na primeira, minha proposta foi aprimorar as ferramentas utilizadas pela Teoria dos Sistemas para explicar o papel que os procedimentos têm na sociedade moderna. Como o diagnóstico de a *Legitimação pelo Procedimento* – segundo o qual os procedimentos são mecanismos que atribuem legitimidade às decisões, porque as tornam aceitáveis para as pessoas – é construído sobre um contexto em que o direito se manifesta como um fenômeno normativo atrelado ao Estado – e isso não acontece no caso da moderação de conteúdo –, foi necessário trazer o funcionalismo estrutural para essa nova realidade do fenômeno jurídico. Assim, e a partir da reinterpretação de Gunther Teubner do trabalho de Luhmann, foi possível utilizar a Teoria dos Sistemas para dar corpo ao pluralismo jurídico, tornando o funcionalismo estrutural capaz de lidar com essa nova realidade jurídica. Em particular, essa ferramenta tanto expôs que a relação entre legitimidade e procedimento permanece válida nesse novo contexto jurídico, quanto trouxe à tona outro papel do procedimento, que é o de conferir autonomia ao direito em relação ao âmbito social que ele regula.

Na segunda parte deste capítulo, minha intenção é utilizar essas ferramentas para entender como funciona e se transforma a moderação de conteúdo das plataformas da internet. Em particular, proponho, a seguir, uma abordagem que testa as hipóteses que apresentei para resolver o problema do Capítulo 3.

Naquele momento, após apresentar uma esquematização para abordar a questão da opacidade da moderação de conteúdo, a proposta foi avançar em problemas mais complexos, algo permitido justamente por esse movimento de esclarecimento quanto a como esses mecanismos funcionam. Em síntese, argumentei que é possível dividir os juízos feitos pelas plataformas na moderação de conteúdo em dois tipos: os de ordem material, que orientam o conteúdo que é permitido publicar nesses ambientes – e que se aproximavam, justificadamente, de uma perspectiva estatal-democrática de proteção à liberdade de expressão; e os de ordem procedimental, que indicam como decisões sobre o conteúdo publicado devem ser tomadas – e que, ao contrário dos anteriores, eram marcadamente arbitrários, isto é, eram tudo menos inspirados no modo como o Estado toma decisões, porque caracterizados pela opacidade, ausência de participação dos usuários e irrecurribilidade.

Em seguida, sustentei, a partir de evidências colhidas nas três principais plataformas da internet – Facebook, Twitter e YouTube –, que esses juízos procedimentais passam por uma transformação, porque cada vez mais a moderação de conteúdo toma de

empréstimo características do devido processo legal dos Estados, mudando a forma como as decisões são tomadas para garantir, aos seus usuários, mecanismos de transparência, oportunidades de participação e instrumentos de recorribilidade. Mais surpreendente que isso – uma vez que o modelo anterior estava alinhado com as necessidades das plataformas, de exercer poder sobre aquilo que é dito no espaço que administram, atuando com liberdade suficiente para excluir conteúdos de seus usuários – é o fato de que essa transformação é inexplicável. A implantação de mecanismos de transparência, participação e recorribilidade na moderação de conteúdo é algo que não faz sentido, porque as plataformas não estão obrigadas a fazê-lo; porque a lógica pela qual operam é atendida pelo modelo arbitrário anterior; e, principalmente, porque as plataformas não se deixam levar facilmente por exigências de grupos organizados da sociedade civil – como é interessante para elas narrar –, o que sugere que há uma espécie de *tradeoff* que, aceito por essas empresas, permanece oculto a quem observa, externamente, o movimento. Por que as plataformas decidiram mudar o seu modelo decisório, transformando-o em um baseado naquele que obriga os Estados? Que razões existem para que elas, depois de desenvolverem um sistema próprio de moderação de conteúdo, agora se empenhem em mudá-lo? Por que – e, em que medida – essas empresas estão abrindo mão do seu poder?

No decorrer deste capítulo e do anterior, elaborei duas hipóteses para responder a essa pergunta, que se baseiam no papel que o procedimento tem na estrutura social, tal como detectado por uma perspectiva funcionalista do problema – sustentada pela Teoria dos Sistemas de Niklas Luhmann, cuja aplicação justifiquei antes<sup>704</sup>. Em síntese, a primeira hipótese é de que as plataformas da internet têm se aproximado do modelo procedimental dos Estados, porque ele é capaz de produzir legitimidade, isto é, porque ele viabiliza que as decisões sejam aceitas com mais facilidade pelos seus usuários, mesmo que elas lhe sejam desfavoráveis. Aqui, o *tradeoff* oculto é revelado por uma relação entre *procedimento* e *legitimidade*. A segunda hipótese é de que é possível ler essa mudança como parte de uma transformação mais ampla que, viabilizada pela internet, permite perceber não só a formação de novos polos decisórios não-estatais – como a moderação de conteúdo –, mas também a constitucionalização dessa forma de decidir – pela estruturação de formatos procedimentais responsáveis por autonomizar essas decisões das estruturas corporativas em que estão inseridas, algo que auxilia a compreender, com mais precisão,

---

<sup>704</sup> Cf. *supra*, Cap. II.3, 3.2.

os passos seguintes desse modelo regulatório. Aqui, o *tradeoff* oculto é revelado por uma relação entre *procedimento* e *autonomia*.

No que segue, testo essas hipóteses, jungindo a esquematização que esclarece como a moderação de conteúdo funciona – e que desenvolvi na primeira parte da dissertação – com as ferramentas da Teoria dos Sistemas – nos termos em que reconstruídas aqui – parte no Capítulo 4, parte no início deste capítulo. Faço isso em três seções: nas duas primeiras (3.2 e 3.3), ofereço argumentos que respectivamente sustentam as minhas duas hipóteses; na terceira (3.4), desenvolvo três propostas de continuidade deste trabalho, a partir dos seus resultados.

### 3.2 O procedimento como legitimação

O primeiro diagnóstico do funcionalismo estrutural quanto ao papel desempenhado pelos procedimentos pode ser sintetizado da seguinte forma: procedimentos são úteis porque eles são mecanismos que produzem decisões legítimas. Como esclareci antes<sup>705</sup>, essa definição tem dois pressupostos conceituais: primeiro, o termo “*procedimento*” designa um mecanismo que administra, por meio de estruturas, a incerteza da decisão de uma forma regrada. Ele define, portanto, uma *forma de produzir decisões*, em que uma variabilidade inicial é convertida em uma decisão final, mantendo-se, entre um e outro ponto, a incerteza quanto a qual será o resultado. Segundo, “*legitimidade*” é a prontidão generalizada das pessoas a aceitarem uma decisão, mesmo que ela lhe seja desfavorável. Portanto, os procedimentos são mecanismos que constroem uma situação de incerteza regrada quanto a uma decisão, e, por isso, geram uma prontidão generalizada para que ela seja aceita.

Como argumentei na seção anterior, esse diagnóstico, inicialmente extraído de um âmbito local, em que direito e Estado estão vinculados, é trazido para um contexto em que esse vínculo não existe mais – o do direito transnacional – pelo pluralismo funcionalista de Teubner. Segundo ele – e a partir de exemplos com a *lex mercatoria* e os códigos corporativos –, faz sentido conceituar o direito como um fenômeno comunicativo que, primeiro, atribui a acontecimentos sociais um dos lados da característica lícito / ilícito; e, segundo, age reflexivamente, ou seja, reproduz pela existência de instituições que fazem com que os seus critérios normativos sejam balizados por eles mesmos. Por isso, para que

---

<sup>705</sup> Cf. *supra*, Cap. II.4, 2.2.2.

a relação entre procedimento e legitimidade seja aplicável à moderação de conteúdo na internet, é preciso perguntar antes: faz sentido interpretar aquilo que as plataformas fazem quando definem o que alguém pode dizer como uma expressão do direito transnacional, no sentido que Teubner atribui a esse termo?

Uma primeira indicação de que a resposta para essa pergunta é afirmativa é retirada do próprio Teubner. Isso porque, no desenvolvimento do pluralismo funcionalista – e mesmo que de forma mais incidental que a *lex mercatoria* e os códigos corporativos –, ele se refere à internet como um campo cujas normatizações são explicáveis a partir do direito transnacional. Segundo ele, “*Hoje, os mais proeminentes regimes jurídicos são a lex mercatoria da economia internacional e a lex digitalis da internet*”<sup>706</sup>. Essa abordagem fica mais clara em um texto em particular, em que Teubner discute se há um direito de acesso ao mundo digital a partir de um caso judicial concreto, em que pessoas processaram um provedor da internet porque ele bloqueou o acesso de seus usuários a páginas na internet “*em relação às quais ele considerava que o risco de ser responsabilizado criminalmente ou civilmente era muito alto*”<sup>707</sup>.

O artigo é importante para a discussão da aplicabilidade do pluralismo funcionalista à internet porque ele evidencia características da era digital que tanto a distanciam da esfera de atuação do Estado, quanto expõem as compatibilidades de suas normatizações com essa nova forma de ver o direito. Elas podem ser resumidas em três: a digitalização (“*digitisation*”), a privatização e a globalização. O argumento de Teubner, em resumo, é de que a presença dessas características na realidade da internet mostra que, como aconteceu com o comércio internacional – via *lex mercatoria* – e as empresas multinacionais – via os códigos corporativos – faz mais sentido interpretar as normatizações que surgem na internet como uma parte da transnacionalização do direito, do que tentar vinculá-la à estrutura política local – algo que já tinha ficado evidente no

---

<sup>706</sup> “*Today, the most prominent private legal regimes are the lex mercatoria of the international economy and the lex digitalis of the Internet*”. TEUBNER, G. *Fragmented Foundations: Societal Constitutionalism Beyond the Nation State*, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, 2010, p. 332.

<sup>707</sup> “*Darauf sperrt der provider mit elektronischen Mitteln alle die websites, für die er das Prozessrisiko strafrechtlicher oder zivilrechtlicher Natur als zu hoch einschätzt*”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.). Trata-se do caso LICRA v. Yahoo!, julgado pelo *Tribunal de grande instance de Paris* (TGI). Cf. a íntegra do julgamento em LICRA ET UEJF CONTRE YAHOO! Inc. et YAHOO FRANCE, disponível em <<http://www.lapres.net/ya2011.html>>, acesso em 8.9.2019.

fracasso da tentativa de analisar a moderação de conteúdo a partir das categorias jurisprudenciais da Suprema Corte dos Estados Unidos<sup>708</sup>.

Primeiro, e quanto à digitalização, a tecnologia associada à internet viabiliza que as informações que antes eram dispostas em suportes físicos – como um jornal ou uma revista – sejam traduzidas em uma linguagem compreensível por um computador. O cerne dessa transformação não está na mudança do local onde essa informação passa a ser armazenada – da estante em que guardado o livro para um disco rígido – mas sim em como esse armazenamento impacta na forma como a informação circula. Isso porque, à medida em que ela é convertida e produzida em meios digitais, perdem força os instrumentos políticos capazes de restringir a sua circulação a um espaço territorial determinado – como uma fronteira física. Se um documento em papel é transformado em um arquivo .pdf, ele perde as restrições de circulação que tinha antes – a necessidade de enviá-lo pelo correio, a possibilidade ser interceptado no meio do caminho, etc. – em favor de uma realidade digital que, ao menos em situações normais<sup>709</sup>, não toma conhecimento da realidade política: uma vez anexada a um e-mail, essa manifestação antes física pode ser enviada para o outro lado do mundo com a mesma facilidade em que um papel é entregue a um colega de trabalho localizado na sala vizinha. Aquilo que era um *discrímen* importante para formas tradicionais de se transmitir informações – a distância física e suas barreiras políticas – não são mais relevantes. Por características que lhe são inerentes, a comunicação é, nesse sentido, naturalmente transnacional.

Segundo, é característico desse movimento a tendência à privatização, no sentido de que, com a internet, ganham protagonismo atores que Teubner denomina “*não-estatais*”<sup>710</sup>. Como a digitalização viabilizou a transmissão das informações por uma esfera própria, com divisões que não coincidem necessariamente com as que compartimentam politicamente o mundo, os atores privados – que, por sua vez, também não estão

---

<sup>708</sup> Cf. *supra*, Cap. I.2, 3.2.1.

<sup>709</sup> Existem exceções a isso, como expõem os trabalhos a respeito das decisões judiciais que bloqueiam determinadas plataformas em um país como uma tentativa de fazê-los colaborar em processos judiciais; cf., a esse respeito, ABREU, J. Disrupting the disruptive: making sense of app blocking in Brazil. **Internet Policy Review**, v. 7, n. 3, disponível em <<https://policyreview.info/articles/analysis/disrupting-disruptive-making-sense-app-blocking-brazil>>.

<sup>710</sup> “[...] *nicht-staatliche Akteure* [...]”. TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

necessariamente vinculados a um território – mostram-se mais vocacionados a tomar o protagonismo desse novo âmbito do que os Estados. Assim, a internet engloba às “organizações internacionais, empresas multinacionais, sindicatos internacionais, grupos de interesses e organizações não governamentais, como participantes da construção global de decisões”<sup>711</sup> atores privados que, de forma compatível com a característica transnacional da comunicação digital, também são transnacionais: além das plataformas, que abordo com mais detalhes a seguir, Teubner menciona, por exemplo, o mecanismo arbitral da ICANN, instituição que resolve disputas quanto aos nomes de domínios na internet<sup>712</sup> definindo, por exemplo, quais empresas têm o direito de utilizar um determinado endereço na rede.

Terceiro, é característico da era digital – e, por isso, capaz de fazer com que aquilo que produz seja inserido nessa realidade do direito transnacional – a globalização. Sob uma perspectiva genérica, a globalização é um “*processo policêntrico em que simultaneamente diferentes áreas da vida rompem os seus limites regionais e cada uma constitui setores globais autônomos para si mesmas*”<sup>713</sup>: direito, ciência, economia e outros âmbitos sociais que antes se viam restritas ao âmbito local, delimitando-se, como a política faz, a uma fronteira, ganham o mundo, no sentido de que constroem uma esfera de atuação global. É exatamente isso que acontece, por exemplo, com a *lex mercatoria* e os códigos corporativos, em que, respectivamente, os instrumentos regulatórios do comércio e a forma como empresas disciplinam a conduta de seus funcionários extravasam as fronteiras locais

---

<sup>711</sup> “Die Akteure sind zum einen internationale Organisationen, multinationale Unternehmen, internationale Gewerkschaften, Interessenverbände und Nichtregierungsorganisationen als Beteiligte an weltpolitischen Entscheidungsprozessen [...]”. TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>712</sup> “[...] als interner Selbstorganisationsprozeß des Internet durch Institutionen wie ICANN und deren Schiedsgerichte, Standardisierungsorganisationen wie das World Wide Web Consortium oder die Internet Engineering Task Force und digitale Bürgerbewegungen oder in einem hybriden Prozeß herausbildet, ist eine ganz andere Frage”. TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>713</sup> “Vielmehr ist Globalisierung ein polyzentrischer Prozeß, in dem unterschiedliche Lebensbereiche ihre regionalen Schranken durchbrechen und je für sich autonome Globalsektoren konstituieren”. TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).



do político, e perpassam horizontalmente as estruturas estatais para construir regulamentações que são relativamente indiferentes à localização territorial. Mesmo que uma compra e venda tenha se iniciado no Brasil e terminado nos Estados Unidos, é possível a uma corte arbitral aplicar preceitos acordados entre as partes e que nada tem que ver com as estruturas legislativas de um ou de outro país; o código corporativo de uma empresa alemã aplicado à conduta de um empregado no Chile pode não ter relação nem com um país, nem com o outro. O “*estar em um país*” deixa de ser relevante para definir uma transação comercial, a verdade de uma afirmação científica ou a validade de uma norma.

Não é difícil ver, nesse sentido, como a internet, pela construção de uma comunicação digital e pelo protagonismo dos atores privados, é parte neste processo de globalização: com ela, os limites regionais da transmissão de informação são ignorados, e nesse sentido, constrói-se o que Teubner chama de uma “*vila global*”<sup>714</sup>, ou seja, um ambiente à parte da realidade política regionalizada. Esse aspecto globalizatório, que congrega as duas primeiras características, é particularmente perceptível quando alguém navega pela internet em ambientes virtuais construídos e hospedados em diferentes países, sem que isso, via de regra<sup>715</sup>, traduza-se em algo de relevante para o usuário. Ninguém apresenta um passaporte para sair de um site brasileiro e ir a um hospedado na Inglaterra; não faz sentido exigir um visto para acessar um site proveniente dos Estados Unidos. Isso não significa que a internet não tenha as suas próprias fronteiras: alguém que queira, por exemplo, acessar o Facebook precisa criar um perfil na plataforma, identificável por meio de uma senha que lhe garante acesso àquele ambiente restrito. O ponto, nessa direção, não é que a internet tenha abolido as fronteiras, mas sim que *ela tenha construído as suas próprias*, de modo que, ao não fazê-las coincidir com as políticas, ela oferece a principal evidência de que é parte desse processo globalizatório que extravasa os limites locais, construindo uma realidade que deve ser interpretada de uma maneira independente.

Essas três características da internet – digitalização, privatização e globalização – são agregadas por Teubner em um diagnóstico que *detecta na realidade criada pela internet e, principalmente, nas normatizações que a regem uma expressão do pluralismo jurídico*. Teubner traça, nesse sentido, quanto àquilo que ocorre na internet, um

---

<sup>714</sup> “[...] *global villages* [...]”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>715</sup> Cf. *supra*, nota n. 692.

caminho paralelo ao da *lex mercatoria* e ao dos códigos corporativos: primeiro, ele identifica as limitações que um conceito de direito – e, como exponho adiante, de constituição – vinculado à “*arquitetura do Estado-nação*”<sup>716</sup> têm para abranger fenômenos normativos associados à internet. Assim como no caso da *lex mercatoria*, em que predominava a perspectiva estatal que não conferia valor normativo a esse “*fantasma conjurado por alguns professores especuladores da Sorbonne*”<sup>717</sup>, no caso da internet ele menciona os constitucionalistas que não conferem credibilidade às “*fantasias jurídicas empolgadas de professores de Harvard*”<sup>718</sup>. Segundo, Teubner oferece uma resposta para essa dificuldade ampliando o conceito de direito para a sua versão funcionalista do pluralismo jurídico, segundo a qual “*as regras jurídicas não são produzidas apenas dentro da regulação de conflitos por cortes oficiais nacionais ou internacionais, mas também dentro de estruturas não-políticas de resolução de disputas sociais, organizações internacionais, cortes de arbitragem, corpos de mediação, comitês éticos e sistemas de tratados*”<sup>719</sup>.

Esse paralelismo de caminhos entre *lex mercatoria*, códigos corporativos e internet – ao ponto em que o texto *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*, no qual discutido o caso do provedor de internet<sup>720</sup>, ser praticamente uma versão aplicada à internet de *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*, em que introduzido o pluralismo funcionalista como

---

<sup>716</sup> “[...] *Architektur des Nationalstaats* [...]”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>717</sup> “[...] *die lex mercatoria als "Phantom spekulativer Sorbonne-Professoren" zu attackieren*”. TEUBNER, G. *Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus*. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>718</sup> “[...] *Rechtspantasien überspannter Harvard-Professoren* [...]”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>719</sup> “*Und Rechtsnormproduktion aus Konfliktlösung vollzieht sich nicht nur innerhalb von nationalen und internationalen Gerichten, sondern auch innerhalb von nichtpolitischen gesellschaftlichen Konfliktlösungsinstanzen, internationalen Organisationen, Schiedsgerichten, Mediationsinstanzen, Ethikkommissionen, Vertragsregimes*”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>720</sup> TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie*. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28

uma forma de resolver o problema da *lex mercatoria*<sup>721</sup>, não significa identidade entre ambos: como Teubner posiciona, há especificidades entre esses âmbitos que particularizam tanto eles, quanto o direito que produzem<sup>722</sup>. O importante, entretanto, para o objetivo desta seção, é que o pluralismo funcionalista que explica a *lex mercatoria* e os códigos corporativos, é o mesmo que se aplica àquilo que acontece na internet, o que abre um flanco para entender em que medida aquilo que as plataformas – um ator não-estatal que surge justamente nesse mundo digital construído pela internet – fazem via moderação de conteúdo – um tipo de regulação transnacional da liberdade de expressão – pode ser considerado direito – e, dessa forma, ter os seus movimentos esclarecidos por essa versão da Teoria dos Sistemas.

O primeiro passo para fazer isso é abordar, com o pluralismo funcionalista, as plataformas da internet. Como defini antes<sup>723</sup>, plataformas da internet são serviços e páginas na internet que (i) “*hospedam, organizam e circulam, para os seus usuários, o conteúdo que eles compartilham e as interações que fazem*”; (ii) “*sem ter produzido ou comissionado (o grosso) desse conteúdo*”; (iii) “[fazendo isso] *com base em uma infraestrutura que [...] realiza o processamento de dados para prestar serviços ao usuário, fazer anúncios e lucrar*”<sup>724</sup>. No que segue, argumento, a partir do que reuni na primeira parte do trabalho, que elas concretizam as três características que Teubner utiliza para vincular a internet à realidade transnacional, o que permite compreender os fenômenos normativos que acontecem ali – e, em particular, a moderação de conteúdo – como uma expressão do direito, no sentido do pluralismo funcionalista.

Primeiro, quanto à digitalização, Teubner afirma que, na era da internet, a informação é armazenada e circulada em um formato compreensível por computadores, o

---

<sup>721</sup> TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, 1996, pp. 255/290 (n. p.).

<sup>722</sup> Cf., por exemplo, as considerações de Teubner quanto à facilidade de implementação das auto-regulações da internet pelos atores privados, uma vez que eles podem interferir diretamente no código, mitigando os problemas de *enforcement* encontrados em âmbitos como a *lex mercatoria* (“*Das selbstgemachte Recht des Internet profitiert nicht nur von den Schwierigkeiten der beiden anderen Regulierungsformen, sondern besonders von den technischen Vorteilen, welche die Architektur des Code für eine hocheffiziente Regulierung bietet*”) em TEUBNER, G. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, 2003, pp. 1/28 (n. p.).

<sup>723</sup> Cf. *supra*, Cap. I.2, 2.2.

<sup>724</sup> Cf. *supra*, nota n.º 238.

que a torna compartilhável em uma escala global e fundamentalmente independente dos limites políticos. As plataformas, nesse sentido, são o exemplo perfeito de como isso funciona na prática, porque a digitalização da comunicação é algo tão inerente a elas que faz parte do modo como são definidas, isto é, como um serviço que “*hospeda, organiza e circula*” aquilo que as pessoas produzem na internet. Nesse sentido, o modo como as publicações são feitas nas plataformas, algo que descrevi com mais detalhes acima<sup>725</sup>, viabiliza que os seus usuários já se manifestem em um formato digital compreensível pela infraestrutura informática utilizada para construir aquele espaço – o que é, no fundo, um pressuposto para que ele seja utilizado –, de modo, portanto, compartilhável em escala global. Assim, quando uma publicação é feita no Facebook, uma mensagem é veiculada no Twitter ou um vídeo é enviado ao YouTube, eles imediatamente se tornam acessíveis a qualquer pessoa com acesso à internet<sup>726</sup>, mesmo que esteja localizada em um país diferente, de modo que, via de regra, a fronteira política não se traduz em uma fronteira comunicativa.

As plataformas, nesse sentido, são instrumentos de manejo da comunicação digital, que Teubner posiciona como algo central ao vínculo entre internet e transnacionalidade. Por isso, o ambiente que elas constroem e administram também é transnacional, o que também se reflete na forma como ele é regulamentado: os Padrões da Comunidade do Facebook, as Diretrizes da Comunidade do YouTube e as Regras do Twitter são válidos para toda a plataforma, o que significa que eles perpassam horizontalmente as regulamentações verticais dos Estados em que essas empresas atuam. Comunicação digitalizada – e, por isso, transnacional – significa, no caso das plataformas, tanto um ambiente transnacional, quanto uma regulamentação transnacional. Isso não se traduz, entretanto, em uma situação na qual os limites políticos são absolutamente irrelevantes para esse tipo de comunicação. Como o caso das publicações que negam o Holocausto – que tratei no primeiro capítulo<sup>727</sup> – expõe, é possível que essas fronteiras

---

<sup>725</sup> Cf. *supra*, Cap. I.2, 3.1.

<sup>726</sup> Existem exceções a essa regra, notadamente as vinculadas às configurações de privacidade da plataforma – um usuário pode restringir o acesso àquilo que publica às pessoas que aceitou em seu perfil. O ponto, entretanto, é que essas fronteiras não são barreiras políticas, mas sim construções da própria plataforma, que o faz de acordo com critérios que ela define. Dizer que não é possível visualizar tudo que os usuários compartilham na plataforma não significa, necessariamente, uma reintrodução das divisões políticas no mundo digital.

<sup>727</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

sejam refletidas nas plataformas: o Facebook, nesse caso, impede que determinados conteúdos sejam exibidos na Alemanha, porque isso é contrário à legislação do país. Essa é, entretanto, a exceção, que deve ser assim tratada: via de regra, o mesmo conteúdo é acessível em todos os países. Mais do que isso: como o próprio exemplo da negação do Holocausto mostra, na maioria das vezes em que há legislações que a proíbem, essa restrição não é levada em consideração ao Facebook, que optou por bloqueá-lo apenas em quatro países, nos quais a plataforma identifica um risco para as suas atividades.

Segundo, Teubner vincula a internet ao pluralismo funcionalista porque identifica nela o protagonismo de atores não-estatais, que, à semelhança dos agentes econômicos da *lex mercatoria* e das empresas dos códigos corporativos, atuam perpassando horizontalmente os limites políticos. Novamente, é exatamente isso que acontece no caso das plataformas: ao longo desta dissertação, forneci uma lista das mais importantes<sup>728</sup>, escolhendo, por razões que expus anteriormente<sup>729</sup>, focar-me em três delas – o Facebook, Twitter e YouTube. Em todos esses casos, as plataformas são atores essencialmente privados: são empresas transnacionais vinculadas à tecnologia que mantêm e administram espaços de manifestação na internet, razão pela qual, inclusive, argumentei que um trabalho que aborde a regulação da liberdade de expressão na era digital não poderia se concentrar apenas naquilo que o Estado faz<sup>730</sup>. No caso das plataformas, a razão exposta por Teubner para vincular a internet e o pluralismo funcionalista se evidencia: como a digitalização da comunicação a faz transpor as fronteiras do político, a administração de ambientes destinados à manifestação das pessoas deve ter como base, também, uma estrutura capaz de ultrapassá-las. As empresas, nesse sentido, são estruturas econômicas que se dissociam de sua origem – o Facebook não é pensado como um embaixador dos Estados Unidos, mesmo que tenha surgido na Califórnia – e, por isso, que conseguem operar a administração dessas informações independentemente dos limites políticos.

Terceiro, o vínculo entre internet e globalização também é concretizado nas plataformas. Sob uma perspectiva geral, Teubner expõe essa característica como a criação de uma “*vila global*”, à parte das separações estatais, em que comunidades transnacionais são construídas por âmbitos comunicativos como o direito, a ciência e a economia, que

---

<sup>728</sup> Cf. *supra*, Cap. I.2, 2.2.

<sup>729</sup> Cf. *supra*, nota n.º 239.

<sup>730</sup> Cf. *supra*, Cap. I.1, 3.3.3.

extravassam os limites políticos locais. Nesse caso, talvez as plataformas sejam a manifestação mais literal desse modo de perceber a globalização na internet. Como expus acima<sup>731</sup>, empresas como o Facebook viabilizam que as pessoas se manifestam criando uma interface capaz de tornar realidade a promessa da Web 2.0: se, no início, a internet reproduzia a assimetria entre um produtor de informações e um receptor, com o tempo essa diferença foi amenizada pela possibilidade de que qualquer pessoa publique algo na rede. Para que isso aconteça, entretanto, foram necessários novos intermediários – as plataformas da internet –, responsáveis por criar interfaces que viabilizassem que qualquer pessoa fosse capaz de utilizar a infraestrutura digital internet para se manifestar, mesmo que sem conhecimento técnico a esse respeito. Quando a solução ficou pronta, ela tomou a forma justamente dessa “*vila global*” – à qual Teubner talvez se referia de forma figurada, mas que se concretiza, no caso das plataformas, exatamente como uma comunidade em que as pessoas se manifestam em um espaço que não é dividido de acordo com as repartições políticas dos Estados. É como se, nessa esfera digital independente, viabilizada pelo caráter digitalizado da comunicação e administrada por agentes privados, as pessoas pudessem assumir uma nova personalidade, destacada da que têm na realidade, compondo um mundo à parte do mundo, no qual as suas manifestações não sofressem as restrições que têm para se propagar no mundo politicamente dividido e, por isso, transcendessem as fronteiras do local, sem tomar conhecimento, por exemplo, das fronteiras que separam um país do outro.

Como exposto no exemplo que forneci antecipadamente na perspectiva geral, isso não significa que as plataformas não tenham fronteiras – mas apenas que elas são construídas de maneira independente às dos Estados. Uma outra forma de ver isso são as restrições à privacidade, impostas por elas ou por seus usuários àquilo que é veiculado nesse novo ambiente. Assim, esse espaço pode não ser repartido pelas divisões políticas entre os Estados, mas, por exemplo, as comunidades que são criadas no seu interior podem ser construídas com níveis diferentes de privacidade, que fazem com que as suas informações estejam disponíveis apenas a quem as acompanha. Ainda, há mecanismos tanto no Twitter quanto no Facebook que restringem a visualização das informações compartilhadas pelo usuário às pessoas que foram aceitas, respectivamente como “*amigo*” ou “*seguidor*” por aquele que as publica; e, ainda, formas pelas quais um usuário pode escolher não ver determinados conteúdos.

---

<sup>731</sup> Cf. *supra*, Cap. I.2, 2.2.

Se a digitalização, a privatização e a globalização são as características da internet que permitem vinculá-la à realidade transnacional, na qual é aplicável o pluralismo funcionalista; e essas três características se apresentam de forma tão evidente no caso das plataformas da internet, então as conclusões acima abrem caminho para que as ferramentas que Teubner utilizou para trazer o direito, como estabelecido pela Teoria dos Sistemas, do âmbito local, para o âmbito transnacional, sejam também aplicadas nas plataformas. Em particular, isso torna viável perguntar, nesse contexto, em que medida as normatizações produzidas pelas plataformas são manifestações do direito, aqui entendido pelas chaves do pluralismo jurídico – que vê esse fenômeno como algo independente do Estado – e do funcionalismo de Teubner – que dá corpo teórico ao pluralismo jurídico. Dito de outra forma: a moderação de conteúdo é direito?

Como reconstruí na seção anterior, o pluralismo funcionalista de Teubner enfrenta o desafio de fazer frente a fenômenos normativos desvinculados ao Estado que, a exemplo da *lex mercatoria* e dos códigos corporativos, são típicos da globalização. Ele faz isso traduzindo esse problema na questão de entender os impactos da globalização para o direito, oferecendo, como resposta, o surgimento de “*objetos normativos não identificados*”<sup>732</sup>, isto é, de fenômenos semelhantes ao direito, mas que não são classificados como jurídicos porque não são produzidos pelo Estado. Ele propõe, então, um novo conceito de direito, capaz de abrangê-los: para Teubner, o direito é (i) “*uma multiplicidade de processos comunicativos que observam a ação social sob o código binário do lícito / ilícito*”<sup>733</sup> (critério do lícito/ilícito), em que (ii) as expectativas normativas que resultam desses processos comunicativos são aplicáveis a si mesmas, fazendo das normas sujeitas a elas mesmas (critério da reflexividade). A pergunta do parágrafo anterior, portanto, pode ser traduzida da seguinte forma: os critérios (i) e (ii) estão presentes na moderação de conteúdo?

Sustento que, para o primeiro critério, a resposta é definitivamente sim. Na primeira parte da dissertação, tentei resolver o problema da opacidade da moderação de conteúdo a partir de uma pesquisa que fornecesse um esquema capaz de explicar como ela

---

<sup>732</sup> ‘UNOs’ (*unidentified normative objects*)”. TEUBNER, G. *Transnationales Recht: Legitimation durch horizontale Grundrechtswirkung* in VIELLECHNER, L. **Transnationalisierung des Rechts**. Velbrück: Weilerswist, 2013, prefácio (n. p.).

<sup>733</sup> “[...] *eine Vielfalt kommunikativer Prozesse, die gesellschaftliches Handeln über den Code Recht/Unrecht beobachten*”. TEUBNER, G. *Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne* in SCHMIDT, E., WEYERS, H. (ed.). **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, 1995, p. 201.

funciona. Conceitualmente, isso se traduziu em definir a moderação de conteúdo em uma *perspectiva geral* – segundo a qual ela é o instrumento utilizado pelas plataformas da internet para exercer algum tipo de controle sobre as manifestações de seus usuários – e em uma *perspectiva restrita* – para a qual moderar conteúdo é a prática de estabelecer e aplicar um conjunto de regras que determinam aquilo que os usuários de uma plataforma podem ou não dizer<sup>734</sup>.

Essa distinção agora se revela importante, porque ela esquematiza com precisão que há práticas de moderação de conteúdo, tomada em sentido geral, que não são fenômenos jurídicos no sentido estabelecido por Teubner. Assim, quando a plataforma escolhe o seu nicho de atuação, definindo o tipo de mídia em que irá concentrar as suas atividades, ela exerce um tipo de controle sobre a manifestação de seus usuários – no YouTube, por exemplo, não são aceitas publicações de imagens, mas apenas de vídeos –, balizando-a, entretanto, sem se utilizar de um instrumento normativo. Moderar conteúdo em sentido amplo não é sinônimo de direito.

A situação é diferente quando a moderação de conteúdo é tomada em sentido estrito – que o tema desta dissertação: aqui, as plataformas constroem estruturas normativas programadas justamente para observar acontecimentos sociais específicos que lhe são internos – as manifestações de seus usuários – sob a chave do lícito / ilícito. Nesse campo, a ideia que utilizei no início deste trabalho, de “*definir aquilo que as pessoas podem dizer*” é traduzida, pela segunda parte da dissertação, por atribuir, àquilo que se diz na plataforma, um dos lados do código lícito / ilícito. Quem publica um vídeo em que há nudez no Facebook, nega o Holocausto em um vídeo no YouTube, ou dissemina o discurso de ódio no Twitter diz o que não pode ser dito – ou, na linguagem de Teubner, gera um acontecimento social que é observado por esse fenômeno normativo pelo lado do “*ilícito*”.

Esse diagnóstico inicial é corroborado pela esquematização do Capítulo 2. Como expus ali, há dois tipos de atividade vinculadas à moderação de conteúdo em sentido estrito<sup>735</sup>: a de definir regras sobre o que os usuários podem dizer, e a de aplicar essas regras em casos concretos.

Sob a perspectiva do primeiro critério de Teubner, ambas as atividades são semelhantes, no sentido de que observam – abstratamente ou concretamente – um acontecimento interno à plataforma como lícito ou ilícito. Essa característica unifica as

---

<sup>734</sup> Cf. *supra*, Cap. I.2, 2.2.

<sup>735</sup> Cf. respectivamente e *supra*, Cap. I.2, 2.3.1 e 2.4.1.



diferentes modalidades de normatização abstrata – não importa se gerais e opacas, gerais e públicas, específicas e opacas ou específicas e públicas –, porque, em todos esses casos, a plataforma fixa uma expectativa normativa quanto àquilo que é lícito ou não dizer no ambiente que administra: operando uma tradução semelhante à do parágrafo anterior, dizer “*Não poste: [...] Vídeos de pessoas agonizando, feridas ou mortas*”<sup>736</sup> é a mesma coisa que afirmar que é ilícito publicar um vídeo de uma pessoa morta.

De forma semelhante, essa atribuição do lícito / ilícito acontece na expressão adjudicatória da moderação de conteúdo: assim como excluir a imagem de Egeland e banir as publicações de Jones significa observá-las como ilícitas, permitir que as pessoas neguem o Holocausto no Facebook é o mesmo que firmar que esse é um posicionamento lícito naquele ambiente. Paralelamente, assim como contratos e cortes arbitrais impõem às partes obrigações quanto àquilo que é lícito fazer na *lex mercatoria*, e códigos corporativos e comitês disciplinares expõem aos empregados de uma empresa o que é considerado ético fazer, os Padrões da Comunidade do Facebook, as Diretrizes da Comunidade do YouTube e as Regras do Twitter, jungidos com seus respectivos procedimentos de moderação de conteúdo, fixam, aos usuários das plataformas, o que é lícito ou ilícito dizer.

Quanto ao segundo critério – o da reflexividade – a resposta precisa ser mais cautelosa. Aqui, o critério de Teubner se refere à existência de normas secundárias, que se concretizam no “*desenvolvimento de atores ou instâncias, que são responsáveis pelo estabelecimento, modificação, interpretação e implementação da formação de normas primárias*”<sup>737</sup>. Como acontece com a *lex mercatoria* e com os códigos corporativos, também na moderação de conteúdo essa é uma característica em desenvolvimento. Em todos esses três casos, a explicação para isso é a mesma – e que vale a pena aqui retomar<sup>738</sup>: na versão local do direito, a reflexividade é garantida por um conjunto de instituições vinculadas ao Estado, a exemplo das cortes de justiça, que são responsáveis por aplicar a interpretar essas regras, analisando-as a partir de seus próprios critérios. Uma

---

<sup>736</sup> Cf. *supra*, nota n.º 279.

<sup>737</sup> “[...] *die Herausbildung von Verfahrensrollen, die für die Setzung, Modifizierung, Interpretation und Implementation der primären Normierungen zuständig sind*”. TEUBNER, G. Codes of Conduct multinationaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. et al (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

<sup>738</sup> Cf. *supra*, Cap. II.5, 2.2.2.

vez que esse fenômeno normativo se desvincula dessa instância política local, as manifestações do direito transnacional se veem obrigadas a construir as suas próprias instituições de reflexividade, um processo que se encontra em diferentes estágios de desenvolvimento: assim, a *lex mercatoria* já conta com câmaras de arbitragem estabelecidas, ao passo em que os códigos corporativos ainda precisam desenvolver melhor esse aspecto, atualmente restrito a comitês disciplinares instituídos dentro das empresas para julgar a conduta de seus funcionários.

No caso da moderação de conteúdo, a gestação desses mecanismos de reflexividade é identificada no movimento de institucionalização que perpassa as plataformas da internet, no sentido de que a atividade de moderar conteúdo foi objeto de transformações que a profissionalizaram, observando e corrigindo a aplicação dessas regras de acordo com os seus próprios critérios. Isso pode ser melhor compreendido acessando uma perspectiva histórica dessa forma de normatização, que reconstruí na primeira parte do trabalho<sup>739</sup>. Como visto, há, na moderação de conteúdo, um primeiro momento pré-profissional, que corresponde à situação, por exemplo, no Facebook antes do ingresso de Dave Willner – responsável pelo movimento que Klonick denomina “*dos padrões às regras*”<sup>740</sup> – e do Google, antes do trabalho desenvolvido por Nicole Wong. Nesse estágio, as plataformas já diziam aquilo que seus usuários podiam ou não dizer – ou seja, o primeiro critério era satisfeito –, mas ainda não havia regras públicas nem gerais; nem se tinha definido um formato procedimental para moderar conteúdo, algo que se resumia a um mecanismo conduzido por poucas pessoas, sem padronizações claras nem critérios – normativos ou não – pelos quais essas regras eram aplicadas. Esse arbítrio é bem exemplificado pelo trecho que citei no Capítulo 2, segundo o qual moderar conteúdo no YouTube era uma atividade realizada por uma equipe de, no máximo, dez pessoas, que passavam o dia vendo vídeos e os classificando de acordo com “*um pedaço único de papel, dobrado em dois para mostrar uma lista de instruções dividida em itens: remova vídeos de*

---

<sup>739</sup> Cf. *supra*, Cap. I.2, 2.3.1.

<sup>740</sup> “*The development at YouTube and Facebook from standards to rules for content moderation reflects these trade-offs.*” (“*O desenvolvimento no YouTube e no Facebook dos padrões às normas reflete esses tradeoffs*”). KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, p. 1.632.

*abuso animal. Remova vídeos mostrando sangue. Remova nudez visível. Remova pornografia*<sup>741</sup>.

Com o trabalho de Willner no Facebook e de Wong no Google – que trouxe, a reboque, as demais plataformas, em particular o Twitter –, a moderação de conteúdo passa a um segundo momento, em um movimento que acima identifiquei como a sua profissionalização – e que, visto sob a perspectiva de Teubner, é iluminado como o início de uma maior reflexividade na aplicação de suas regras. Elas se tornam, ao menos em alguma medida, públicas e gerais; criam-se procedimentos pelos quais elas são aplicadas por fases delimitadas nas quais se detecta o conteúdo, ele é avaliado a partir dessas balizas normativas e, se o caso, o seu responsável é sancionado; por fim, e mais importante, estabelecem-se formas de observar normativamente essa aplicação normativa de acordo com o código jurídico. Assim, no caso da moderação de conteúdo decidida manualmente – isto é, a partir da avaliação direta de um ser humano, na maioria das vezes por meio de uma empresa terceirizada –, são estabelecidos controles por amostragem que detectam a precisão com a qual essas avaliações normativas foram realizadas, algo que é, inclusive, um critério da remuneração recebida por essas empresas<sup>742</sup>. Também é um indício dessa reflexividade progressiva a circunstância de que a aplicação dessas normas é utilizada como um argumento para modifica-las, tornando a sua alteração a consequência delas mesmas – mesmo que ainda não por um procedimento organizado como o político-legislativo. No caso Egeland, por exemplo, da aplicação de regras quanto à nudez a uma imagem da Guerra do Vietnã emergiu uma exceção à norma geral, permitindo que fotos que retratem eventos históricos sejam publicadas no Facebook, mesmo que contenham algum grau de nudez<sup>743</sup>. Da mesma forma, uma reavaliação nos critérios utilizados pela plataforma na aplicação de suas regras levou à conclusão de que o separatismo branco estava indissociavelmente ligado ao supremacismo branco e, por isso, também deveria ser

---

<sup>741</sup> “*Mora-Blanco’s team — 10 people in total — was dubbed The SQUAD (Safety, Quality, and User Advocacy Department). They worked in teams of four to six, some doing day shifts and some night, reviewing videos around the clock. [...] Mora-Blanco sat next to Misty Ewing-Davis, who, having been on the job a few months, counted as an old hand. On the table before them was a single piece of paper, folded in half to show a bullet-point list of instructions: Remove videos of animal abuse. Remove videos showing blood. Remove visible nudity. Remove pornography*”. BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

<sup>742</sup> Cf. *supra*, Cap. I.2, 2.4.2.2.

<sup>743</sup> Cf. *supra*, Cap. I.1, 3.2.2.1.

proibido, como uma regra geral<sup>744</sup>; nesse mesmo sentido, a mudança de política do YouTube, que agora não mais permite vídeos que neguem o Holocausto na plataforma<sup>745</sup>.

Esse movimento de reflexividade se completa, em particular – e, em um passo que, ao menos neste momento, é dado mais à frente das demais plataformas – pelo Facebook. Como afirmei acima<sup>746</sup>, a plataforma é a pioneira na proposta e criação – que anda a passos largos – de uma instituição independente e à parte de sua estrutura empresarial, dedicada à avaliação das decisões da moderação de conteúdo – a “*Suprema Corte do Facebook*”. Pelo que se tem divulgado pela plataforma, a instituição servirá como aplicadora independente das regras do Facebook, submetendo-as, portanto, a elas mesmas – e não ao arbítrio de alguém. Ela representa, nesse sentido, um passo adiante nesse movimento reflexivo, porque, à semelhança do que aconteceu com a *lex mercatoria*, prevê o estabelecimento de uma espécie de tribunal arbitral internacional da liberdade de expressão – ao qual se estende, portanto, o argumento de Teubner quanto à reflexividade que esse tipo de instituição confere ao direito:

*“A despeito de a arbitragem e os contratos-padrões estejam, eles mesmos, baseados no contrato, eles transformam a criação contratual de direitos e deveres em um ‘direito não-oficial’ que é agora controlado e disciplinado pelo ‘direito oficial’ dos corpos arbitrais. A arbitragem privada e a legislação privada se transformam no núcleo de um sistema de decisão que começa a construir uma hierarquia de normas e de corpos organizacionais. Isso faz a reflexividade da lex mercatoria possível”*<sup>747</sup>.

Em conclusão: mesmo que em níveis variáveis de desenvolvimento, a depender da plataforma e do momento que é tomado, *a moderação de conteúdo é o direito das plataformas*. Ela é a manifestação normativa transnacional de um fenômeno jurídico que, impactado pela globalização, extravasa o âmbito político local e, diferenciando-se

---

<sup>744</sup> Cf. *supra*, Cap. I.1, 3.2.2.2.

<sup>745</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>746</sup> Cf. *supra*, Cap. II.3, 2.

<sup>747</sup> “Denn obwohl sowohl Schiedsgerichtsbarkeit als auch Allgemeine Geschäftsbedingungen selbst auf Vertrag beruhen, transformieren sie doch die vertraglichen Rechte und Pflichten in “nicht-offizielles Recht”, das dann vom “offiziellen Recht” der Schiedsgerichte kontrolliert und diszipliniert wird. Private Schiedsgerichte und private Gesetzgebung werden auf diese Weise zum Mittelpunkt eines Entscheidungssystems, das eine Hierarchie von Normen und Entscheidungsstellen zu errichten beginnt. Dies macht die Reflexivität der *lex mercatoria* möglich”. TEUBNER, G. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. *Rechtshistorisches Journal*, v. 15, 1996, pp. 255/290 (n. p.).

funcionalmente, constitui-se em uma escala global tanto (i) atribuindo, a acontecimentos sociais – no caso, às manifestações dos usuários das plataformas – um dos lados do código lícito / ilícito; quanto (ii) desenvolvendo mecanismos de reflexividade, que vão desde a sua profissionalização até a elaboração de mecanismos adjudicatórios externos e independentes.

Assim como a afirmação de que a *lex mercatoria* e os códigos corporativos fazem parte do direito não tem um valor apenas taxonômico, também definir a moderação de conteúdo como o direito das plataformas tem um significado muito além do classificatório. Isso porque é exatamente a confirmação da possibilidade de se conformar o que acontece na moderação de conteúdo a esse conceito de direito do pluralismo funcionalista que permite o último passo do argumento: utilizar as ferramentas e as conclusões da Teoria dos Sistemas, quanto ao papel do procedimento, para resolver o problema apresentado no Capítulo 3. Se o pluralismo funcionalista é aplicável às plataformas; e se a moderação de conteúdo é o direito das plataformas, então faz sentido investigar em que medida os papéis que a Teoria dos Sistemas atribuiu ao procedimento também se aplicam àquilo que acontece nas plataformas – em especial, à transformação que permanece inexplicada e que tem feito do procedimento de moderar conteúdo algo cada vez mais próximo à forma como os Estados produzem decisões vinculantes.

Como expliquei antes<sup>748</sup>, esse diagnóstico se desdobrou, a partir da reconstrução do diálogo Luhmann/Teubner, em duas relações: primeiro, a que existe entre procedimento e legitimidade, de a *Legitimação pelo Procedimento*, que permanece fazendo sentido no direito globalizado; e, segundo, na que há entre procedimento e autonomização, e que é exposta, com mais clareza, na transição feita por Teubner do direito local ao direito transnacional. A partir de cada uma dessas relações, estabeleci hipóteses para responder à pergunta do Capítulo 3, que passo a testar a seguir – a primeira nesta seção, e a segunda na seção seguinte.

A primeira hipótese é de que o movimento identificado nas plataformas da internet, de aproximar os seus procedimentos à forma como os Estados tomam decisões vinculantes, é explicada como uma forma de solucionar um problema de legitimidade, que existe quanto às decisões produzidas pela moderação de conteúdo. Ou seja: o *tradeoff* oculto por trás do discurso tanto dos movimentos sociais, quanto das próprias plataformas, segundo o qual elas mudaram a forma como decidem porque são sensíveis às demandas

---

<sup>748</sup> Cf., respectivamente e *supra*, Cap. II.4, 2.2.2 e Cap. II.5, 2.2.3.

apresentadas por seus usuários, é o de que, com esse movimento, mesmo que abrindo mão de parte do seu poder, elas são capazes de obter, em troca, legitimidade. Para testar se isso é verdade ou não, são necessários dois passos: primeiro, é preciso saber se existe um problema de legitimidade nas plataformas; segundo, se é correto interpretar o movimento que descrevi no Capítulo 3 como uma tentativa de solucionar esse problema, ou seja, como uma aproximação entre a forma como as plataformas decidem e o modelo de procedimento, no sentido em que apresentado por Luhmann.

Primeiro: há, nas plataformas, um problema de legitimidade? Para responder a essa pergunta é preciso, antes, retomar o que Luhmann quis dizer com esse termo: para ele, legitimidade significa uma disposição geral para se aceitar uma decisão, mesmo que ela seja desfavorável. Um problema de legitimidade, portanto, é uma situação em que essa aceitabilidade não existe ou não é suficiente, generalizando o conflito decorrente de uma decisão desfavorável a um nível em ele se torna incontrolável.

Meu argumento, a esse respeito, é de que existem indicativos suficientes ao longo deste trabalho de que a legitimidade é, nesse particular sentido, um problema para as decisões tomadas pela moderação de conteúdo das plataformas da internet. A principal forma de se perceber isso é a constatação de que, em alguns casos, *meios externos às plataformas são mobilizados pelas pessoas para tentar reverter as suas decisões* – o que expõe não apenas que elas não as aceitam, como que essa não aceitabilidade se reflete em uma generalização do conflito, extravasando inclusive o ambiente da plataforma e fazendo disso um problema para a empresa que administra aquele espaço, que é obrigada inclusive a ceder diante das pressões mobilizadas contrariamente à sua decisão. As pessoas não apenas não aceitam a decisão das plataformas; elas conseguem convencer as outras pessoas de que aquela é uma decisão não aceitável, escalando um problema entre elas e as plataformas em controvérsias que, como exponho adiante, mobilizam desde a imprensa internacional até Chefes de Estado.

Isso fica claro nos três casos que utilizei no primeiro capítulo para demonstrar que entender a regulação da liberdade de expressão apenas pelo que fazem o Estado é um erro. No primeiro caso<sup>749</sup>, Tom Egeland, um jornalista norueguês, publicou a imagem *O Horror da Guerra* no Facebook em um artigo que discutia o impacto da fotografia na história da guerra. A imagem foi retirada pela plataforma, porque violava sua política de conteúdo quanto à nudez – algo que foi feito em um contexto de moderação de

---

<sup>749</sup> Cf. *supra*, Cap. I.1, 3.2.2.1.

conteúdo em que inexistiam mecanismos de transparência, participação ou recorribilidade à disposição de Egeland. Não havia, portanto, como saber as razões por trás da exclusão, nem estavam disponíveis vias pela própria plataforma em que esse tipo de questão pudesse ser abordada: a perplexidade de Egeland, que não entendeu bem o que estava acontecendo, é o melhor retrato de como essa forma de moderar conteúdo produzia tudo menos uma decisão aceitável. O resultado não poderia ser outro: não apenas a não aceitação do desfavorável, mas a mobilização dos mecanismos que se tinha à mão para revertê-la. A controvérsia quanto à exclusão da imagem, como detalhei no Capítulo 1, tomou proporções muito além dos limites da plataforma, envolvendo a imprensa local e internacional, e o próprio governo norueguês, após a Primeira-Ministra do país compartilhar – e ter excluída – uma publicação com a imagem. As pressões extraplataforma – e, portanto, frontalmente contrárias a uma posição de aceitação do decidido – escalaram a questão a um ponto em que o Facebook se viu obrigado a recuar, reestabelecendo a imagem e incluindo uma exceção a respeito na sua política quanto a conteúdos com nudez.

No segundo caso, a controvérsia quanto ao banimento de Alex Jones teve um formato semelhante ao do caso Egeland: assim como com a exclusão de *O Horror da Guerra*, a decisão de retirar o ativista das plataformas da internet, foi tomada unilateralmente pelo Facebook, Twitter e YouTube<sup>750</sup>. Como no caso anterior, a posição de fazê-lo, no lugar de resolver os problemas causados pelas publicações do *Infowars*, escalou a situação, gerando uma controvérsia externa à plataforma a respeito da existência de mecanismos de censura por parte dessas empresas. A situação se agravou até o ponto em que o Presidente dos Estados Unidos se manifestou a respeito, colocando em evidência que os mecanismos utilizados pelas plataformas para definir aquilo que poderia ou não ser dito na plataforma – ou, no caso específico, quem pode se manifestar naquele espaço – estavam longe de resolver as controvérsias que surgiam a esse respeito, substituindo-as por problemas ainda mais extensos. Houve quem, por exemplo, concordasse que aquilo que Alex Jones publicava era danoso às pessoas, mas discordasse da posição das plataformas de bani-lo, ao menos da forma como feito. Mesmo que, nesse caso, as plataformas tenham mantido a sua posição, o custo de uma decisão ilegítima se apresentou no desgaste enfrentado por essas empresas, não só em relação ao próprio Alex Jones – que, é possível, nunca aceitaria uma decisão desse tipo –, mas fundamentalmente em relação às outras

---

<sup>750</sup> Cf. *supra*, Cap. I.1, 3.2.2.2.

peessoas, que, ao contrário do que se poderia esperar, não ofereceram um apoio unânime àquilo que tinha sido feito.

No terceiro caso, a discussão quanto à se as pessoas devem poder ou não negar o Holocausto foi reiniciada no Facebook após uma entrevista de Mark Zuckerberg, em que ele esclareceu a posição da plataforma de não excluir esse tipo de publicação<sup>751</sup>. Mais uma vez, ficam claras as deficiências que os mecanismos decisórios da moderação de conteúdo presentes à época têm para manejar esse tipo de problema, em particular as barreiras que essas decisões têm de serem aceitas pelas pessoas que, por não disporem de vias de reclamar do seu conteúdo, levam a uma escalada do problema no âmbito extraplataforma, generalizando-o. Assim, a permissibilidade do Facebook quanto a publicações que negam o Holocausto gerou uma carta aberta de sobreviventes do nazismo e manifestações de repúdio da comunidade sionista, que se recusava a aceitar a posição da plataforma, por entendê-la como uma forma de antisemitismo disfarçada. O conflito, mais uma vez, não teve lugar na plataforma – que não expunha às denúncias feitas em relação a esse conteúdo as suas razões de decidir, não permitia que elas fossem acompanhadas pelas pessoas, nem viabilizava que elas fossem recorríveis – tornando-se incontrolável na medida em que generalizado como uma controvérsia que envolveu grupos da sociedade civil e imprensa.

Esse problema de legitimidade, que apresento nessas três situações – porque elas foram tratadas de forma mais extensa acima e, portanto, dispensam uma digressão maior – aparece também em outros casos que mencionei ao longo da dissertação, nos quais também se faz presente essa não aceitabilidade da decisão, traduzida tanto em uma generalização do conflito, quanto na busca por mecanismos extraplataforma capazes de revertê-la. No caso Lipman<sup>752</sup>, por exemplo, isso aparece na perplexidade do grupo nacionalista israelense quanto à exclusão da imagem contrária à palestina e no teste, feito pelo parlamentar e depois exposto em uma carta aberta em um jornal local, que evidenciou a falta de critério do Facebook – e que redundou no reestabelecimento da publicação. Na situação envolvendo manifestações de *drag queens* que foram banidas da plataforma porque utilizavam o seu nome profissional<sup>753</sup>, a ausência de mecanismos que justificassem e permitissem impugnar a decisão do Facebook levou a uma mobilização da imprensa

---

<sup>751</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>752</sup> Cf. *supra*, Introdução.

<sup>753</sup> Cf. *supra*, Cap. I.2, 2.4.2.1.



internacional para que o erro fosse corrigido – descobriu-se, depois, que a moderação de conteúdo havia sido induzida por um radical religioso –, mais uma vez expondo tanto a não aceitação da decisão, quanto a generalização de um conflito que poderia ter sido resolvido na própria plataforma, por meio de um canal institucional recursal próprio – que não existia na época.

Com esses exemplos, minha intenção não é expor que os problemas de legitimidade acontecem todas as vezes em que se modera conteúdo, mas sim que, em algumas situações, eles existem e repercutem na forma de uma generalização das controvérsias que decorrem da moderação de conteúdo das plataformas da internet, expondo, com clareza, a ausência de aceitabilidade de suas decisões e de mecanismos capazes de dar conta dessa insatisfação antes que ela escale e envolva atores externos como a imprensa e, em alguns casos, até mesmo autoridades governamentais. A ausência da legitimidade, portanto, na forma de uma aceitação generalizada que viabiliza à plataforma tomar decisões e tê-las cumpridas, é um problema não só em si – ou seja, como um inconformismo pessoal do usuário em relação ao Facebook, Twitter ou YouTube –, mas porque ela leva a uma generalização do conflito, ocasionando situações em que, tal como ocorreu no caso Alex Jones, essas empresas são vistas como censoras da liberdade de expressão das pessoas.

Se existe um problema de legitimidade no direito das plataformas, o passo seguinte é entender se faz sentido dizer que a mudança nos procedimentos de moderação de conteúdo é uma forma de tentar solucioná-lo. Aqui, é importante retomar algo dito no Capítulo 3: o funcionalismo estrutural não se dedica a descobrir intencionalidades, mas sim em entender o sentido de transformações sociais a partir da ideia de *função*. Dito de outra forma: a questão que se coloca não é identificar a partir de entrevistas feitas com os atores desse movimento, por exemplo, qual era a intenção que se tinha ao mudar a forma como as decisões são tomadas nas plataformas; ou encontrar, em algum documento vazado do Facebook, uma pista que leve a essa conclusão. A proposta do funcionalismo estrutural é exatamente oferecer uma alternativa a essas abordagens – que têm sido as mais comuns –, dedicadas mais em obter informações a respeito da moderação de conteúdo com os seus atores do que em organizá-las e lhes conferir corpo teórico, porque elas têm limitações. Em particular, e quanto ao problema do Capítulo 3, pode não convir, a um ator envolvido com uma plataforma, expor aquilo que está por trás desse movimento; além disso, mesmo que fosse possível obter essa informação de modo fidedigno, nada garante que os responsáveis por esse movimento tenham a plena consciência do que fazem. Por isso, a minha proposta

não é acessar esse tipo de informação, mas sim apresentar e testar hipóteses capazes de revelar o papel que essa mudança procedimental tem no contexto das plataformas.

Nesse sentido, o Capítulo 3 reconstruiu o movimento das plataformas a partir dos *Santa Clara Principles On Transparency and Accountability in Content Moderation*, um documento elaborado por advogados e acadêmicos que reúne exigências para que o modelo da moderação de conteúdo se aproximasse ao do devido processo legal utilizado nos Estados. Essa transformação foi decomposta em três critérios – “*números*”, “*participação*” e “*recorribilidade*” – que expõem essa transformação procedimental. A questão, portanto, é em que medida essas secções que compõem o movimento principal apontam para uma modificação que transforma o formato utilizado inicialmente pelas plataformas em algo semelhante a um procedimento, em termos funcionalistas.

Quanto ao primeiro critério, a mudança na moderação de conteúdo se caracteriza como uma transição entre *opacidade* e *transparência* procedimentais. Concretamente, isso significou que, antes da mudança, não havia informações disponíveis sobre como as decisões sobre aquilo que as pessoas podem dizer eram tomadas: não se sabia, também, quando, como e se a plataforma havia agido sobre uma publicação; qual era o volume de sua atuação sobre aquilo que os usuários diziam. Também não eram claras as regras e a forma como o procedimento caminhava de um momento inicial, em que algo era denunciado aos moderadores – ou por eles detectado – até a decisão final, de manter ou não algo na plataforma. Após a transformação, plataformas como o Facebook, Twitter e YouTube passaram a disponibilizar tanto informações específicas aos usuários quanto à sua atuação específicas em suas publicações, quanto relatos gerais quanto à forma como agiam nas publicações, publicizando, também, as regras pelas quais atuavam sobre o conteúdo.

Quanto ao segundo critério, a mudança na moderação de conteúdo se caracteriza como uma transição entre *ausência de participação* e *participação* dos seus usuários. No modelo anterior, os usuários eram um item à parte da decisão, recebendo-a com a perplexidade exposta por Egeland quando soube da exclusão da fotografia de guerra, ou com a incompreensão do grupo político que acionou Lipman após ter a manifestação sobre a palestina retirada do Facebook. Não eram de conhecimento público as regras e os procedimentos adotados pela moderação de conteúdo; não havia como acompanhar se algo estava sob decisão da plataforma; e, sobretudo, sequer se tinha ideia das razões que haviam embasado uma decisão. Após a transformação, o usuário foi chamado a participar desse procedimento decisório, conferindo-se, a ele, acesso às normas

que regem a moderação de conteúdo – inclusive com seções procedimentais de seus padrões de comunidade, em que indicadas as fases do procedimento; a possibilidade de acompanhar denúncias a respeito do que outra pessoa publicou na plataforma; e, por fim, o acesso à exposição dos motivos que a plataforma utilizou para decidir o caso. Migrou-se, nesse sentido, de um estado em que o usuário era o expectador de um procedimento decisório, para outro em que mecanismos estimulam que ele tome parte dessa decisão, tanto acionando a plataforma, quanto o acompanhando e argumentando a respeito de seu resultado, que, agora, é apresentado como uma definição fundamentada – e recorrível, como exponho a seguir – a respeito do que ele pode dizer.

Pelo terceiro critério, fica perceptível a mudança de uma moderação de conteúdo em que *não há recorribilidade* para outra em que *há recorribilidade* – ou seja, em que as decisões tomadas pela plataforma não são mais definitivas. Antes, portanto, uma vez definida pela plataforma que o um determinado conteúdo não era compatível com os padrões da comunidade, ele era retirado sem que houvesse uma via interna de administrar a irresignação do usuário. A única possibilidade, como os exemplos acima expõem, era de atuar generalizando a controvérsia, trazendo-a para um âmbito externo das plataformas em que ela poderia ser tratada por instrumentos como a imprensa ou autoridades governamentais, generalizando a divergência entre plataforma e usuário e traduzindo-a, por exemplo, em uma controvérsia entre o governo norueguês e o Facebook – no caso Egeland –; o Presidente dos Estados Unidos e as plataformas em geral – no caso Alex Jones –; e a comunidade sionista e o Facebook – no caso da negação do Holocausto. Com a transformação detectada e a consequente implantação de mecanismos internos de recorribilidade, tornou-se possível canalizar a controvérsia para vias internas, resolvendo-a como uma disputa entre plataforma e usuário, diminuindo o seu potencial generalizador. Um exemplo disso é o que ocorreu quando o YouTube resolveu modificar a sua política quanto a publicações que neguem o Holocausto, proibindo-as<sup>754</sup>. Como a detecção do conteúdo potencialmente violador dessas novas diretrizes não ficou a cargo da ação humana – ou seja, na classificação do Capítulo 2<sup>755</sup>, trata-se de uma moderação de conteúdo com detecção via ação automatizada –, ocorreram erros, porque vídeos educativos sobre o nazismo foram confundidos como instrumentos de propagação do

---

<sup>754</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

<sup>755</sup> Cf. *supra*, Cap. I.2, 2.4.2.1.

discurso de ódio. Mesmo que a questão tenha sido noticiada em alguns veículos<sup>756</sup>, o problema acabou sendo rapidamente resolvido após os responsáveis pelas publicações utilizarem do mecanismo de recorribilidade do YouTube, que reestabeleceu os vídeos e encerrou a questão, que permaneceu restrita à plataforma na forma de um conflito não generalizado.

Como interpretados essas transformações? Sob a perspectiva da Teoria dos Sistemas, esses movimentos não são percebidos reflexos de um *modelo jurídico* – como o das constituições que garantem o devido processo legal –, mas sim como aproximações a um *modelo funcional* – no caso, o que relaciona procedimento e legitimidade. Assim, transparência, participação e recorribilidade, tal como garantidas pelas plataformas aos seus usuários, ganham um significado diferente: elas não são a aplicação da Décima Quarta Emenda pelas plataformas – algo que não faz sentido, porque elas não estão obrigadas a isso e, fazê-lo é algo, à primeira vista, desvantajoso – mas passam a ser entendidas como algo funcionalmente útil a essas empresas, porque conferem, ao resultado do procedimento de moderação de conteúdo, legitimidade.

Isso acontece porque, ao estabelecer uma forma transparente, participativa e recorrível de tomar decisões, as plataformas dão um passo no sentido tanto de fazer com que o usuário participe do procedimento, quanto o de tornar a sua decisão incerta, legitimando-a. No primeiro sentido desse movimento, traz-se o usuário para dentro do processo que constrói a decisão que irá definir se aquilo que ele publicou deve ou não permanecer na plataforma. Atribui-se, a ele, um papel nessa forma de decidir: ele passa a ter acesso a informações como as normas aplicáveis ao conteúdo disponibilizado na plataforma, os motivos que levaram à sua retirada e os caminhos procedimentos para reverter esse posicionamento. Em troca, o usuário se vê na contingência de aceitar o resultado desse procedimento, para o qual ele contribuiu: o conflito que antes se colocava nos jornais, como no caso Egeland – em que a mudança da decisão foi obtida a partir de sua repercussão na imprensa – é internalizado na plataforma, que o traz para um ambiente controlado em que se viabiliza ao usuário influenciar na tomada de decisão, cobrando-se, entretanto, o preço da aceitação. A divergência que era uma disputa entre usuário e plataforma se transforma em uma discordância jurídica, quanto à correta interpretação de um dispositivo: aquilo que, no exemplo de Egeland, era um debate sobre como as plataformas censuram os seus usuários, seria traduzido, hoje, em uma controvérsia quanto

---

<sup>756</sup> Cf. *supra*, nota n.º 361.

à melhor interpretação do artigo dos Padrões da Comunidade que vedam a nudez. Como disse antes, “*O descontentamento é especificado, traduzido tecnicamente e submetido a uma decisão jurídica, que o resolve de forma cirúrgica*”<sup>757</sup>, moldando-o a um padrão resolvível por uma decisão que, é, ao fim, aceita pelas partes. Entram em ação, em síntese, as características funcionalmente atribuídas por Luhmann ao processo judicial – o que mais se assemelha à moderação de conteúdo: a criação de fronteiras entre o que faz parte e o que não faz parte do procedimento – ao permitir a recorribilidade, por exemplo, a plataforma tem um argumento para exigir que o usuário utilize das vias próprias para manifestar o seu inconformismo, e não simplesmente vá à imprensa expô-lo; a autonomia em relação àquilo que lhe é externo – torna-se possível decidir no procedimento de acordo com os argumentos apresentados pelo usuário; o estabelecimento de formas de contato entre as partes – no caso, plataforma e usuário; a adoção de papéis uns dos outros – o usuário, que agora tem acesso às razões que levaram à atuação da moderação de conteúdo, pode compreender melhor porque a plataforma agiu de uma determinada maneira; a construção, pelas partes, de representações coerentes de si mesmas – a plataforma, por exemplo, é impelida a adotar posicionamentos coerentes, sob pena de isso se transformar em um argumento contra ela, como o foi no caso Lipman; a construção de conflitos permitidos – porque restritos, agora, ao procedimento; a limitação à generalização dos conflitos – que agora têm um *locus* interno na plataforma, diminuindo a força de mecanismos que os escalem para instâncias governamentais ou da própria imprensa; a garantia de mecanismos de representatividade àqueles que não fazem parte do procedimento – tornando, assim, a decisão aceitável inclusive para as pessoas que não fazem parte da controvérsia, algo que se mostraria importante em casos como o de Alex Jones; e a aplicação de programas condicionais que restringem a crítica a um escopo técnico, minimizando a divergência – que agora se adstringe à melhor interpretação das regras da plataforma. A moderação de conteúdo é transformada, nesse sentido, em um mecanismo capaz de produzir aceitação – e, por consequência legitimidade.

Essa aceitação tem a sua chave, como afirmei antes – e como aparece, agora, nas plataformas – pela segunda direção proporcionada pela transformação descrita acima. Em síntese, ao proporcionar uma decisão mais transparente, participativa e recorrível, a moderação de conteúdo das plataformas da internet oferece aos seus usuários uma garantia de que ela é incerta, ou seja, de que ela é um resultado do próprio

---

<sup>757</sup> Cf. *supra*, Cap. II.4, 2.3.1.

procedimento, e não algo que já estava definido antes de ele ocorrer. Assim como aconteceu na formação dos Estados e de seus procedimentos, esse é um passo em desenvolvimento, e que é exposto em sua versão mais completa nos mecanismos decisórios do Facebook: ao estabelecer uma instância externa e independente à qual o usuário pode recorrer, ele deu um passo final no sentido de garantir aos seus usuários que não domina o resultado do procedimento. Isso explica por que esse é um movimento tão celebrado: ao mostrar que a decisão final não está mais nas suas mãos, o Facebook avançou definitivamente para garantir a incerteza que caracteriza o procedimento – no sentido de Luhmann – e, nesse sentido, construir um mecanismo capaz de gerar aceitabilidade decisória.

Isso acontece porque, como expliquei antes, à medida que as pessoas têm assegurado que a decisão que resulta do procedimento é incerta – seja, em um momento inicial, pela sua participação e recorribilidade; seja, em uma perspectiva mais avançada, pela estruturação de um órgão externo à plataforma –, elas se veem compelidas a tomar parte desse procedimento, em que lhes é oportunizada a possibilidade de influenciar, efetivamente, no seu resultado. Como é justamente essa participação do usuário que o coloca, ao final, em uma posição de aceitar a decisão produzida, a incerteza que a proporciona aumenta essa aceitabilidade e confere, assim, legitimidade à decisão.

Em conclusão: tanto existe um problema de legitimidade nas plataformas, quanto esse problema é solucionável a partir da transformação observada no procedimento adotado pela moderação de conteúdo para definir aquilo que as pessoas podem dizer, uma perspectiva que confere sentido à modificação que as empresas por trás desses espaços têm implantado, sem toma-las como simples efetivadoras da vontade de seus usuários. Juntas, essas constatações permitem afirmar, no quadro da relação que existe entre legitimidade e procedimento, que *as plataformas da internet passaram a adotar modelos decisórios próximos ao do Estado porque, com isso, vislumbram a possibilidade de resolver um problema de legitimidade, isto é, de tornar as decisões que tomam mais aceitáveis e, por consequência, tratar as controvérsias provenientes da moderação de conteúdo em um ambiente controlado, que não as generalize como têm acontecido.*

Se, portanto, de início, o movimento do Capítulo 3 parecia inexplicável, porque contrário aos interesses dessas empresas, ele agora se mostra coerente com os desafios que elas enfrentam ao moderar conteúdo. É esse, ao menos em parte, o *tradeoff* que se ocultava sob um discurso de atendimento às intenções dos usuários: é verdade que as plataformas abrem mão de parte do seu poder de definir aquilo que as pessoas podem

dizer nos ambientes que administram ao adotarem mecanismos como de transparência, participação e recorribilidade; mas não é verdade que isso acontece sem uma contrapartida. Pelo contrário: à perda de poder que ocorre instituindo um procedimento decisório mais próximo do adotado pelo Estado, corresponde um ganho de legitimidade, que se traduz em uma probabilidade maior de aceitação dessas decisões, mesmo que elas sejam – inevitavelmente – desfavoráveis a alguns de seus usuários.

### 3.3 O procedimento como constitucionalização

A aplicação do método funcionalista ao problema de se definir o papel que os procedimentos têm na estrutura social resultou em duas respostas diferentes, ambas trazidas para a realidade das plataformas pela transição feita por Teubner. Primeiro, procedimentos são instrumentos que produzem legitimidade; segundo, procedimentos são instrumentos da constitucionalização do direito, porque o autonomizam. Na seção anterior, abordei a hipótese que relaciona a primeira conclusão ao problema de se entender as transformações da moderação de conteúdo, argumentando que a adoção de um modelo decisório próximo ao do Estado se justifica como uma forma de resolver um problema de legitimidade – que se manifestava tanto na resistência dos usuários em aceitar as decisões das plataformas, quanto na tendência que essas controvérsias tem de se generalizar. Nesta seção, testo a segunda hipótese – de que essa aproximação procedimental também faz sentido como uma forma de conferir, à moderação de conteúdo, autonomia, em um processo de constitucionalização das plataformas.

De início, é preciso retomar o argumento de Teubner. De acordo com ele, a transição do direito da esfera local para a esfera transnacional gerou um problema: uma vez que o fenômeno jurídico se desvincula do político, ao qual associado nos Estados, ele também deixa para trás os mecanismos que foram desenvolvidos ao longo do tempo para garantir a sua autonomia – notadamente os limites que as constituições impõem à atuação política. Com isso, o direito se vê exposto, no âmbito transnacional, a intervenções por parte daquilo que ele regula, o que o faz replicar os mecanismos utilizados em âmbito local para tentar se manter independente em relação a, por exemplo, o âmbito econômico – regulamentado transnacionalmente via *lex mercatoria*. Vem à tona, portanto, uma espécie de memória institucional do direito, em que os mecanismos utilizados para lhe conferir autonomia sob o julgo da política são reimaginados para fazê-lo em um ambiente em que a política não existe – ao menos na forma como estruturada localmente.

Sob a perspectiva teórica, Teubner explica isso aplicando ao conceito de constituição um método semelhante ao utilizado com o de direito: ele argumenta que o fenômeno constitucional local é apenas uma manifestação de um mecanismo mais complexo, que pode ser funcionalmente compreendido como um *processo de constitucionalização*, no qual o direito constrói instrumentos de autonomia em relação àquilo que ele regula. Essa processo ocorre de duas maneiras: primeiro, como uma *constitucionalização procedimental*; segundo, como uma *constitucionalização material*. Tratarei a segunda adiante, em uma perspectiva de continuidade deste trabalho; quanto à primeira, ela diz respeito ao tema desta dissertação na medida em que é justamente o procedimento – no sentido sistêmico da palavra – que viabiliza ao direito que se construa de forma independente àquilo que ele regula, constitucionalizando-se ao formar esferas própria de produção e avaliação de razões jurídicas.

De que forma essa relação entre procedimento e autonomização esclarece o problema apresentado no Capítulo 3? Uma primeira aproximação a esse problema é compreender em que medida a ideia de constitucionalização mencionada por Teubner se aplica às normatizações que emergem no espaço da internet. Novamente, um indicativo no sentido de que a resposta para essa pergunta é afirmativa é fornecido pelo próprio Teubner: paralelamente ao que ocorreu com a ideia de direito transnacional – que é exemplificada com a *lex mercatoria* e os códigos corporativos, mas também citada como uma forma de explicar as normatizações que surgem na internet –, a constitucionalização também é primeiro abordada nesses dois exemplos paradigmáticas, mas é frequentemente associada, por Teubner, aos fenômenos normativos que têm lugar na internet. Assim, nesse espaço, segundo ele, “*os direitos constitucionais, tais como a liberdade de expressão, não são mais direcionados contra o Estado, mas contra atores privados (não apenas o Facebook, Google, Amazon e a Apple, mas também o regime da internet)*”<sup>758</sup>, isto é, “*os protestos da sociedade civil são cada vez mais levantados não (apenas) contra as instituições do Estado, mas seletivamente, diretamente e intencionalmente [direcionados] contra os atores corporativos*”<sup>759</sup> – algo que ficou claro no Capítulo 1, em que esse foi o argumento

---

<sup>758</sup> “*Constitutional rights such as free speech are no longer directed against the state but against private actors (not only Facebook, Google, Amazon and Apple, but also the Internet regime – ICANN and its subsidiaries)*”. TEUBNER, G. Horizontal Effects of Constitutional Rights in the Internet: A Legal Case on the Digital Constitution. **Italian Law Journal**, v. 3, n. 1, 2017, p. 195.

<sup>759</sup> “*Civil society protests are raised increasingly not (only) against institutions of the state, but selectively, directly, and intentionally, against corporate actors, which are accused of violating their social*



utilizado para demonstrar a insuficiência de uma abordagem da regulação da liberdade de expressão voltada exclusivamente à ação do Estado<sup>760</sup> –; fala-se na “*existência de elementos constitucionais [...] dentro dos regimes digitais da internet*”<sup>761</sup>; em particular, “*constituições podem igualmente ser encontradas, e, de fato, assumem uma importância crescente [...] nos intermediários da internet*”<sup>762</sup> – dentre os quais se localizam, inclusive, as plataformas da internet.

A presença do processo de constitucionalização – que, assim como o direito transnacional, é inicialmente concebida quanto à *lex mercatoria* e os códigos corporativos – na internet e, em particular, nas plataformas, é explicada por duas razões. Primeiro, porque a internet está inserida no contexto do direito transnacional, como expuseram as características de digitalização, privatização e globalização na seção anterior. Isso significa que, assim como na *lex mercatoria*, o direito da internet – a exemplo da moderação de conteúdo – é construído deixando para trás os mecanismos de autonomização que tinha em relação ao Estado, o que o torna vulnerável àquilo que ele regula. Segundo, porque esse risco à autonomia do direito se concretiza, na internet, a partir da constatação de que o fenômeno jurídico emerge nesse contexto transnacional em âmbitos essencialmente privados, uma consequência do protagonismo que os atores não-estatais têm nesse contexto – que Teubner identifica sob a ideia de privatização. Aplica-se, à internet, a constatação inicialmente endereçada aos códigos corporativos, no sentido de que “*empresas transnacionais têm se transformado nas autoridades constitucionais reais*”<sup>763</sup>.

---

*responsibilities*”. TEUBNER, G. Corporate Codes in the Varieties of Capitalism: How their Enforcement Depends Upon the Difference Among Production Regimes, in BECKERS, A. (ed.). Enforcing Corporate Social Responsibility. **Indiana Journal of International Law**, v. 24, pp. 81-97, 2017 (n. p.).

<sup>760</sup> Cf. *supra*, Cap. I.1, 3.3.3.

<sup>761</sup> “[...] *existence of constitutional elements—in the strict sense of the term—within economic regimes, within the academic system and within digital regimes of the Internet*”. TEUBNER, G. Fragmented Foundations: Societal Constitutionalism Beyond the Nation State, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, 2010, p. 333.

<sup>762</sup> “*Ebenso und mit steigender Bedeutung seien Bereichsverfassungen in Wirtschaftsunternehmen, Märkten, privaten Universitäten, Stiftungen, Presseunternehmen, Intermediären des Internets und anderen ‘privates’ Institutionen zu finden*”. TEUBNER, G. Transnationaler Verfassungsppluralismus: Neun Variationen über ein Thema von David Sciulli. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 76, 2016, p. 3.

<sup>763</sup> “[...] *transnational enterprises have become the real constitutional authority, because it is they who create corporate codes through their unilateral public declarations of self-obligation*”. TEUBNER, G. Corporate Codes in the Varieties of Capitalism: How their Enforcement Depends Upon the Difference

Isso levanta, ao menos em abstrato, o problema de como o direito das plataformas lida com a estrutura econômica em que surge e, em particular, com a possibilidade de se ver premido a ignorar os próprios critérios, por exemplo, em nome de uma decisão que gera mais lucro à corporação: surgem, portanto, *problemas constitucionais* – no sentido específico que Teubner atribui a esses termos, isto é, como *problemas de autonomia* – que exigem um *processo de constitucionalização* – um *processo de autonomização* – para serem resolvidos.

Como na seção anterior, entretanto, é insuficiente dizer que o problema da constitucionalização se coloca nos regimes jurídicos da internet. Para que ele possa ser capaz de esclarecer o problema do Capítulo 3, é preciso entender em que medida ele explica a transformação que ocorre nos procedimentos adotados pela moderação de conteúdo. Para isso, são necessários dois passos adicionais – e paralelos aos anteriores –, a serem dados a partir da esquematização feita a esse respeito na primeira parte da dissertação: primeiro, é preciso saber se existe um problema de autonomia no direito das plataformas – isto é, na moderação de conteúdo; segundo, é preciso saber se faz sentido interpretar o movimento de transformação procedimental da moderação de conteúdo como uma tentativa de solucioná-lo.

Primeiro: há, nas plataformas, um problema de autonomia? Argumento, a seguir, que a resposta para essa pergunta é afirmativa. Faço isso sustentando que essa questão se coloca de duas maneiras: primeiro, entre a moderação de conteúdo e as plataformas; segundo, entre a moderação de conteúdo e regimes normativos externos às plataformas. Como exponho nos próximos parágrafos, perceber a questão a partir dessa dualidade é importante, uma vez que explica por que as plataformas se movimentam no sentido de conferir autonomia à moderação de conteúdo, mesmo que isso signifique perder uma parte da influência que têm sobre as suas decisões.

Assim, inicialmente, o problema se posiciona na relação entre a plataforma e os instrumentos de moderação de conteúdo que ela cria. É essa a questão levantada diretamente por Teubner, que afirma, por exemplo, que “*em um exato paralelo com o direito global econômico, a lex electronica traz consigo o problema da corrupção estrutural, isto é, a massiva e não filtrada influência de interesses ‘privados’ na*

---

Among Production Regimes, in BECKERS, A. (ed.). Enforcing Corporate Social Responsibility. **Indiana Journal of International Law**, v. 24, pp. 81-97, 2017, (n. p.).

*elaboração do direito*”; por isso, segundo ele, “*é aqui que a questão constitucional da internet surge*”<sup>764</sup>.

No caso da moderação de conteúdo, a autonomia se coloca como um problema porque as estruturas que decidem aquilo que as pessoas podem dizer ou surgem no interior da empresa que administra esse espaço, ou, como expus acima, são terceirizadas pela contratação de prestadores de serviços especializados em fazer essa avaliação, em uma espécie de *callcenter* da liberdade de expressão<sup>765</sup>. De um modo ou de outro, o vínculo que existe entre plataforma e moderação de conteúdo coloca em dúvida se as decisões tomadas são um juízo neutro quanto a se um usuário violou ou não uma política de conteúdo ou se são apenas uma posição arbitrária que atende aos interesses econômicos das plataformas disfarçada de direito.

Nada garante, assim, que essas sejam decisões de um direito autônomo, e não expressões avançadas de decisões que atendam ao propósito lucrativo das plataformas, cuidadosamente vendidas como proteção à liberdade de expressão. Mais do que isso: a versão anterior da moderação de conteúdo oferecia indicativos claros de que era exatamente isso que acontecia. Como expus acima<sup>766</sup>, a combinação entre, primeiro, juízos materiais que repercutem o discurso estatal da liberdade de expressão, deixando, entretanto, um campo aberto para atuação da moderação de conteúdo – viabilizando que as plataformas encontrem um “*ponto ótimo*” em que maximizavam a interação; e, segundo, juízos procedimentais, que implantaram um modo de decidir arbitrário, capaz de conferir à plataforma espaço para decidir como quiser, traduziu-se em um modelo que era tudo menos autônomo, e, assim, que se prestava a refletir os interesses econômicos das plataformas.

Nesse sentido, o problema da perspectiva que vê a autonomia como um problema apenas entre plataforma e moderação de conteúdo é que ela não ajuda a compreender o problema do Capítulo 3. No limite, ela apenas desenvolve melhor a constatação de que, para as plataformas, não faz sentido ter procedimentos capazes de

---

<sup>764</sup> “*Und in genauer Parallele zum globalen Wirtschaftsrecht entsteht mit der lex electronica das Problem der strukturellen Korruption, also des massiven ungefilterten Einflusses von "privaten" Interessen auf den Rechtsbildungsprozeß. Hier stellt sich die Verfassungsfrage*”. TEUBNER, G. *Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. Zeitschrift für ausländisches öffentliches Recht und Völkerrecht*, v. 63, 2003, pp. 1/28 (n. p.).

<sup>765</sup> *Cf. supra*, Cap. I.2, 2.4.2.2.

<sup>766</sup> *Cf. supra*, Cap. II.3, 2.

conferir autonomia à moderação de conteúdo, porque isso as faz perder parte do poder que têm sobre aquilo que é veiculado na plataforma. É por isso que a transformação surpreende e intriga: se, no formato procedimental atual, os interesses das plataformas estão sendo atendidos – ao ponto de que uma das principais especialistas no assunto ter visto a transição que ocorre agora como improvável, porque “*as plataformas não tem razões para abrir mão do poder [...] ou para adotar as regras e ideias de transparência dos requisitos do devido processo legal [...]*”<sup>767</sup> –, então por que mudar? Por que estabelecer mecanismos que conferem autonomia à moderação de conteúdo?

A resposta para isso está na segunda maneira como o problema da autonomia se coloca para as plataformas. A autonomia da moderação de conteúdo não é algo que se constrói apenas diante dos interesses das empresas que estão por trás desses ambientes – e que podem tentar instrumentalizar o direito –, mas também contra os regimes jurídicos que são externos às plataformas. Aquilo que a moderação de conteúdo decide pode ser contrário tanto ao que interessa economicamente para a plataforma, quanto àquilo que determina o regime jurídico vigente em um Estado no qual ela atua. A autonomia que o direito busca, nesse sentido, vale tanto “*para dentro*” quanto “*para fora*” da plataforma: na medida em que as decisões da moderação de conteúdo são passíveis de serem fragilizadas pela intervenção da lógica econômica da empresa por trás da plataforma, elas também podem sê-lo por intervenções externas; na medida em que o direito das plataformas ganha autonomia contra uma, também se resguarda quanto à outra. Uma moderação de conteúdo independente, isto é, autônoma em relação às decisões da companhia à qual vinculada, é também uma moderação de conteúdo autônoma em relação às influências que lhe são externas, fazendo do que era uma questão subordinação de um regime jurídico a outro um problema de colisão entre regimes jurídicos nacionais e transnacionais.

Há, aqui, uma espécie de *tradeoff*: quanto maior a autonomia que as plataformas conferem à moderação de conteúdo, mais ela se constitui como um direito propriamente dito e, nesse sentido, vocaciona-se a ser autônomo também em relação às influências que lhe são externas, ganhando corpo para evitar que, por exemplo, os Estados determinem aquilo que essas empresas devem fazer. Isso explica porque faz sentido para

---

<sup>767</sup> “*Of course, these private platforms have little motivation to surrender power as in a public benefit corporation, or to adopt the rules and transparency ideas of Citron’s technological due process requirements*”. KLONICK, K. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, 2017, nota n.º 11.

as plataformas da internet estabeleceram mecanismos de autonomia do direito: mesmo que eles signifiquem um menor controle quanto àquilo que a moderação de conteúdo decide, pode valer a pena abrir mão desse poder na medida em que isso torna as decisões de conteúdo menos sujeitas a outros regimes jurídicos.

Uma vez que a questão da autonomia é colocada nesses termos, fica mais fácil ver como a adoção de procedimentos mais próximos ao do Estado se traduz em uma resposta a esse problema e, nesse sentido, a uma forma encontrada pelas plataformas para que conseguissem imunizar o que acontece no espaço que administram de intervenções externas – mesmo que a custa de perder um pouco do poder que tem sobre ele.

Como Teubner sustentou sob uma perspectiva geral, procedimentos são parte do processo de constitucionalização do direito, porque eles constroem mecanismos de autonomia do fenômeno jurídico, tornando-o imune àquilo que lhe é externo. Isso aparece, de forma mais clara, em duas características do procedimento: primeiro, na sua capacidade de construir fronteiras que diferenciam aquilo que acontece dentro e fora do procedimento; e, segundo, na sua autonomia, no sentido de que o ambiente procedimental é construído de uma tal forma que ele avança apenas de acordo com os seus próprios critérios, e não aquilo que lhe é externo.

Essas duas características aparecem, em diferentes níveis de clareza, no movimento de transformação procedimental observada no Facebook, Twitter e YouTube – e que consolidei em uma análise sustentada nos *Santa Clara Principles On Transparency and Accountability in Content Moderation*. Por meio deles, ficou claro que essas plataformas se movimentaram no sentido de conferir transparência, participação e recorribilidade aos usuários quanto às decisões sobre a moderação de conteúdo. O ponto é que esses mesmos mecanismos que, na seção anterior, fizeram – e justificaram – o procedimento como uma forma de produzir legitimidade, nesta seção têm como efeito também gerar, nesse novo formato decisório, um campo de autonomia própria, tanto em relação à estrutura empresarial da plataforma, quanto em relação a outros regimes jurídicos que lhe sejam externos.

Como afirmado na seção anterior, o mecanismo de produção de legitimidade pelo procedimento funciona trazendo as partes para dentro de um ambiente controlado, em que a controvérsia deixa de ser generalizada para ser submetida a um conjunto de etapas que culmina em uma decisão que, a princípio, é incerta – razão pela qual as pessoas se comprometem a participar do procedimento – e que as pessoas se veem na contingência de aceitar. A questão é que, para que o procedimento funcione dessa

forma, ele constrói mecanismos que o separam das influências externas, filtrando-as de modo que elas só passam impactar na decisão na medida em que seguem as regras do procedimento. Isso acontece, como disse acima<sup>768</sup>, porque se as influências externas pudessem definir a decisão que é tomada via procedimento, ninguém se interessaria em tomar parte na sua construção – uma vez que ela já seria previamente definida, não faria sentido se envolver na sua construção. A imunização das influências externas, nesse sentido, é uma espécie de efeito colateral da necessidade que o procedimento tem de garantir a incerteza da definição final – e que, no contexto de constitucionalização do direito, torna-o autônomo, fazendo com que o resultado decisório dependa apenas dele mesmo.

No caso específico da transformação detectada na moderação de conteúdo, a implantação de mecanismos de transparência, participação e recorribilidade, que trazem as partes para dentro do procedimento que define aquilo que elas podem dizer, cumpre o papel de, simultaneamente, fazê-las aceitar o seu resultado e criar barreiras que tornam essa decisão cada vez mais independente daquilo que não faz parte do procedimento. Assim, por exemplo, ao viabilizar que o usuário tenha acesso às razões que levaram uma decisão a ser tomada, trazendo-o para o procedimento, a plataforma limita a sua capacidade de influenciar arbitrariamente o seu resultado, porque agora ele precisa ser minimamente baseado nas políticas da comunidade que teriam sido violadas no caso. A legitimidade pelo procedimento anda paralela à autonomia pelo procedimento.

Como na seção anterior, o caso mais avançado dessa autonomização é o do Facebook: assim como a criação de um corpo externo e independente para analisar recursos de usuários das plataformas levou a incerteza – e a legitimidade – a outro patamar, aqui ele também é um avanço definitivo no sentido de se conferir autonomia à moderação de conteúdo em relação à plataforma, porque a desvincula estruturalmente do órgão responsável por essas decisões. Não é por acaso, portanto, que essa nova estrutura tem sido referida como um momento de constitucionalização do Facebook<sup>769</sup>, ou como

---

<sup>768</sup> Cf. *supra*, Cap. II.4, 2.3.1.

<sup>769</sup> ALANG, N. Inside Facebook's global constitutional convention. **The Week**, 12.5.2019, disponível em <<https://theweek.com/articles/840591/inside-facebooks-global-constitutional-convention>>, acesso em 9.9.2019.

uma Suprema Corte da liberdade de expressão na plataforma<sup>770</sup>: ela é a representante por excelência dessa memória institucional que o direito traz do Estado às plataformas, como um mecanismo cuja função, à semelhança de uma constituição local, é permitir que decisões sejam tomadas independentemente a respeito daquilo que as pessoas podem dizer na plataforma, imunizando-as, portanto, das influências da estrutura econômica da empresa por trás desse espaço.

Esse propósito autonomizador não precisa sequer ser depreendido desse corpo externo, mas, antes, é algo que surge textualmente das tratativas para que ele seja implantado. No último relatório disponibilizado pelo Facebook, em que reunidas as contribuições colhidas quanto a como construir essa “*Suprema Corte*”, a necessidade de se construir um corpo capaz de tomar decisões autônomas em relação à plataforma, imunizando-se em relação às influências que ela possa ter no conteúdo das decisões é a tônica dos debates. Isso aparece, em particular, pela preocupação que perpassa todo o trabalho, no sentido de que subsistam mecanismos que tornem esse órgão completamente independente da plataforma, algo semelhante com o que ocorre entre o direito e a política na estrutura dos Estados, regulando de maneira estrita as relações entre o Facebook e a sua “*Suprema Corte*”.

Assim, refletindo essa linha, discutiu-se, primeiro, como seriam escolhidas as pessoas para compor esse órgão julgador; algo sensível justamente porque isso poderia acabar conferindo ao Facebook o poder de escolher pessoas alinhadas às suas posições. Isso se colocou como uma dificuldade tanto se a plataforma escolhesse diretamente os julgadores, quanto se eles fossem escolhidos por um comitê: “*um ‘comitê de seleção intermediário para escolher o conselho poderia garantir um input externo, mas ainda deixaria com o Facebook a tarefa de ‘escolher quem escolhe’*”<sup>771</sup>. A solução encontrada foi estabelecer métodos de seleção transparente, em que as escolhas – sejam elas feitas pelo Facebook ou por terceiros – possam ser questionados em termos de qualificação

---

<sup>770</sup> READ, M. Facebook Is Going to Have a Supreme Court. Will It Work? **New York Magazine**, 30.1.2019, disponível em <<http://nymag.com/intelligencer/2019/01/facebooks-new-oversight-board-is-a-supreme-court.html>>, acesso em 9.9.2019.

<sup>771</sup> “*An intermediate ‘selection committee’ to pick the Board could ensure external input, but would still leave Facebook with the task of ‘picking the pickers.’*” FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, pp. 17/18, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

técnica – e não de alinhamento ideológico. Segundo, também foi objeto de debate se pessoas que trabalham ou não no Facebook poderiam ser parte do conselho, o que foi rejeitado precisamente porque “há uma necessidade por independência e credibilidade”<sup>772</sup> do conselho, que não seria satisfeita se alguém da própria plataforma estivesse envolvido com essas decisões – exatamente o que se quer evitar com a criação de um órgão externo.

Terceiro, ficou também definido que o conselho deve ter o poder de escolher quais casos irá decidir, e quais critérios deve adotar para priorizar uns sobre os outros. Isso é algo fundamental para garantir a independência decisória desse novo órgão, porque garante que agentes externos – sejam eles as plataformas ou outros atores – não terão o poder de definir o que ele irá julgar; antes, é o próprio colegiado – e o procedimento que ele definir – que terá o poder para isso, o que, mais uma vez, imuniza as suas decisões daquilo que lhe é externo. Assim, “o *Feedback* [dos participantes da consulta] apoiou, de uma maneira geral, que o Conselho deve reter um ‘*poder certiorari*’ de selecionar quais requerimentos de revisão ele deve julgar – em outras palavras, um Conselho que tenha a discricionariedade para ‘controlar o seu acervo’”<sup>773</sup>.

Quarto, também ficou definida a necessidade de que sejam posicionados limites estritos na relação entre o conselho e o Facebook, em particular nas situações em que esse contato é inevitável, como quando informações devem ser fornecidas pela plataforma para permitir o julgamento. Assim, argumentou-se que “os membros do Conselho [...] não devem ter autorização para antecipar as suas decisões com empregados do Facebook, e não se deve permitir que os empregados do Facebook e seus executivos interfiram em casos antes que as deliberações tenham sido concluídas e uma decisão tenha sido atingida”<sup>774</sup>. Quinto, no mesmo sentido de regular essa relação entre julgador e

---

<sup>772</sup> “This is a position that finds general acceptance, given the need for independence and credibility”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 21, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

<sup>773</sup> “Feedback has generally supported a Board that retains ‘certiorari power’ to select which requests for review that it wishes to hear — in other words, a Board that has the discretion “to control its docket”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 23, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

<sup>774</sup> “Board members,” it was argued, “should not be permitted to preview their decisions with Facebook employees ... and Facebook employees and leadership should not be permitted to provide input on cases before deliberations have concluded and a decision has been reached”. FACEBOOK. Global Feedback &



plataforma, também ficou clara a necessidade de que as decisões tomadas pelo conselho impactem no desenvolvimento das políticas de conteúdo do Facebook, vinculando-as em algum nível.

Sexto, e talvez deixando de maneira mais clara a forma como essa constitucionalização procedimental se traduz em uma autonomização do direito das plataformas em relação àquilo que ele regula – isto é, as plataformas –, foi trazido ao debate da construção do órgão julgador e posicionado como “*crítico para a sua eficiência e legitimidade*” a necessidade de que ele seja capaz de “*exercer um julgamento independente*”<sup>775</sup>. É interessante perceber como o debate é aqui colocado não apenas em termos de uma independência em relação ao Facebook, mas também quanto a quaisquer outras influências externas – algo que corrobora a hipótese de que esse movimento se justifica enquanto uma forma de fortalecimento do direito das plataformas, que se torna imune não apenas quanto às próprias plataformas, mas também quanto ao que lhe é externo. A menção aos regimes dos Estados, inclusive, é explícita:

*“O esboço da Carta questionou como o Facebook pode garantir um julgamento independente, o que ele define como [evitar] ‘influência não apropriada ou indevida do Facebook ou de qualquer outra fonte’. Feedbacks a esse respeito tiveram como foco primário três questões principais: a compensação dos membros do Conselho, a relação do Conselho com o Facebook e a influência pelos Estados”*<sup>776</sup>.

---

Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 23, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf>>, acesso em 9.9.2019.

<sup>775</sup> “*The Oversight Board’s ability to exercise independent judgment will be critical for its efficacy and legitimacy*”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 30, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf>>, acesso em 9.9.2019.

<sup>776</sup> “*The Draft Charter has asked how Facebook can ensure independent judgment, which it defines as “inappropriate or undue influence from Facebook or any other external sources.”. Feedback in this regard has focused primarily on three main issues: compensation of Board members, the Board’s relationship with Facebook, and influence by States*”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 31, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf>>, acesso em 9.9.2019.

A questão, portanto, aparece já posta nesse momento de formação de um órgão decisório externo como dividida em assegurar a sua independência, por um lado, em relação à plataforma; e, por outro, em relação aos regimes dos Estados. Quanto ao primeiro, o período de consulta indicou opiniões no sentido de que devem haver regulações – isto é, uma constitucionalização procedimental – que esclareçam como as relações entre plataforma e órgão decisório devem ocorrer; quanto ao segundo, chegou-se apenas à conclusão de que podem haver problemas de “*influência por governos de Estados poderosas e mesmo a possibilidade de captura estatal*”<sup>777</sup> – um ponto que, como detalho a seguir, em termos de continuidade deste trabalho, permanece aberto a discussões e constitui um dos problemas centrais ainda não resolvidos dos regimes jurídicos transnacionais: os seus conflitos com os regimes não-transnacionais.

Há outros exemplos de como a construção de um organismo decisório à parte, que leva ao limite a transformação procedimental da moderação de conteúdo, é uma evidência do papel autonomizador do procedimento – que explica, em um contexto de constitucionalização da moderação de conteúdo, porque essa transformação está acontecendo: discussões sobre a possibilidade de o órgão decisório definir as bases normativas que utilizará para julgar<sup>778</sup>, ou de definir a sua própria competência<sup>779</sup>. O que detalhei acima, entretanto, é o suficiente para deixar claro que a transformação imposta, portanto, pelo Facebook, em maior grau, e por outras plataformas, como o Twitter e o YouTube, é a de um procedimento que gera autonomia tanto internamente, quanto externamente, o que explica o porquê ele está sendo adotado: por meio dele, espera-se ser possível imunizar as decisões da plataforma em relação àquilo que lhe é externo, virando o jogo dessa relação de uma subordinação para um conflito entre regimes jurídicos.

Em conclusão: além da relação entre procedimento e legitimidade, o movimento de constitucionalização do direito transnacional – ou seja, da criação de

---

<sup>777</sup> “[...] *influence from powerful state governments and even the possibility of state capture*”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 32, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

<sup>778</sup> FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, pp. 33/35, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

<sup>779</sup> FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, pp. 35/36, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

mecanismos de autonomização do fenômeno jurídico – explica o que acontece nas plataformas da internet, na medida em que identifica, por um lado, um *problema de autonomização* – que se desdobra tanto em relação à estrutura empresarial das plataformas, quanto em relação a outros regimes jurídicos, como o do Estado –; e, por outro, em uma *resposta procedimental*, em que as ferramentas do procedimento são adotadas como uma forma de autonomizar a moderação de conteúdo em relação a todo o resto. Assim, *a transformação observada no sentido de se mudar a forma como decisões são tomadas pela moderação de conteúdo, aproximando-as do formato estatal, também é explicada como a replicação de um mecanismo de autonomia do direito no âmbito transnacional.*

### 3.4 Os próximos passos

Com a construção das relações entre procedimento e legitimidade e procedimento e autonomia, as duas conclusões das seções anteriores ajudam a entender o movimento atual de transformação da moderação de conteúdo das plataformas da internet. Nesta seção, meu objetivo é propor, como uma forma de encerrar este trabalho, um passo além: entender em que medida essas conclusões permitem aventar cenários futuros quanto à moderação de conteúdo, isto é, caminhos que, a partir do quadro teórico funcionalista, podem ser trilhados daqui em diante para explicar o direito das plataformas.

Primeiro: ao longo deste trabalho, aventei e testei a hipótese de que as plataformas da internet passam por um processo de constitucionalização, que explica, ao menos em parte, a adoção de procedimentos semelhantes aos do Estado para tomar decisões vinculantes, porque funcionam como mecanismos de autonomização do fenômeno jurídico. Como Teubner argumenta, entretanto, esse é apenas um dos lados do processo de constitucionalização, que, também na esfera transnacional, traduz-se “*de um lado, em disposições sobre o estabelecimento e exercício da elaboração de decisões (regras organizacionais e procedimentais) e, de outro, a definição de liberdades individuais e autonomias sociais (direitos fundamentais)*”<sup>780</sup>. A constitucionalização, nesse sentido, é tanto um projeto *procedimental* quanto *material*.

---

<sup>780</sup> “[...] *provisions on the establishment and exercise of decision making (organisational and procedural rules) on the one hand and definitions of individual freedoms and societal autonomies (fundamental rights) on the other*”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 19, disponível em

Assim, se a primeira parte desse processo de constitucionalização – a construção procedimental da autonomia do direito das plataformas, baseada na adoção de um modelo semelhante ao que garantia essa independência em relação aos Estados – está em curso, é de se esperar que os seus existam movimentos relacionados à segunda parte do processo, ou seja, a constitucionalização de caráter material. Novamente, o pioneiro nesse movimento é o Facebook – algo compreensível, uma vez que ele também é a plataforma que mais avançou no processo de constitucionalização procedimental. Assim, o último documento disponível a respeito do mecanismo externo de revisão de suas decisões reconhece que uma parte importante do trabalho de implantá-lo será a elaboração de uma espécie de constituição, em que dispostos os valores que irão orientar a decisão:

*“O Conselho irá precisar ter uma fundamentação para o seu processo de elaboração de decisões. O esboço da Carta indicou que isso indicaria um conjunto de valores, mas não determinou quais seriam. Em resposta, alguns pediram por uma ‘constituição’ ou um ‘compromisso de valores’ – alguma forma de afirmação de princípios de ordem mais geral. Outros urgiram que qualquer lista de valores deve ser clara em estabelecer prioridades, para a hipótese de que diferentes valores venham a entrar em conflito. Ainda, um grande contingente de organizações de direitos humanos pediram que o Conselho simplesmente adote o direito internacional dos direitos humanos como uma base para a elaboração das decisões. Outros expressaram que isso não forneceria a necessária clareza, e apontaram para a existência de diferentes interpretações de normas de direitos humanos por diferentes regiões e países”<sup>781</sup>.*

Sob a perspectiva da autonomia, isso parece acontecer porque possibilitar que um procedimento seja adotado como uma forma de autonomizar o direito das

---

<<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

<sup>781</sup> “The Board will need to have a foundation for its decision-making. The Draft Charter indicated that this would include a set of values, but did not determine what those would be. In response, some have called for a “constitution” or “values commitment” — some form of statement of higher-order principles. Others urged that any list of values be clear on prioritization, in cases where different values come into conflict. Meanwhile, a strong contingent of human rights organizations urged that the Board simply adopt international human rights law as its basis for decision-making. Others felt this would not provide the necessary clarity and pointed to different interpretations of human rights norms across different regions and countries”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 33, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

plataformas daquilo que lhe é externo faz pouco sentido se os critérios normativos não são, eles mesmos, criados independentemente dos demais. Ou seja: é importante para as plataformas fixar, elas mesmas, os valores constitucionais que serão aplicados nesse novo procedimento, justamente para que ninguém mais preencha esse espaço. Autonomia, nessa direção, se traduz tanto em uma perspectiva procedimental – mais avançada, e justificada neste trabalho – quanto material – ainda em desenvolvimento, mas que deve vir à luz junto com esse novo corpo decisório do Facebook.

Perceber isso é importante, porque coloca uma série de desafios – alguns deles expostos na consulta feita pelo Facebook – já enfrentados pelos processos de constitucionalização locais, mas que terão que ser ressignificados nessa esfera transnacional – e, por isso, devem pautar os próximos movimentos desse debate. Quem deve escolher os valores que serão protegidos nessa constituição? Como contrabalançar a atuação dos principais atores envolvidos na moderação de conteúdo – as plataformas e os usuários? Deve haver mecanismos de alteração desses valores? Qual deve ser o impacto de instrumentos normativos externos nessa constituição? Plataformas como o Facebook, ao elaborá-la, estão vinculadas a compromissos internacionais semelhantes aos do Estado, como os relacionados à proteção dos direitos humanos?

Segundo: como argumentei neste capítulo, o processo de constitucionalização das plataformas não é um movimento isolado, mas parte de um problema maior, vinculado à transição de direito da esfera local para a esfera transnacional. Na medida em que o direito perde os seus laços com a estrutura política local e passa a regular âmbitos sociais diferenciados em esfera global – como é o caso da moderação de conteúdo –, ele tem a sua autonomia ameaçada, porque ele deixa para trás tanto o Estado, quanto os mecanismos que tinha desenvolvido para se imunizar das influências políticas. Esses mecanismos permanecem, entretanto, com uma memória institucional, que precisa ser reimaginada em âmbito transnacional – algo que se tem presenciado na formação do órgão decisório externo do Facebook, que concretiza um movimento de constitucionalização *procedimental* – que analisei neste trabalho – e *material* – que sugeri como perspectiva de continuidade.

A constitucionalização, entretanto, é apenas um dos instrumentos locais disponíveis à esfera transnacional do direito, e que é aqui aplicada, especificamente, à moderação de conteúdo. Ela é um modelo inicial, porque “*tiveram sucesso em*

*efetivamente limitar a totalização do poder político*”<sup>782</sup>, mas nada indica que esse movimento seja o suficiente para legitimar e autonomizar o direito das plataformas diante das pressões que enfrenta da estrutura econômica em que inserido e dos regimes que lhe são externos, como os do Estado. Faz sentido, portanto, perguntar: quais são os próximos movimentos desse direito transnacional? De que forma ele mobilizará novos elementos dessa memória institucional nas plataformas para garantir a sua autonomia? Essas tentativas terão sucesso?

Os instrumentos que construí neste trabalho, nesse sentido, são úteis também para analisar esses novos desdobramentos dessa questão, que podem tomar forma, por exemplo, no avanço da constitucionalização para outros âmbitos das plataformas – na elaboração do último relatório sobre a “*Suprema Corte do Facebook*”, já surgiram questionamentos quanto a se “*o conselho poderia decidir também sobre outros assuntos da política [da plataforma], como o ranking criado por algoritmos, privacidade, direito local, inteligência artificial, monetização, propaganda política e viés [bias]*”<sup>783</sup>, rapidamente respondidos negativamente pelo Facebook. Em um sentido mais radical, novos instrumentos de legitimidade e autonomização podem ser trazidos pelas plataformas aos Estados, notadamente a construção democrática das decisões. Como no caso dos procedimentos judiciais, mimetizados na moderação de conteúdo, a existência de procedimentos eleitorais e legislativos é algo que, simultaneamente, confere legitimidade e autonomia à decisão, porque viabiliza, paulatinamente, escolhas independentes da plataforma e dos demais atores interessados em sua política de conteúdo – como os Estados. Assim como no caso analisado neste trabalho, entretanto, isso traz consigo um custo, materializado na necessidade de que as plataformas estejam dispostas a abrir mão de parte do seu poder – como o fizeram quanto à moderação de conteúdo em sentido estrito –, porque só na medida em que essas decisões forem autônomas em relação a si mesma, elas também serão autônomas em relação aos demais atores. A questão, portanto, é saber se essas empresas aceitarão ou não pagar esse custo.

---

<sup>782</sup> Cf. *supra*, nota n.º 695.

<sup>783</sup> “[...] *whether the Board could also hear other policy issues, such as algorithmic ranking, privacy, local law, AI, monetization, political ads, and bias*”. FACEBOOK. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, p. 35, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 9.9.2019.

Terceiro: até então, os conflitos que surgiram entre o direito das plataformas e o direito dos Estados eram resolvidos de acordo com a lógica da empresa – isto é, econômica – e não a jurídica. Um exemplo disso é o caso da negação do Holocausto, que opôs, de um lado, a regra do Facebook de manter esse tipo de publicação; e, de outro, as leis de países que proibiam que esse tipo de conteúdo fosse veiculado. Como o vazamento de informações a respeito da plataforma indicou<sup>784</sup>, a solução encontrada teve contornos pragmáticos e econômicos: nos países em que a existência de regra oferecia um risco à operação do Facebook – ou seja, poderia impedir que ele atuasse e, por consequência, tivesse lucro –, foram implantados mecanismos de geobloqueio, impedindo que os usuários daqueles lugares tivessem acesso ao conteúdo proibido. Nos países em que, mesmo havendo a proibição, ela não oferecia riscos ao lucro da plataforma; bem como nos demais, em que inexistiam disposições normativas a esse respeito, o conteúdo permanece disponível.

Como esse tipo de problema será tratado à medida em que avança o ciclo de constitucionalização da moderação de conteúdo das plataformas da internet? Em particular: uma vez implantados instrumentos de autonomia procedimental e material, o que irá acontecer se a conclusão a que chegar a plataforma, quanto à permissibilidade de um conteúdo, for contrária à de um Estado? Diferente do que ocorreu no caso da negação do Holocausto, a tendência com a implantação de mecanismos de constitucionalização procedimental e material é que os regimes jurídicos das plataformas sejam robustecidos, colocando-os cada vez mais em pé de igualdade com os dos Estados – com a vantagem de que, diferente dos regimes locais, eles não são limitados por fronteiras para atuar. Isso pode escalar a um nível ainda maior se, como propus no item anterior, forem estabelecidos mecanismos de representatividade democrática para as decisões das plataformas, surgindo, com isso, uma colisão entre níveis de legitimidade decisória: o que é mais legítimo, o valor escolhido pelos usuários de uma plataforma ou pelos nacionais de um país?

É difícil responder a essas perguntas: esse cenário pode se desdobrar tanto na utilização, por parte dos Estados, de mecanismos de coação contra as plataformas, obrigando-as a respeitar as legislações locais – como ocorreu, recentemente, quanto às ordens judiciais de acesso às conversas no aplicativo WhatsApp, que ocasionaram o bloqueio do aplicativo no Brasil –; quanto em graus cada vez maiores de resistência da empresa de tecnologia, firmando que detêm um regime jurídico próprio e independente do

---

<sup>784</sup> Cf. *supra*, Cap. I.1, 3.2.2.3.

estatal – um argumento robustecido pela autonomia que lhe é conferida pelo processo de constitucionalização e que, como disse, pode ser ainda fortalecido por mecanismos de legitimidade democrática, que podem fazer frente à conferida pelos nacionais de um Estado.

À semelhança do problema de constitucionalização, também aqui Teubner parece indicar um caminho: em um texto de 2004<sup>785</sup>, ele discute com Fischer-Lescano precisamente as consequências da inviabilidade de se ter, em uma realidade transnacional, uma unidade de normatizações. Na medida em que o fenômeno jurídico deixa de ser local, os conflitos entre ordens jurídicas deixam de ser uma realidade exclusiva dos Estados e passam a acontecer também entre fenômenos jurídicos que, tais como a *lex mercatoria*, os códigos corporativos e a moderação de conteúdo, regulam não um território, mas um setor social funcionalmente diferenciado. Em síntese, “*um formato totalmente diferente de direitos dos conflitos de irá emergir; um ‘direito dos conflitos intersistêmicos’, derivado não de colisões entre as distintas nações do direito internacional privado, mas de colisões entre diferentes setores sociais globais*”<sup>786</sup>. Uma saída possível para esse problema – aventada por Teubner e Fischer-Lescano – é, tal como no pluralismo transnacional e na constitucionalização do direito transnacional, mobilizar uma forma adaptada da Teoria dos Sistemas para compreender que tipo de solução pode ser construída nesses casos – e, principalmente, quais são as suas limitações. Se, por um lado, isso significa encarar a realidade e constatar a inviabilidade de se estabelecer uma autoridade hierárquica superior capaz de organizar esses conflitos – Teubner fala em um “*lasciate ogni speranza*”<sup>787</sup> quanto a essa possibilidade –, por outro, a Teoria dos Sistemas ilumina a necessidade de que esses regimes fragmentários construam, dentro deles mesmos, mecanismos que expressem uma necessidade mínima de compatibilização entre si. Isso não significa esperar que esses regimes abram mão de suas respectivas posições antagônicas, mas sim

---

<sup>785</sup> FISCHER-LESCANO, A., TEUBNER, G. Regime-Collisions: The Vain Search for Legal Unity in the Fragmentation of Global Law. **Michigan Journal of International Law**, v. 25, n. 4, pp. 999/1046, 2004.

<sup>786</sup> “[...] *a wholly new form of conflicts law will emerge; an ‘intersystemic conflicts law,’ derived not from collisions between the distinct nations of private international law, but from collisions between distinct global social sectors*”. FISCHER-LESCANO, A., TEUBNER, G. Regime-Collisions: The Vain Search for Legal Unity in the Fragmentation of Global Law. **Michigan Journal of International Law**, v. 25, n. 4, 2004, p. 1.000.

<sup>787</sup> FISCHER-LESCANO, A., TEUBNER, G. Regime-Collisions: The Vain Search for Legal Unity in the Fragmentation of Global Law. **Michigan Journal of International Law**, v. 25, n. 4, 2004, pp. 1.017 e ss.



que elas sejam em alguma medida harmonizáveis por ferramentas que eles próprios dispõem – como é o caso da possibilidade das plataformas da internet de terem as próprias políticas, ao mesmo tempo em que aplicam, em determinadas porções territoriais, regras dos Estados locais, reintroduzidas na forma, por exemplo, do geobloqueio de certos conteúdos.

Quarto – e em continuidade com o item anterior: com o processo de constitucionalização, que confere graus crescentes de autonomia às plataformas e aos seus respectivos regimes jurídicos, surge a questão de como aquilo que se desenvolve externamente a esses ambientes – como a realidade política, que persiste como um fenômeno essencialmente local e territorialmente delimitado – se relaciona com o mundo digital.

Um exemplo de como isso se coloca como um problema é a forma como os espaços digitais são organizados. Se é verdade que a globalização é uma “*oportunidade para o direito institucionalizar constituições para vilas globais de setores socialmente autônomos, com uma relativa distância da política e da economia*”, é um fato a ser observado de que forma esses novos ambientes se compartimentam, isto é, quais são os critérios utilizados para dividi-los, em particular porque, segundo Teubner, “*dentro dos setores autônomos, há potenciais para a re-politicização, a re-regionalização e a reindividualização do processo de elaboração de normas*”<sup>788</sup>. A internet, nesse sentido, e, em especial, as plataformas, trazem um conjunto novo de discussões, na medida em que estabelecem critérios de divisão do espaço virtual global por critérios não associados às fronteiras políticas e econômicas, mas que têm, por exemplo, relação com a privacidade: em todas as plataformas que mencionei neste trabalho, há a possibilidade de que sejam alteradas as configurações de visualização das informações pessoais dos usuários, restringindo-as. Assim, é possível estabelecer, no Facebook, que apenas pessoas autorizadas irão visualizar as publicações de um usuário; no Twitter, que apenas aqueles que seguem um determinado perfil irão ter acesso às suas mensagens e, no YouTube, que os vídeos ficarão disponíveis para um público restrito. Cada usuário tem um espaço de liberdade em que pode construir fronteiras em relação àquilo que diz e as outras pessoas, criando uma secção dessa vila global para si.

---

<sup>788</sup> “*Innerhalb der Autonomiebereiche gäbe es Potentiale für eine Repolitisierung, Reregionalisierung und Reindividualisierung von Normbildungsprozessen*”. TEUBNER, G. Privatregimes: Neo-Spontanes Recht und duale Sozialverfassungen in der Weltgesellschaft, in SIMON, D., WEISS, M. (Ed.) **Zur Autonomie des Individuums: Liber Amicorum Spiros Simitis**. Baden-Baden: Nomos, 2000, pp. 437/453 (n. p.).

Simultaneamente a esses critérios específicos da tecnologia da internet, é possível, como menciona Teubner, que aspectos políticos e econômicos sejam reinseridos nesses ambientes, reproduzindo clivagens que lhe são externas e, assim, produzindo o que ele chama de “*re-politicização e re-regionalização*”. Isso pode acontecer tanto espontaneamente, a partir dos próprios usuários, quanto partindo da plataforma: por um lado há comunidades no Facebook – ou seja, espaços de discussão da plataforma – que aceitam apenas pessoas de uma determinada nacionalidade, reinserindo no ambiente virtual uma separação política que, ao menos a princípio, não faz parte dela; outras, são restritas a pessoas com determinado poder aquisitivo, reintroduzindo divisões de caráter econômico no mundo digital.

Por outro, como o caso do geobloqueio das publicações que negam o Holocausto expôs, é possível que as plataformas restrinjam o acesso de usuários localizados em um país a um determinado conteúdo, criando, quando às manifestações que vão à plataforma, um espaço à parte. Como esses espaços irão se relacionar entre si? De que modo essas clivagens políticas e econômicas, reinseridas no ambiente virtual, se submetem à realidade digital, ordenada pelas plataformas? Faz sentido, por exemplo, que uma plataforma dissolva uma comunidade que só aceita israelenses, mas não palestinos? O mesmo raciocínio se aplica a uma comunidade que só aceita gente rica? Como julgamentos a respeito dessas separações, feitos pela moderação de conteúdo, serão vistos – e, talvez, influenciados – pelas estruturas políticas e econômicas localizadas além da plataforma?

Essas perguntas – e as que fiz mais acima – fogem do escopo deste trabalho, mas expõem perspectivas para a sua continuidade, em particular porque problemas como esses – a constitucionalização material, a mobilização e a reimaginação de outros instrumentos da memória institucional jurídica, a colisão de regimes nacionais e transnacionais, e a reinserção de divisões externas no espaço transnacional – são questões tratadas pelo pluralismo funcionalista e, por isso, assim como o papel do procedimento na estrutura social moderna, podem ser reinterpretadas por esse quadro teórico, fornecendo indicativos do que está por vir. Nesse sentido, mais do que explicar a transformação que ocorre atualmente quanto aos procedimentos adotados para se moderar conteúdo, a expectativa deste trabalho foi também a de contribuir para mostrar como esse direito das plataformas pode ser interpretado a partir de uma estrutura sociológica robusta – como o funcionalismo estrutural –, e como fazer isso colabora para esclarecer os caminhos que a moderação de conteúdo percorre – e está ainda por percorrer.

#### 4. CONCLUSÃO

Iniciei o capítulo com um objetivo: oferecer respostas ao problema proposto na segunda parte da dissertação. Para fazer isso, propus duas etapas: primeiro, apurar as ferramentas do funcionalismo acessadas no Capítulo 4, tornando-as aptas a analisar um contexto diferente daquele que foi tomado como base para a sua elaboração. Segundo, utilizar essas ferramentas para testar as hipóteses que apresentei a respeito da transformação procedimental pela qual passa a moderação de conteúdo das plataformas da internet – que, analisada com o que se tinha de disponível, permanecia inexplicada.

Na primeira etapa, que corresponde à primeira seção, a proposta foi fornecer uma versão aperfeiçoada do funcionalismo estrutural. Em particular, minha intenção foi apurar as ferramentas da Teoria dos Sistemas que se demonstraram promissoras para solucionar o problema da segunda parte desta dissertação mas que, como expus antes, têm limitações: como, ao menos à primeira vista, elas partem de uma associação entre o direito e o Estado, – contexto que Luhmann tinha à mão na época –, mostrou-se necessário, no mínimo, reinterpretá-las, para compreender em que medida elas teriam alcance para apreender uma situação que envolvesse decisões não vinculadas ao Estado – como é o caso das plataformas da internet.

Esse passo adiante do funcionalismo foi dado desdobrando, em três etapas, a abordagem que Gunther Teubner faz da Teoria dos Sistemas. Na primeira, expus o problema que ele se propõe a enfrentar: a necessidade de se explicarem fenômenos normativos que, a exemplo da *lex mercatoria* e dos códigos corporativos, decorriam dos impactos que da globalização no direito. Em específico, Teubner identifica nessas situações casos em que fenômenos muito próximos ao jurídico – em que há regras, processos, competências definidas, órgãos fixados para tomar decisões vinculantes, etc. – surgiram na esfera transnacional, de forma desvinculada dos Estados. A questão passa a ser saber, nesse contexto, se o funcionalismo estrutural é capaz de fazer frente a esses fenômenos, ou, então, se é limitado a interpretar o direito como um acontecimento essencialmente local.

Teubner propõe, então, uma solução a esse problema, que baseei, acima, em uma reconstrução que demonstra que tanto o conceito inicial de direito na Teoria dos Sistemas, ainda desvinculado da teoria social de Luhmann; quanto a sua versão posterior, já inserida em uma leitura comunicativa da sociedade, não exigem que direito e Estado estejam vinculados para que façam sentido. Fica claro, assim, que a relação percebida em

*Legitimação pelo Procedimento* entre ambos é contingente, o que abre um flanco de atuação, no sentido de se acoplar a Teoria dos Sistemas a uma perspectiva do direito não centralizada no Estado, viabilizando que as suas ferramentas sejam estendidas para contextos como os da *lex mercatoria* e dos códigos corporativos.

Em seguida, Teubner cumpre essa tarefa jungindo Teoria dos Sistemas e pluralismo jurídico em um *pluralismo funcionalista*, que conceitua o direito a partir de duas características: primeiro, pela sua capacidade de atribuir a acontecimentos sociais a marca do lícito / ilícito; e, segundo, pela sua capacidade de atuar reflexivamente, isto é, de submeter as regras que cria a elas mesmas, de modo que o próprio direito seja capaz de definir, por exemplo, em que medida uma norma sua é ou não válida. Ao mesmo tempo em que esse conceito capta o essencial do direito, ele não depende de um vínculo com uma estrutura política para fazer sentido – e, por isso, ele é capaz de conferir corpo teórico ao pluralismo jurídico, fazendo dele um instrumento para compreender esses fenômenos normativos transnacionais – exemplificados por Teubner como a *lex mercatoria* e os códigos corporativos e, trazidos por mim, na seção seguinte, na forma da moderação de conteúdo das plataformas da internet.

Por fim, na terceira etapa, colhi os benefícios dessa solução, especificando-os para o problema que pretendo resolver. Para tanto, tentei entender em que medida a relação entre procedimento e legitimidade, extraída por Luhmann de um contexto em que direito e Estado ainda estavam fortemente vinculados, era transformada pelo pluralismo funcionalista. Esse diagnóstico começou com a percepção de que os fenômenos jurídicos transnacionais estão em formação, e que têm as suas características tensionadas pela ausência da estrutura política do Estado, que é substituída, por exemplo, por estruturas econômicas – como os contratos empresariais, na *lex mercatoria*, e as empresas, no caso dos códigos corporativos.

Assim, por um lado, a relação entre procedimento e legitimidade se mantém, com alguns tensionamentos. Como acontecia nos contextos políticos locais, no direito transnacional é preciso explicar a razão pela qual as pessoas obedecem, e, também aqui, essa prontidão generalizada à aceitação da decisão alheia – isto é, a legitimidade – pode ser compreendida a partir dos procedimentos. Isso porque, também em contextos transnacionais, a coação e o consenso não se mostram como alternativas viáveis – sobretudo porque o fenômeno jurídico agora não se refere mais a uma comunidade territorialmente delimitada, que por vezes compartilhava valores –, e o procedimento se apresenta como uma forma viável de gerir a aceitação, trazendo as partes a uma posição

em que elas estejam na contingência de acolher uma decisão, mesmo que ela lhes seja desfavorável.

Como expus acima, isso não significa que essa relação permanece incólume. Isso porque parte dos mecanismos procedimentais de legitimidade estavam ancorados em formas de representatividade política que não existem mais. A legitimidade pelo procedimento de Luhmann é, na verdade, uma legitimidade pelos *procedimentos* – que dependem, ao menos quanto à escolha eleitoral e legislativa, de uma estrutura política representativa que não existe, via de regra, no direito transnacional. Há uma sobrecarga sobre o direito, que não tem mais como externalizar esse processo de legitimação para a política e, por isso, se vê em uma “*situação de auto-legitimidade*”<sup>789</sup>, em que depende dos próprios procedimentos para produzir a legitimidade.

Por outro, o papel do procedimento é ressignificado. Como o direito se encontra em uma situação de vulnerabilidade, sem o seu correspondente político, ele fica exposto a influências dos outros âmbitos aos quais se acopla na realidade transnacional, que colocam em risco a sua autonomia. Para se resguardar em relação a isso, ele replica na esfera transnacional mecanismos de constitucionalização, que são reinterpretados por Teubner. Aqui, ele opera um movimento semelhante ao feito em relação ao direito, retirando o fenômeno constitucional de seus limites territoriais, e o reinterpretando funcionalmente para que faça sentido também fora do que ocorre no Estado. Isso é feito, basicamente, entendendo a constitucionalização como uma forma de autonomização do direito em relação àquilo que ele regula, construindo um passo adiante em relação ao pluralismo funcionalista em que o procedimento tem um papel central. Como ele constrói barreiras em relação ao que não faz parte dele, ele viabiliza que o direito se construa de forma independente àquilo que ele regula, constitucionalizando-se ao se tornar uma esfera própria de produção e avaliação de razões jurídicas. Isso significa, por um lado, fazer aquilo que ele regula independente, mas, por outro, constituir-se também de maneira independente.

A segunda parte da estratégia foi trazer essas duas relações para a realidade das plataformas, entendendo em que medida elas são aplicáveis nesse contexto e, principalmente, são aptas a testar as hipóteses que posicionei a respeito da transformação pela qual passa a moderação de conteúdo. Fiz isso, primeiro, inserindo a internet na realidade transnacional, a partir de três características que a marcam – digitalização,

---

<sup>789</sup> Cf. *supra*, Cap. II.5, 2.2.3.

privatização e globalização. Em seguida, a partir de um paralelo entre, de um lado, *lex mercatoria* e códigos corporativos; e, de outro, a moderação de conteúdo, foi possível defini-la como o direito das plataformas, isto é, um fenômeno jurídico transnacional que observa acontecimentos sociais a partir do código lícito / ilícito e que tem níveis diferentes de reflexividade.

A partir dessa constatação, foi possível trazer as relações entre procedimento e legitimidade e procedimento e autonomia como formas de explicar a transformação implantada pelas plataformas da internet no procedimento de moderação de conteúdo. Concretamente, isso foi compatibilizado com a realidade da moderação de conteúdo, por um lado, mostrando que há, nas plataformas, um problema de legitimidade; e que esse problema de legitimidade é solucionado, ao menos em parte, pela instituição de procedimentos capazes de conferir legitimidade às suas decisões e, principalmente, evitar a generalização das controvérsias envolvendo as suas decisões de moderação de conteúdo. Por outro, e de forma paralela, a segunda relação foi demonstrada expondo que há, nas plataformas, também uma questão de autonomia – aqui relacionada não apenas à atuação das plataformas na moderação de conteúdo, mas também quanto à moderação de conteúdo e aquilo que lhe é externo, bem como que os procedimentos são um indicativo de uma constitucionalização procedimental, que confere às decisões, autonomia.

Por fim, em último movimento, essas constatações foram trazidas como indicativos dos próximos passos das plataformas, primeiro, quanto à existência de um processo de constitucionalização material, em que valores terão que ser definidos como base para a moderação de conteúdo; segundo, quanto à possibilidade de que novos instrumentos locais sejam reimaginados para a realidade transnacional, promovendo uma constitucionalização mais ampla, ou, por exemplo, o início de movimentos de legitimidade democrática; terceiro, quanto à forma como os conflitos entre o regime jurídico das plataformas e o regime jurídico dos Estados serão resolvidos; e, quarto, quanto à forma como organizado esse espaço das plataformas, em particular quanto à existência de movimentos de re-politização que inserem, no mundo digital, os conflitos e as divisões da realidade política local.

## CONCLUSÃO

Em um sentido amplo, o objetivo deste trabalho foi entender como as plataformas da internet regulam a liberdade de expressão. Para fazer isso, minha proposta – também entendida em um sentido geral – foi mobilizar instrumentos que esclarecessem tanto *como funciona* o mecanismo regulatório que elas construíram para isso – a moderação de conteúdo – quanto *por que* ele tem se transformado nos últimos anos, transitando de um modelo decisório arbitrário para um baseado na forma como o Estado toma decisões vinculantes – mesmo que, à primeira vista, isso não faça sentido.

Essa proposta foi organizada em um trabalho de duas partes: “*Por uma descrição*” e “*Por uma reinterpretação*”. A primeira parte dedicou-se a resolver um *problema de opacidade*: como as plataformas da internet são estruturas privadas, os instrumentos regulatórios da liberdade de expressão utilizados por elas – consolidados na ideia de moderação de conteúdo – não estão sujeitos aos mesmos protocolos de transparência que, via de regra, são observados na regulação estatal. Por isso, não há um referencial único e seguro a partir do qual as normatizações e os procedimentos da moderação de conteúdo possam ser compreendidas; antes, essa forma de regulação é marcada por uma *opacidade* – algo que é um problema em si, na medida em que inviabiliza que problematizações e hipóteses mais consequentes a respeito do funcionamento desse mecanismo sejam elaboradas. Disso decorreu a proposta da primeira parte do trabalho: explicar tanto *por que* faz sentido estudar a moderação de conteúdo (Capítulo 1) quanto *e o que é* a moderação de conteúdo (Capítulo 2).

O Capítulo 1, nesse sentido, oferece uma justificativa para este trabalho. Ele se dedica a responder à seguinte pergunta: por que faz sentido tomar a moderação de conteúdo das plataformas da internet como o tema de uma dissertação? Minha proposta foi reconstruir o percurso traçado por um dos autores de referência na regulação da liberdade de expressão: Jack Balkin. Em síntese, Balkin escreve, em 2014, o texto *Old-School/New-School Speech Regulation*, em que ele oferece uma contribuição à pergunta de como as mudanças tecnológicas impactam na regulação da liberdade de expressão. Ele argumenta, nesse sentido, que as transformações nas telecomunicações fizeram surgir uma “*era digital*”, que, se reflete na liberdade de expressão na medida em que provoca uma transição específica entre as *infraestruturas da liberdade de expressão*. Para Balkin, a infraestrutura da liberdade de expressão é o contexto institucional que viabiliza às pessoas se manifestarem, algo que, no mundo pré-digital, era associado às mídias impressas, à

disponibilização de espaços de manifestação aos editores e ao protagonismo de conglomerados de mídias impressas; e, no mundo digital, passa a ser vinculado à internet, à disponibilização de espaços de manifestação para os usuários de serviços como as plataformas digitais e, por isso, é protagonizado por conglomerados de tecnologia.

Como os instrumentos regulatórios utilizados pelo Estado para restringir a liberdade de expressão das pessoas frequentemente atuam sobre essa infraestrutura, a era digital significou não apenas uma transformação nesse pano de fundo institucional, mas uma mudança também na própria forma como a liberdade de expressão é regulada. Há, nas palavras de Balkin, uma transição entre uma regulação da liberdade de expressão “*old-school*” – caracterizada pelo direcionamento à mídia tradicional, utilização de restrições prévias e mecanismos repressivos, que permaneciam visíveis – para uma “*new-school*” – direcionada para os intermediários da internet, via restrições prévias, mecanismos repressivos e retributivos, que permanecem, preferencialmente, invisíveis.

Esse, entretanto, é um diagnóstico que parte de um recorte insuficiente: a era digital não significou apenas o surgimento de novos *espaços de manifestação*, mas também a criação de novos *espaços de decisão*, localizados dentro das empresas que administram as plataformas da internet, e que definem, com relativa independência, aquilo que as pessoas podem ou não dizer. Isso ficou claro a partir de três casos, em que controvérsias decorrentes da moderação de conteúdo na avaliação de uma publicação de um jornalista norueguês – o caso Egeland –, do banimento de um ativista das principais plataformas da internet – o caso Alex Jones – e da resistência do Facebook em excluir publicações que negam o Holocausto, levaram a divergências entre as plataformas e os Estados, evidenciando que elas não agem como agentes estatais – tal como pressupõe Balkin –, mas, antes, como reguladores independentes da liberdade de expressão, estabelecendo e fazendo valer as suas próprias políticas de conteúdo. Por isso, ao restringir a análise dos impactos da era digital para a regulação da liberdade de expressão àquilo que fazem os Estados – deixando de lado, portanto, a atuação de agentes privados na moderação de conteúdo das plataformas da internet –, Balkin deixou de lado a verdadeira revolução que decorreu dessas mudanças tecnológicas: não o surgimento de novos espaços de manifestação, mas a criação de novos espaços de decisão.

Isso fica evidente quando, retomando o percurso traçado pelo próprio Balkin, ele reconhece, no texto *Free Speech is a Triangle* – de 2018 – que a transição *old-school / new-school* reflete apenas uma parte das transformações que a era digital ensejou na regulação da liberdade de expressão, algo que justifica entender também – como foi a



proposta deste trabalho – de que forma agentes privados que atuam na internet passaram a definir, por critérios que lhe são próprios, aquilo que as pessoas podem dizer. Em síntese, entender a regulação da liberdade de expressão atualmente traduz-se, também, em entender como agentes privados passaram a estruturar mecanismos decisórios para exercer poder sobre as manifestações veiculadas por seus usuários nas plataformas que administram.

Esclarecidas as razões pelas quais faz sentido escrever um trabalho sobre a moderação de conteúdo, o passo seguinte é entender *o que ela é*. Esse foi o objeto do Capítulo 2, em que propus uma esquematização para esclarecer tanto como funciona esse mecanismo regulatório, quanto quais são as interpretações que temos disponíveis a seu respeito. O primeiro passo para isso foi fornecer conceitos capazes de delimitar o tema do trabalho – isto é, a moderação de conteúdo das plataformas da internet. Fiz isso em duas etapas: primeiro, definindo o que são plataformas, contextualizando-as com a evolução da internet; segundo, conceituando, dentro das plataformas, o que é a moderação de conteúdo.

Em síntese, sustentei que a internet nasce reproduzindo a assimetria que existe nas mídias tradicionais, posicionando, de um lado, um produtor de conteúdo – o dono de uma página na rede, por exemplo – e, de outro, aquele que o consome – o usuário comum. À medida que ela evolui, entretanto, essa dualidade é mitigada por um movimento que ficou conhecido como um transitar da Web 1.0 para a Web 2.0: surgem, no ambiente da rede, instrumentos que aproximam a internet de sua promessa inicial, no sentido de viabilizar um “*espaço colaborativo em que você pode se comunicar por meio do compartilhamento das informações*”<sup>790</sup>, porque permitem ao usuário comum dizer aquilo que ele pensa. Esse movimento, entretanto, teve o seu custo: ao contrário do que se pensava inicialmente, a internet não eliminou os intermediários – atores que se posicionavam entre o indivíduo e a possibilidade de ele se fazer ouvido em uma audiência mais ampla –, mas apenas promoveu uma espécie de troca de guarda, produzindo, na verdade, novos intermediários. Isso acontece porque, para que as pessoas pudessem se manifestar na internet sem deter o conhecimento técnico necessário para, por exemplo, construir uma página na rede, foram criadas as plataformas – espaços administrados por empresas de tecnologia que constroem a interface rede / indivíduo, transformando códigos de programação complexos em ferramentas intuitivas de expressão. Defini plataformas, nesse particular sentido, como serviços e páginas na internet que (i) “*hospedam, organizam e circulam, para os seus usuários, o conteúdo que eles compartilham e as interações que*

---

<sup>790</sup> Cf. *supra*, nota n.º 220.

*fazem*”; (ii) “*sem ter produzido ou comissionado (o grosso) desse conteúdo*”; (iii) “[*fazendo isso*] *com base em uma infraestrutura que [...] realiza o processamento de dados para prestar serviços ao usuário, fazer anúncios e lucrar*”<sup>791</sup>.

A segunda etapa foi definir, a partir desse conceito de plataformas, o que é a moderação de conteúdo. Começando com uma comparação com os dois paradigmas disponíveis – de um lado, as empresas de mídia, que se responsabilizam pelo conteúdo que produzem; e, por outro, as empresas de transmissão de informações, que não são por ele responsáveis, foi preciso definir a moderação de conteúdo como algo à parte, porque as plataformas nem produzem, nem apenas transmitem conteúdo. O que acontece nesses ambientes é, de alguma forma, diferente dessas situações: seja por um senso de responsabilidade corporativa, seja por verem nesse tipo de conduta uma forma de criar um ambiente mais propício à interação – e, por consequência, que lhes gera mais lucro –, o fato é que o trajeto da informação em uma plataforma não é uma linha reta entre usuários, mas um percurso de obstáculos, em que esse intermediário específico – a empresa de tecnologia que administra o espaço virtual – define onde, como e qual conteúdo será ali veiculado.

Essa noção permitiu, sob uma perspectiva geral, definir a moderação de conteúdo como o instrumento pelo qual as plataformas da internet exercem algum nível de controle sobre as manifestações de seus usuários, dispondo quanto à forma como elas virão à luz na rede sem nunca, entretanto, tomá-las como suas. Essa definição, que abrange tanto as decisões mais gerais tomadas pelas plataformas – que estabelecem, por exemplo, que irão se constituir como serviços de compartilhamento de vídeos, imagens, ou textos –; quanto as mais específicas – trazidas à luz nas controvérsias dos três casos que mencionei antes –, é complementada por outra, que traduz uma perspectiva restrita sobre a moderação de conteúdo. Sob essa visão, que é objeto da dissertação, defini a moderação de conteúdo como a prática de estabelecer e aplicar um conjunto de regras que determinam aquilo que os usuários de uma plataforma podem ou não veicular – algo que abriu caminho, na seção seguinte, para tratar de cada um desses aspectos.

Para expor o funcionamento da moderação de conteúdo em sentido estrito, tomei, como um recorte que persistiu ao longo da dissertação, aquilo que aconteceu com o Facebook, Twitter e YouTube, as maiores e principais plataformas da internet da atualidade. Assim, reconstruí a história normativa abstrata desses espaços, expondo que, à

---

<sup>791</sup> Cf. supra, nota n.º 238.

medida que a moderação de conteúdo se profissionalizou, elas tenderam a uma maior variabilidade normativa, adotando regras tanto opacas e públicas, quanto gerais e específicas. Em seguida, adotei uma abordagem semelhante quanto ao viés da moderação de conteúdo que aplica concretamente essas normas, distinguindo três fases nesse processo: a de detectar o conteúdo a ser analisado, a de tomar uma decisão a seu respeito, e a de, se o caso, impor uma sanção como consequência da violação de suas políticas de conteúdo. A partir da literatura disponível, de casos documentados em publicações especializadas em tecnologia e em vazamentos de documentos internos das plataformas, propus uma classificação que subdivide a primeira fase (detectar) de acordo com a origem – extraplataforma / intraplataforma –, o momento – *ex-ante* / *ex-post* –, o tipo de agente – proativa / reativa – e o modo – ação humana / ação automatizada; a segunda fase (decidir) de acordo com o modo – centralizado / descentralizado – e o agente decisório – manual / automatizado; e a terceira fase (sancionar) de acordo com as consequências – externas / internas.

Em seguida, reconstruí as interpretações que temos disponíveis a respeito da moderação de conteúdos, subdividindo-as em duas correntes: primeiro, a que a vê pela lente de categorias jurisprudenciais, notadamente as que existem nos precedentes da Suprema Corte dos Estados Unidos – algo que demonstrei ser problemático, porque deixa de lado especificidades importantes dessa forma de regular a liberdade de expressão; e, segundo, a que a entende como uma forma de governança privada, que avança no sentido de reconhecer a existência de um marco regulatória fora do Estado – mas que ainda não tem a robustez suficiente para responder a questões mais consequentes a respeito de como a moderação de conteúdo é e de que forma ela se transforma. Juntas, as esquematizações a respeito de como as plataformas regulam a liberdade de expressão e como essa regulação é interpretada forneceram um quadro geral da questão, capaz de tornar a moderação de conteúdo um mecanismo menos obscuro e, nesse sentido, tanto resolver o seu problema de opacidade, quanto abrir caminho para problematizações mais consequentes – como à que dediquei a segunda parte da dissertação.

Assim, o Capítulo 3, que inicia a Parte II do trabalho (“*Por uma reinterpretção*”), aproveitou-se dos aportes trazidos pela primeira parte da dissertação, construindo um problema de pesquisa sustentado na possibilidade de se ter mais claro, a partir dos capítulos anteriores, como a moderação de conteúdo funciona. Esse problema, em síntese, é de que as plataformas da internet têm transformado o modo como moderam conteúdo e isso não tem explicação nas interpretações temos disponíveis a respeito desse

mecanismo. Para expô-lo com mais clareza, propus separar os juízos que as plataformas fazem quando moderam conteúdo, em sentido estrito, em dois tipos: primeiro, em juízos materiais, ou seja, decisões que definem aquilo que os usuários poderão ou não publicar; segundo, em juízos procedimentais, as decisões que definem como é formatado o procedimento que determina se uma publicação violou ou não a política de conteúdo da plataforma. Uma vez que esses juízos são colocados lado a lado, fica perceptível a forma como as plataformas se aproximam e se afastam do da forma como o Estado regula a liberdade de expressão, de acordo com aquilo que lhes convém. Assim, se nas decisões materiais as plataformas se sustentam em um quadro geral de defesa democrática da liberdade de expressão, posicionando como a regra a ausência de intervenção naquilo que dizem as pessoas – porque assim se favorece a permanência e a interação entre pessoas; nas decisões procedimentais as plataformas se baseiam em um modelo que é tudo menos democrático, construindo procedimentos arbitrários caracterizados pela opacidade, ausência de participação de seus usuários e irrecorribilidades. Como mencionei, as plataformas sustentavam que protegiam os direitos da Constituição dos Estados Unidos tomando decisões como um Tribunal da Arábia Saudita.

O problema, nesse contexto, é que o procedimento pelo qual as plataformas moderam conteúdo está se transformando – e ninguém sabia explicar o porquê. A partir de uma análise sustentada nos *Santa Clara Principles On Transparency and Accountability in Content Moderation* – uma carta de exigências feitas por acadêmicos e membros da sociedade civil para mudar a forma como decisões são tomadas na moderação de conteúdo –, foi possível diagnosticar que, cada vez mais, as plataformas têm adotado mecanismos que aproximam o procedimento da moderação de conteúdo daquele utilizado pelos Estados para tomar decisões vinculantes, implantando instrumentos de transparência, participação dos usuários e recorribilidade. Essa mudança, entretanto, não faz sentido, tanto porque as plataformas não estão obrigadas a implantá-la, quanto porque o modelo anterior garante às empresas que administram esses espaços mais poder sobre a manifestação dos usuários.

A proposta da segunda parte da dissertação, nessa direção, foi compreender as razões por trás desse movimento, em particular desvendando o que há por trás da posição das plataformas de transformar o procedimento pelo qual moderam conteúdo. Argumentei, nesse sentido, que tanto a primeira corrente quanto à moderação de conteúdo, que a interpreta pela lente das categorias jurisprudenciais, quanto a segunda, que a vê como uma forma de governança privada, não detêm os instrumentos suficientes para resolver esse problema. Para solucioná-lo, seria necessário ir além das abordagens

vigentes, que se concentram mais em obter informações do que em organizá-las em um corpo teórico consistente, e *reinterpretar a moderação de conteúdo e seus movimentos com as ferramentas da sociologia*. Sustentei, nessa direção, que mais do que entrevistar os atores desse processo ou tentar descobrir algo em documentos vazados, seria o caso de se mobilizar os instrumentos sociológicos – em particular, os construídos pela Teoria dos Sistemas – para identificar a *função* dos procedimentos na estrutura social moderna. A ideia, nesse sentido, foi tentar entender por que o formato decisório baseado em um devido processo legal se desenvolveu no Estado para, em seguida, levar esse diagnóstico dele às plataformas, identificando qual é o seu papel na moderação de conteúdo.

Dediquei o Capítulo 4 ao primeiro passo dessa estratégia: identificar o papel que os procedimentos têm na estrutura política moderna – ou seja, nos Estados. Fiz isso reconstruindo um texto de Niklas Luhmann – *Legitimação pelo Procedimento* –, em que a sociologia jurídica é utilizada para descobrir justamente a função dos procedimentos. Após descartar perspectivas dogmático-jurídicas a respeito do procedimento, Luhmann se propõe a testar a hipótese de que os procedimentos funcionam como uma forma de se conferir legitimidade às decisões vinculantes tomadas pelo Estado. Para fazer isso, ele se baliza por uma redefinição tanto do que é um procedimento, quanto do que é legitimidade. Procedimentos, para Luhmann, são mecanismos que constroem uma decisão vinculante a partir de uma variabilidade inicial de possibilidades, fazendo isso por estruturas que garantem que o resultado desse processo é necessariamente incerto – isto é, resulta do próprio procedimento. Legitimidade, por sua vez, é a prontidão generalizada das pessoas a aceitar uma decisão, mesmo que ela lhes seja desfavorável. Luhmann constrói, nesse contexto, a hipótese de que a função dos procedimentos na estrutura política da sociedade moderna é produzir decisões legítimas, ou seja, que viabilizem a obediência das pessoas àquilo que determina o Estado mesmo quando fatores como o uso da força não estejam presentes.

O passo seguinte foi mostrar como essa hipótese é testada na prática, isto é, na realidade dos três procedimentos apontados por Luhmann como centrais na sociedade moderna: o de decisão judicial, o de escolha eleitoral e o de deliberação legislativa. Em todos esses casos, Luhmann argumenta que o procedimento ocasiona legitimidade porque constrói mecanismos que garantem a incerteza de seu resultado e que, por isso, atraem as partes interessadas para participarem de seu desenvolvimento, aproximando-os em uma lógica que as coloca em posições nas quais elas se veem, ao final, na contingência de aceitar o resultado da decisão, mesmo que ela lhe seja desfavorável. Como isso só acontece

se as partes se veem na posição de influenciar o resultado do procedimento, é indispensável para esse papel legitimador do procedimento que ele seja capaz de estabelecer fronteiras entre o que acontece dentro dele e todo o resto, de modo que o seu resultado dependa apenas daquilo que acontece dentro dele próprio.

Procedimentos, em síntese, produzem legitimidade – mas constatar isso não é o suficiente para resolver o problema das plataformas. Isso porque a moderação de conteúdo da internet ocorre em um contexto diferente daquele em que Luhmann estruturou seu argumento a respeito da relação entre legitimidade e procedimento; em particular, *Legitimação pelo Procedimento* é um texto que pressupõe que o direito é o um fenômeno essencialmente local e vinculado ao Estado. Por isso, essa conclusão não pode ser simplesmente transposta para a moderação de conteúdo, tomando-a como uma solução; antes, é preciso que sejam acionados mecanismos que façam a Teoria dos Sistemas ir dos Estados às plataformas.

Nesse sentido, o Capítulo 5 – que se propõe a testar as hipóteses quanto ao problema da segunda parte da dissertação – é dividido em duas partes. Na primeira, o diagnóstico funcionalista é levado do Estado às plataformas, a partir de uma reconstrução da releitura que faz Gunther Teubner da Teoria dos Sistemas; no segundo, essa versão própria às plataformas do funcionalismo é utilizada para testar essas hipóteses.

Quanto à primeira parte, estruturei ela ao redor da proposta, feita por Teubner, de utilizar a Teoria dos Sistemas para explicar os impactos da globalização no direito. Em síntese, ele identifica um problema na forma como alguns fenômenos normativos – notadamente a *lex mercatoria* e os códigos corporativos – são tratados: mesmo que eles se assemelhem com o direito, eles não são tratados como direito, porque não produzidos pelo Estado. Teubner propõe, então, que o pluralismo jurídico – que vê o direito como um fenômeno independente do Estado – seja robustecido teoricamente por uma versão adaptada da Teoria dos Sistemas – que não relaciona um ao outro, exceto contingentemente, como em *Legitimação pelo Procedimento*, e, por isso, é o instrumento mais adequado para esse papel. O resultado é um pluralismo funcionalista, que parte de uma definição não-estatal do direito: para Teubner, o direito é um fenômeno comunicativo que atribui a acontecimentos sociais um dos lados da característica lícito / ilícito; e que, reflexivamente, define ele próprio os seus critérios normativos.

Com essa definição em mãos, foi possível retornar à relação trazida à luz por Luhmann entre legitimidade e procedimento, para verificar em que medida ela se mantém hígida em um contexto não estatal, como o das plataformas. Assim, por um lado,

continua fazendo sentido dizer que o procedimento produz legitimidade, porque tanto o problema, quanto a solução que são endereçadas por essa relação permanecem presentes no contexto do direito transnacional. Por outro, como no contexto transnacional não há um correspondente político ao Estado, a desvinculação do direito a esse âmbito local de referência retesa e ressignifica o papel do procedimento. Mais do que um produtor de legitimidade, ele também assume, nesse novo contexto, a função de autonomizar o direito em relação às influências que lhe são externas, tomando parte em um processo de *constitucionalização procedimental*, pelo qual o fenômeno jurídico tenta se resguardar das intervenções daquilo que ele regula. Procedimentos, portanto, tanto são mecanismos produtores de legitimidade, quanto são instrumentos garantidores de autonomia do direito.

Quanto à segunda parte do Capítulo 5, ela se dedicou a trazer essa relação entre legitimidade e procedimento, agora adaptada às plataformas, para o problema da dissertação, primeiro, expondo como a internet e as normatizações que as regulam se inserem nesse contexto do direito transnacional; segundo, demonstrando que a moderação de conteúdo é uma expressão jurídica, isto é, é o direito das plataformas; e, terceiro, demonstrando que as relações entre procedimento e legitimidade e procedimento e autonomia se aplicam às evidências colhidas na primeira parte da dissertação. Com isso, o movimento das plataformas da internet passa a ser funcionalmente explicado como uma transformação procedimental voltada, por um lado, a produzir legitimidade, isto é, garantir aceitação às decisões produzidas pela moderação de conteúdo; e, por outro, em assegurar autonomia a essa produção jurídica, não apenas em relação à plataforma – algo que é, na realidade, um efeito colateral – mas fundamentalmente quanto a outros regimes jurídicos, como os dos Estados, escalando a relação Estado / plataforma, que antes era de uma mera subordinação, para uma colisão entre regimes.

Com essas conclusões, foi possível, ao final, projetar discussões que podem ser desenvolvidas em continuidade a este trabalho: primeiro, fica em aberto de que modo se desenvolverão os mecanismos de constitucionalização material das plataformas, isto é, como elas fixarão valores para decidir em um nível capaz de lhes garantir autonomia em relação aos demais regimes jurídicos. Segundo, é preciso ainda compreender como esse movimento de autonomização do direito avançará, seja promovendo a constitucionalização de novos âmbitos da atuação das plataformas – como os relacionados à inteligência artificial, aos algoritmos que modulam aquilo que as pessoas veem ou à política de anúncios; seja mobilizando a memória institucional do direito em um sentido diferente, resgatando outros mecanismos, como o da legitimidade democrática. Terceiro, também

fica a ser solucionada a questão de como se dará a relação entre Estados e plataformas nesse contexto de crescente autonomia da moderação de conteúdo – ou seja, o que acontecerá, agora, à medida que começarem a surgir conflitos entre aquilo que se decide nas plataformas e aquilo que se decide no Estado, particularmente se, como exposto no item anterior, as plataformas decidirem adquirir legitimidade democrática. Quarto, permanece aberta a questão de saber em medida as divisões políticas e econômicas, supostamente superadas pela característica mundial das plataformas, não serão nelas reinscridas, na forma de comunidades divididas tal como é o aspecto político global, reproduzindo, no ambiente digital, as divisões que ele prometeu apagar.

Em 2019, Bruce Schneier argumentou que, para fazer frente aos desafios regulatórios da atualidade, os atores responsáveis por formatar a estrutura tecnológica por trás da internet tinham que tomar parte de um debate mais amplo, capaz de apreender as implicações de sua criação para os direitos das pessoas<sup>792</sup>. Segundo ele – na frase que serviu de epígrafe para este trabalho –, dois passos seriam necessários para isso: primeiro, quem trabalha na estruturação de uma plataforma precisa ter a consciência de que, ao fazê-lo, constrói um mundo à parte do nosso. Segundo, aqueles que são responsáveis por regular as plataformas precisam entender esse mundo à parte.

Este foi um trabalho sobre esse mundo à parte: em particular, sobre como mecanismos localizados no interior de empresas de tecnologia definem aquilo que as pessoas podem dizer, regulando a liberdade de expressão nas plataformas da internet. Ele se propôs, nesse sentido, tanto a esclarecer como essas estruturas funcionam, conceituando e sistematizando a moderação de conteúdo; quanto a entender como ela se transforma, construindo uma estrutura funcionalista capaz de analisá-la. Essa, espero, tenha sido a sua contribuição mais significativa: não a de explicar uma única transição, mas a de elaborar um quadro sociológico capaz de, daqui em diante, fazer frente às transformações que certamente estão por vir. Isso permite que as conclusões expostas acima sejam avaliadas não por aquilo que elas explicam – mas sim por aquilo que elas ainda podem explicar, deixando abertos os espaços de sua própria continuidade.

---

<sup>792</sup> SCHNEIER, B. Why technologists need to get involved in public policy? 2018 (14m52s). Disponível em: <<https://www.youtube.com/watch?v=U2jn4pXDZn0>>, acesso em 12.1.2020.



## REFERÊNCIAS

ABREU, J. Disrupting the disruptive: making sense of app blocking in Brazil. **Internet Policy Review**, v. 7, n. 3, pp. 1/16, 2018, disponível em <<https://policyreview.info/articles/analysis/disrupting-disruptive-making-sense-app-blocking-brazil>>.

ADL. Quantifying Hate: A Year of Anti-Semitism on Twitter. Disponível em <<https://www.adl.org/resources/reports/quantifying-hate-a-year-of-anti-semitism-on-twitter#holocaust-denial>>, acesso em 20.6.2019.

AGRE, P. The Internet and Public Discourse. **First Monday**, v. 3, n. 3, 1998 (n. p.).

ALAMO, K. Facebook handled Alex Jones just right. **CNN**, 5.5.2019, disponível em <<https://edition.cnn.com/2019/05/05/opinions/facebook-banish-alex-jones-alaimo/index.html>>, acesso em 19.6.2019.

ALANG, N. Inside Facebook's global constitutional convention. **The Week**, 12.5.2019, disponível em <<https://theweek.com/articles/840591/inside-facebooks-global-constitutional-convention>>, acesso em 9.9.2019.

ALCOVER, P. G. **El Derecho en La Teoría de la Sociedad de Niklas Luhmann**. Barcelona: J. M. Bosch, 1993, 400 pp..

ALEMANHA. Netzwerkdurchsetzungsgesetz vom 1. September 2017 (BGBl. I S. 3352), disponível em <<https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>>, acesso em 20.6.2019.

\_\_\_\_\_. Strafgesetzbuch in der Fassung der Bekanntmachung vom 13. November 1998 (BGBl. I S. 3322), § 130 Volksverhetzung, disponível em <[https://www.gesetze-im-internet.de/stgb/\\_130.html](https://www.gesetze-im-internet.de/stgb/_130.html)>.

ALEXANDER, J. YouTube looks to demonetization as punishment for major creators, but it doesn't work. **The Verge**, 25.6.2019, disponível em <<https://www.theverge.com/2019>>

[/6/25/18744246/youtube-demonetization-steven-crowder-patreonadvertising-merch](https://www.youtube.com/watch?v=6/25/18744246)>, acesso em 22.7.2019.

AMATO, L. F., BARROS, M. A. (org.) **Teoria Crítica dos Sistemas? Crítica, teoria social e direito**. Porto Alegre: Fi, 2018, 430 pp.

AMMORI, M. The 'New' New York Times: Free Speech Lawyering in the Age of Google and Twitter. **Harvard Law Review**, v. 127, n. 8, pp. 2.259/2.295, 2014.

ANDERSON, J., STENDER, M., WEST, S., YORK, J. Unfriending Censorship: Insights from four months of crowdsourced data on social media censorship. **Online Censorship**, 31.3.2016, disponível em <<https://www.onlinecensorship.org/en/news-and-analysis/44>>, acesso em 30.7.2019.

ANDREWS, N., SEETHARAMAN, D. Facebook Steps Up Efforts Against Terrorism. **The Wall Street Journal**, 11.2.2016, disponível em <<https://www.wsj.com/articles/facebook-steps-up-efforts-against-terrorism-1455237595>>, acesso em 23.7.2019.

ANDREWS, T. Turkish man, 22, fatally shoots himself on Facebook Live. **The Washington Post**, 13.10.2016, disponível em <[https://www.washingtonpost.com/news/morning-mix/wp/2016/10/13/turkish-man-22-fatally-shoots-himself-on-facebook-live/?utm\\_term=.4a95c3b1bd37](https://www.washingtonpost.com/news/morning-mix/wp/2016/10/13/turkish-man-22-fatally-shoots-himself-on-facebook-live/?utm_term=.4a95c3b1bd37)>, acesso em 22.7.2019.

ARMIJO, E. Meet the New Governors, Same as the Old Governors. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, disponível em <<https://knightcolumbia.org/content/meet-new-governors-same-old-governors>>, acesso em 30.7.2019.

BACHUR, J. P. **Às portas do labirinto: para uma recepção crítica da teoria social de Niklas Luhmann**. Rio de Janeiro: Azougue, 2010, 300 pp..

BALKIN, J. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **UC Davis Law Review**, v. 51, pp. 1.149/1.210, 2018.

BALKIN, J. Free Speech is a Triangle. **Columbia Law Review**, v. 118, n. 7, pp. 2.011/2.056, 2018.

\_\_\_\_\_. Old-School / New-School Speech Regulation. **Harvard Law Review**, v. 127, n. 8, pp. 2.296/2.342, 2014.

\_\_\_\_\_. The Future of Free Expression in the Digital Age. **Pepperdine Law Review**, v. 36, pp. 427/444, 2009.

BANKSTON, K. SINGH, S. Facebook And Google Finally Take First Steps On Road To Transparency About Content Moderation. **Techdirt**, 26.4.2018, disponível em <<https://www.techdirt.com/articles/20180426/10164939724/facebook-google-finally-take-first-steps-road-to-transparency-about-content-moderation.shtml>>, acesso em 30.7.2019.

BARALDI, C., CORSI, G., ESPOSITO, E. **Glossar zu Niklas Luhmann**. Frankfurt: Suhrkamp, 1998, 217 pp.

BBC. Web's inventor gets a knighthood. **BBC**, 31.12.2003, disponível em <<http://news.bbc.co.uk/2/hi/technology/3357073.stm>>, acesso em 22.7.2019, *apud* ORLY, L. The Law of the Platform. **Minnesota Law Review**, San Diego Legal Studies Paper n.º 16-212, pp. 87/166, 2016.

BEARAK, M. Facebook apologizes for taking down a pro-Israel post. **The Washington Post**, 15.6.2016, disponível em <[https://www.washingtonpost.com/news/worldviews/wp/2016/06/15/facebook-apologizes-for-taking-down-a-pro-israelpost/?noredirect=on&utm\\_term=.867baac48fec](https://www.washingtonpost.com/news/worldviews/wp/2016/06/15/facebook-apologizes-for-taking-down-a-pro-israelpost/?noredirect=on&utm_term=.867baac48fec)>, acesso em 30.7.2019.

BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12 n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>> (n. p.).

BELLI, L., FRANCISCO, P. A., ZINGALES, N. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police *in* BELLI, L., ZINGALES, N. (ed.)

**Platform Regulators: How Platforms are Regulated and How They Regulate Us.** Rio de Janeiro: FGV, 2017, pp. 41/64.

BENKLER, Y. A Free Irresponsible Press: Wikileaks and the Battle over the Soul of the Networked Fourth Estate. **Harvard Civil Rights-Civil Liberties Law Review**, v. 46, n. 2, pp. 311/397, 2011.

BERGEN, M. YouTube's New Moderators Mistakenly Pull Right-Wing Channels. **Bloomberg**, 28.2.2018, disponível em <<https://www.bloomberg.com/news/articles/2018-02-28/youtube-s-new-moderators-mistakenly-pull-right-wing-channels>>, acesso em 30.7.2019.

BERGER, K. **The Creeping Codification of the New Lex Mercatoria.** Alphen aan den Rijn: Kluwer Law International, 2010, 464 pp..

BEVER, L. A pediatrician exposes suicide tips for children hidden in videos on YouTube and YouTube Kids. **The Washington Post**, 24.2.2019, disponível em <[https://www.washingtonpost.com/technology/2019/02/24/pediatrician-exposes-suicide-tips-children-hidden-videos-youtube-youtubekids/?noredirect=on&utm\\_term=.b9389fe41524](https://www.washingtonpost.com/technology/2019/02/24/pediatrician-exposes-suicide-tips-children-hidden-videos-youtube-youtubekids/?noredirect=on&utm_term=.b9389fe41524)>, acesso em 23.7.2019.

BICKERT, M. Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process. **Facebook Newsroom**, 24.4.2018, disponível em <<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>>, acesso em 30.7.2019.

BLAIR, O. Mark Zuckerberg accused of ‘abusing power’ after Facebook deletes iconic image of ‘napalm girl’. **The Independent**, 9.6.2016, disponível em <<https://www.independent.co.uk/news/people/mark-zuckerberg-facebook-accused-abusing-power-napalm-girl-vietnam-war-image-a7233431.html>>, acesso em 19.6.2019.

BLOGGER. **Política de Conteúdo.** Disponível em <<https://www.blogger.com/content.g?hl=pt&bpli=1>>, acesso em 23.7.2019.

BLUM, R. Israeli 'Lawfare' Group Says Experiment Reveals Anti-Jewish Bias on Facebook. **The Algemeiner**, 05.01.2016, disponível em <<https://www.algemeiner.com/2016/01/05/israeli-lawfare-group-experiment-reveals-anti-jewish-bias-on-facebook/>>, acesso em 30.7.2019.

BOHN, D. How Facebook makes money: 85 percent from ads, 12 percent from Zynga. **The Verge**, 1.2.2012, disponível em <<https://www.theverge.com/2012/2/1/2764825/facebook-revenue-ads-12-percent-zynga>>, acesso em 23.7.2019.

BRANDOM, R. New rules challenge Google and Facebook to change the way they moderate users. **The Verge**, 7.5.2018, disponível em <<https://www.theverge.com/2018/5/7/17328764/santa-clara-principles-platform-moderation-ban-google-facebook-twitter>>, acesso em 30.7.2019.

BRANNON, V. Free speech and the regulation of social media content. **Congressional Research Service**, 27.3.2019, disponível em <<https://fas.org/sgp/crs/misc/R45650.pdf>>, acesso em 19.6.2019.

BRASIL. Supremo Tribunal Federal. Acórdão no Habeas Corpus n.º 82.424/RS. Relator: ALVES, M. Publicado no DJ de 19.3.2004, p. 24. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=2052452>>, acesso em 13.6.2019.

\_\_\_\_\_. Supremo Tribunal Federal. Acórdão no Recurso Extraordinário n.º 330.817/RJ. Relator: TOFFOLI, D. Publicado no DJe de 27.11.2017. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=1984213>>, acesso em 16.6.2019.

\_\_\_\_\_. Supremo Tribunal Federal. Acórdão no Recurso Extraordinário n.º 511.961/SP. Relator: MENDES, G. Publicado no DJ de 13.11.2009, p. 692. Disponível em <<http://portal.stf.jus.br/processos/detalhe.asp?incidente=2469175>>, acesso em 17.6.2019.

BROOMFIELD, M. Twitter shuts down 125,000 Isis-linked accounts. **The Independent**, 6.2.2016, disponível em <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/125000-isis-linked-accounts-suspended-by-twitter-a6857371.html>>, acesso em 23.7.2019.

BUNI, C., CHEMALY, S. The Secret Rules of the Internet. **The Verge**, 13.4.2016, disponível em <<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>>, acesso em 22.7.2019.

CAMPILONGO, C. F. **O Direito na Sociedade Complexa**. São Paulo: Saraiva, 2011, 188 pp.

CHANDER, A. Facebookistan. **North Carolina Law Review**, v. 90 n. 5, pp. 1.807/ 1.844, 2012.

CHEN, A. The laborers who keep dick pics and beheadings out of your Facebook feed. **Wired**, 23.10.2014, disponível em <<https://www.wired.com/2014/10/content-moderation/>>, acesso em 24.7.2019.

CITRON, D. Extremist Speech, Compelled Conformity and Censorship Creep. **Notre Dame Law Review**, v. 93, n. 3, pp. 1.035/1.072, 2018.

\_\_\_\_\_. Technological Due Process. **Washington University Law Review**, v. 85 n. 6, pp. 1.249/1.313, 2008.

CLAM, J. **Droit et Societé dans la sociologie de Niklas Luhmann: Fondés en contingence**. Paris: Presses Universitaires de France, 1997, 337 pp..

CLEGG, N. Charting a Course for an Oversight Board for Content Decisions. **Facebook Newsroom**, 28.1.2019, disponível em <<https://newsroom.fb.com/news/2019/01/oversight-board/>>, acesso em 30.7.2019.

COOPER, R. Copyright laws are breaking YouTube. Here's how to fix the problem. **The Week**, 7.3.2016, disponível em <<https://theweek.com/articles/608700/copyright-laws-are-breaking-youtube-heres-how-fix-problem>>, acesso em 24.7.2019.

COX, C. (Chris Cox). Sem título. Menlo Park, 1 de outubro de 2014. **Facebook**: chris.cox. Disponível em <<https://www.facebook.com/chris.cox/posts/10101301777354543.>>, acesso em 23.7.2019.

COX, J. Leaked Documents Show Facebook's 'Threshold' for Deleting Pages and Groups. **Vice**, 18.7.2018, disponível em <[https://www.vice.com/en\\_us/article/ne5nxz/leaked-documents-facebook-threshold-de-lete-pages-groups](https://www.vice.com/en_us/article/ne5nxz/leaked-documents-facebook-threshold-de-lete-pages-groups)>, acesso em 23.7.2019.

DALE, C. Taking a harder look at harassment. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/taking-harder-look-at-harassment.html>>, acesso em 30.7.2019.

DAS, S., KRAMER, A. Self-Censorship on Facebook. **Proceedings of the Seventh International Conference on Weblogs and Social Media**, 2013, disponível em <<https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6093/6350>>, acesso em 23.7.2019.

DEARDEN, L. New Zealand attack: How nonsensical white genocide conspiracy theory cited by alleged gunman is spreading poison around the world. **The Independent**, 16.3.2019, disponível em <<https://www.independent.co.uk/news/world/australasia/new-zealand-christchurch-mosque-attack-white-genocide-conspiracy-theory-a8824671.html>>, acesso em 19.6.2019.

DELCKER, J. Germany to Zuckerberg: There won't be Holocaust denial on German Facebook. **Politico**, 19.7.2018, disponível em <<https://www.politico.eu/article/germany-to-zuckerberg-there-wont-be-holocaust-denial-on-german-facebook/>>, acesso em 19.6.2019.

DENARDIS, L., HACKL, A. M. Internet governance by social media platforms. **Telecommunications Policy**, v. 39, n. 9, pp. 761/770, 2015.

DEWEY, M. Wenn die Entscheidungsakzeptanz scheitert in BIRLE, P., DEWEY, M., MASCAREÑO, A. (org.). **Durch Luhmanns Brille: Herausforderungen an Politik und Recht in Lateinamerika und in der Weltgesellschaft**. Wiesbaden: Springer, 2012, pp. 59/74.

DUNN, C. Applying the Constitution to Private Actors. **New York Civil Liberties Union**, 28.4.2009, disponível em <<https://www.nyclu.org/en/publications/column-applying-constitution-private-actors-new-york-law-journal>>, acesso em 19.6.2019.

DWOSKIN, E. ROMM, T. Twitter adds labels for tweets that break its rules — a move with potentially stark implications for Trump’s account. **The Washington Post**, 27.6.2019, disponível em <[https://www.washingtonpost.com/technology/2019/06/27/twitter-adds-labels-tweets-that-break-its-rules-putting-president-trump-companys-crosshairs/?utm\\_term=.284dbf6b9605](https://www.washingtonpost.com/technology/2019/06/27/twitter-adds-labels-tweets-that-break-its-rules-putting-president-trump-companys-crosshairs/?utm_term=.284dbf6b9605)>, acesso em 22.7.2019.

EGELAND, T. (Tom Egeland). Sem título. Oslo, 26 de agosto de 2016. **Facebook**: tom.egeland. Disponível em <<https://www.facebook.com/tom.egeland/posts/10153691636277750>>, acesso em 13.6.2019.

ELECTRONIC FRONTIER FOUNDATION. 88 rights groups call for Facebook to implement appeals process for removed content. **IFEX**, 13.11.2018, disponível em <[https://ifex.org/digital\\_rights/2018/11/13/facebook\\_eff\\_letter/](https://ifex.org/digital_rights/2018/11/13/facebook_eff_letter/)>, acesso em 30.7.2019.

\_\_\_\_\_. **Tossed Out**, disponível em <<https://www.eff.org/tossedout>>, acesso em 8.1.2020.

ELLICKSON, R. **Order without Law: How Neighbors Settle Disputes**. Cambridge: Harvard University Press, 1991, 316 pp..

ESTADOS UNIDOS DA AMÉRICA. Suprema Corte. *Grosjean v. American Press Co.*, 297 U.S. 233 (1936).

\_\_\_\_\_. Suprema Corte. *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46 (1988).

\_\_\_\_\_. Suprema Corte. *Lochner v. New York*, 198 U.S. 45 (1905)

\_\_\_\_\_. Suprema Corte. *Marsh v. Alabama*, 326 U.S. 501 (1946).

\_\_\_\_\_. Suprema Corte. *Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241 (1974).

\_\_\_\_\_. Suprema Corte. *Nebraska Press Association v. Stuart*, 427 U.S. 539 (1976).

\_\_\_\_\_. Suprema Corte. *New York Times Co. v. Sullivan*, 376 US 254 (1964).



ESTADOS UNIDOS DA AMÉRICA. Suprema Corte. *New York Times Co. v. United States*, 403 U.S. 713 (1971).

\_\_\_\_\_. Suprema Corte. *Packingham v. North Carolina*, 582 U. S. \_\_\_\_ (2017).

\_\_\_\_\_. Suprema Corte. *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997).

\_\_\_\_\_. Suprema Corte. *Richmond Newspapers Inc. v. Virginia*, 448 U.S. 555 (1980).

\_\_\_\_\_. Suprema Corte. *Time, Inc. v. Firestone*, 424 U.S. 448 (1976).

FACEBOOK NEWSROOM (Facebook). Sem título. Sem localização, 17 de março de 2019.

**Twitter:** fbnewsroom. Disponível em <[https://twitter.com/fbnewsroom/status/1107117981358682112?ref\\_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref\\_url=https%3A%2F%2F](https://twitter.com/fbnewsroom/status/1107117981358682112?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1107117981358682112%7Ctwgr%5E393039363b74776565745f6d65646961&ref_url=https%3A%2F%2F)>, acesso em 23.7.2019.

FACEBOOK. Community Standards Enforcement Report. **Facebook Transparency**, 2018, disponível em <<https://transparency.facebook.com/community-standards-enforcement#hate-speech>>, acesso em 30.7.2019.

\_\_\_\_\_. Como faço para banir ou cancelar o banimento de alguém da minha Página? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/289207354498410?helpref=about\\_content](https://www.facebook.com/help/289207354498410?helpref=about_content)>, acesso em 23.7.2019.

\_\_\_\_\_. **Facebook Live**. Disponível em <<https://live.fb.com>>, acesso em 23.7.2019.

\_\_\_\_\_. Global Feedback & Input on the Facebook Oversight Board for Content Decisions. **Facebook Newsroom**, 27.6.2019, disponível em <<https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf/>>, acesso em 29.7.2019.

\_\_\_\_\_. **Padrões da Comunidade**. Disponível em <<https://www.facebook.com/communitystandards/>>, acesso em 23.7.2019.

FACEBOOK. Por que vejo um aviso antes de poder visualizar uma foto ou um vídeo? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/814083248683500?helpref=uf\\_permalink](https://www.facebook.com/help/814083248683500?helpref=uf_permalink)>, acesso em 23.7.2019.

\_\_\_\_\_. Q & A on Transparency for Electoral and Issue Ads. **Facebook Newsroom**, 24.5.2018, disponível em <<https://newsroom.fb.com/news/2018/05/q-and-a-on-ads-transparency/>>, acesso em 30.7.2019.

\_\_\_\_\_. Quais são as diversas funções das Páginas e o que elas podem fazer? **Central de Ajuda do Facebook**, disponível em <[https://www.facebook.com/help/289207354498410?helpref=about\\_content](https://www.facebook.com/help/289207354498410?helpref=about_content)>, acesso em 23.7.2019.

\_\_\_\_\_. Standing Against Hate. **Facebook Newsroom**, 27.3.2019, disponível em <<https://newsroom.fb.com/news/2019/03/standing-against-hate/>>, acesso em 19.6.2019.

\_\_\_\_\_. Stats. **Facebook Newsroom**, 2018, disponível em <<https://newsroom.fb.com/company-info/>>, acesso em 23.7.2019.

\_\_\_\_\_. **Termos de Serviço**. Disponível em <<https://www.facebook.com/legal/terms>>, acesso em 23.7.2019.

FAERAS, A., NORDBY, K. OLSEN, T. GJERDE, R. Erna Solberg svarer sensur av historisk Vietnam-fotografi med å legge ut svartsladdede versjoner av flere kjente og historiske bilder. **Aftenposten**, disponível em <<https://www.aftenposten.no/norge/i/qEjdL/Slik-svarer-Erna-Solbergs-pa-Facebook-sensur>>, acesso em 13.6.2019.

FARIA, J. E. A Globalização Econômica e sua Arquitetura Jurídica (10 tendências do Direito Contemporâneo). **Revista da Academia Judicial**, v. 1, pp. 41/59, 2010.

FARREL, H., LEVI, M., O'REILLY, T. Mark Zuckerberg runs a nation-state and he's the king. **Vox**, 10.4.2018, disponível em <<https://www.vox.com/the-big-idea/2018/4/9/17214752/zuckerberg-facebook-power-regulation-data-privacy-control-political-theory-data-breach-king>>, acesso em 13.6.2019.

FERRAZ JR., T. S. **Introdução ao estudo do direito: técnica, decisão, dominação**. São Paulo: Atlas, 2018, (n. p.).

FINKELSTEIN, D. Don't change laws to mop up after YouTube. **The Times**, 29.5.2018, disponível em <<https://www.thetimes.co.uk/article/don-t-change-laws-to-mop-up-after-youtube-c9lhrkwrg>>, acesso em 19.6.2019.

FISCHER-LESCANO, A. Critical Systems Theory. **Philosophy and Social Criticism**, pp 1/21, 16.11.2011, disponível em <<http://psc.sagepub.com/content/early/2011/11/15/0191453711421600>>, acesso em 4.1.2020;

FISCHER-LESCANO, A., TEUBNER, G. Regime-Collisions: The Vain Search for Legal Unity in the Fragmentation of Global Law. **Michigan Journal of International Law**, v. 25, n. 4, pp. 999/1046, 2004.

FISHER, M. Inside Facebook's Secret Rulebook for Global Political Speech. **The New York Times**, 27.12.2018, disponível em <<https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>>, acesso em 22.7.2019.

FLICKR. **Sobre o Flickr**. Disponível em <<https://www.flickr.com/about>>, acesso em 23.7.2019.

FRIENDLY, H. Some Kind of Hearing. **University of Pennsylvania Law Review**, v. 123, n. 6, pp. 1.267/1.317, 1975.

FUNG, B. PayPal bans Alex Jones, saying Infowars 'promoted hate or discriminatory intolerance'. **The Washington Post**, 29.9.2018, disponível em <[https://www.washingtonpost.com/technology/2018/09/21/paypal-bans-alex-jones-saying-infowars-promoted-hate-or-discriminatory-intolerance/?utm\\_term=.6f6931d5660d](https://www.washingtonpost.com/technology/2018/09/21/paypal-bans-alex-jones-saying-infowars-promoted-hate-or-discriminatory-intolerance/?utm_term=.6f6931d5660d)>, acesso em 19.6.2019.

GADDE, V. Twitter Executive: Here's How We're Trying to Stop Abuse While Preserving Free Speech. **The Washington Post**, 16.4.2015, disponível em <<https://www.washington>

[post.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/?utm\\_term=.e8fd3bcd5193](https://posteverything.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/?utm_term=.e8fd3bcd5193)>, acesso em 30.7.2019.

GARTENBERG, C. YouTube updates channel strike system with one-time warning and more consistent punishments. **The Verge**, 19.2.2019, disponível em <<https://www.theverge.com/2019/2/19/18231382/youtube-strike-system-warning-punishment-community-guidelines>>, acesso em 29.7.2019.

GASCA, D., TOIZER, S. More clarity on reported Tweets and enforcement. **Twitter Blog**, 17.10.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/product/2018/more-clarity-on-reported-tweets-and-enforcement.html](https://blog.twitter.com/en_us/topics/product/2018/more-clarity-on-reported-tweets-and-enforcement.html)>, acesso em 29.7.2019.

GIBBONS, K. Antisemitic hate posts allowed by Facebook. **The Times**, 27.7.2018, disponível em <<https://www.thetimes.co.uk/article/antisemitic-hate-posts-including-holocaust-denial-allowed-by-facebook-7crdvw2x>>, acesso em 19.6.2019.

GILLESPIE, T. **Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018, Kindle (n. p.).

GOLBECK, J. On Second Thought... **Slate**, 13.12.2013, disponível em <<https://slate.com/technology/2013/12/facebook-self-censorship-what-happens-to-the-posts-you-dont-publish.html>>, acesso em 23.7.2019.

GOLDMAN, B. Les frontières du droit et 'lex mercatoria'. **Archives de Philosophie du Droit**, n. 9, pp. 177/192, 1964.

GOOGLE. Aplicação das diretrizes da comunidade do YouTube. **Google Transparency Report**, 2018, disponível em <<https://transparencyreport.google.com/youtube-policy/removals>>, acesso em 30.7.2019.

\_\_\_\_\_. Como funciona o ContentID. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2797370?hl=pt-BR>>, acesso em 24.7.2019.

GOOGLE. Conceitos básicos sobre Avisos das diretrizes da comunidade. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2802032>>, acesso em 22.7.2019.

\_\_\_\_\_. Contestar ações das diretrizes da comunidade. **Ajuda do YouTube**, disponível em <[https://support.google.com/youtube/answer/185111?hl=pt-BR&ref\\_topic=2803138](https://support.google.com/youtube/answer/185111?hl=pt-BR&ref_topic=2803138)>, acesso em 29.7.2019.

\_\_\_\_\_. Políticas de Conteúdo Violento ou Explícito. **Ajuda do YouTube**, disponível em <<https://support.google.com/youtube/answer/2802008?hl=pt-BR/>>, acesso em 23.7.2019.

\_\_\_\_\_. **Termos de Serviço**. Disponível em <<https://policies.google.com/terms?hl=pt-BR>>, acesso em 23.7.2019.

\_\_\_\_\_. Ver seu histórico de denúncias. **Ajuda do YouTube**, disponível em <[https://support.google.com/youtube/answer/7687979?hl=pt-BR&ref\\_topic=2803138](https://support.google.com/youtube/answer/7687979?hl=pt-BR&ref_topic=2803138)>, acesso em 29.7.2019.

GRAHAM, J. Did social media do the right thing by banning Alex Jones? **USA Today**, 6.8.2018, disponível em <<https://www.usatoday.com/story/tech/talkingtech/2018/08/06/did-facebook-youtube-and-apple-do-right-thing-banning-alex-jones/916763002/>>, acesso em 19.6.2019.

GRIMMELMANN, J. The Virtues of Moderation. **Yale Journal of Law and Technology**, v. 17, n. 1, pp. 42/109, 2015.

GUERRA FILHO, W. Luhmann and Derrida: Immunology and Autopoiesis *in* LA COUR, A., PHILIPPOPOULOS-MIHALOPOULOS, A. **Luhmann Observed**. Houndmills: Palgrave Macmillan, 2013, pp. 227/242.

HANSEN, E. Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture. **Aftenposten**, 8.9.2016, disponível em <<https://www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this->

to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture>, acesso em 13.6.2019.

HARRIS, B. An Update on Building a Global Oversight Board. **Facebook Newsroom**, 12.12.2019, disponível em <<https://about.fb.com/news/2019/12/oversight-board-update/>>, acesso em 8.1.2020.

HARVEY, D. Making our rules easier to understand. **Twitter Blog**, 7.6.2019, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2019/rules-refresh.html](https://blog.twitter.com/en_us/topics/company/2019/rules-refresh.html)>, acesso em 29.7.2019.

HATMAKER, T. RealNamePolicy and the real story behind Facebook's name policy tumble. **The Daily Dot**, 3.10.2014, disponível em <<https://www.dailydot.com/debug/realnamepolice-facebook-real-names-policy/>>, acesso em 23.7.2019.

HAYES, C. Trump attacks social media sites after James Woods, Alex Jones other extremist conservatives banned. **USA Today**, 4.5.2019, disponível em <<https://www.usatoday.com/story/news/politics/2019/05/04/donald-trump-attacks-twitter-facebook-james-woods-banned/1102418001/>>, acesso em 19.6.2019.

HEINS, M. The Brave New World of Social Media Censorship. **Harvard Law Review Forum**, v. 127, n. 8, pp. 325/330, 2014.

HENNEBEL, L., HOCHMANN, T. **Genocides Denials and the Law**. New York: Oxford University Press, 2011, 380 pp..

HERN, A. Facebook moderators tell of strict scrutiny and PTSD symptoms. **The Guardian**, 26.2.2019, disponível em <<https://www.theguardian.com/technology/2019/feb/26/facebook-moderators-tell-of-strict-scrutiny-and-ptsd-symptoms>>, acesso em 24.7.2019.

\_\_\_\_\_. Mark Zuckerberg's remarks on Holocaust denial 'irresponsible'. **The Guardian**, 19.7.2018, disponível em <<https://www.theguardian.com/technology/2018/jul/19/mark-zuckerberg-holocaust-denial-facebook-remarks-offensive>>, acesso em 20.6.2019.

HOPKINS, N. Revealed: Facebook's internal rulebook on sex, terrorism and violence. **The Guardian**, 21.5.2017, disponível em <<https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence>>, acesso em 30.7.2019.

IBRAHIM, Y. Facebook and the Napalm Girl: Reframing the Iconic as Pornographic. **Social Media + Society**, pp. 1/10, 2017, disponível em <<https://journals.sagepub.com/doi/pdf/10.1177/2056305117743140>>, acesso em 13.6.2019.

ICC. **ICC Rules of Arbitration**. Disponível em <<https://iccwbo.org/dispute-resolution-services/arbitration/rules-of-arbitration/>>, acesso em 8.9.2019.

INFOWARS. (Infowars). Página no Twitter. Sem localização, sem data. **Twitter**: infowars. Disponível em <<https://twitter.com/account/infowars>>, acesso em 19.6.2019.

INGRAM, D. Facebook to keep wraps on political ads data despite researchers' demands. **Reuters**, 22.6.2017, disponível em <<https://www.reuters.com/article/us-usa-politics-facebook/facebook-to-keep-wraps-on-political-ads-data-despite-researchers-demands-idUSKBN19D1CN>>, acesso em 30.7.2019.

INTERNET LIVE STATS. **Twitter Usage Statistics**. Disponível em <<https://www.internetlivestats.com/twitter-statistics/>>, acesso em 23.7.2019.

INTERNET.ORG. **A focus on efficiency: a white paper from Facebook, Ericsson and Qualcomm**. 16.9.2013, disponível em <<https://pt.scribd.com/document/168671112/A-Focus-On-Efficiency-Internet-org-Whitepaper>>, acesso em 23.7.2019.

JOHNSON, E. Full Q&A: Former Google lawyer and deputy U.S. CTO Nicole Wong on Recode Decode. **Vox**, 12.9.2018, disponível em <<https://www.vox.com/2018/9/12/17848384/nicole-wong-cto-lawyer-google-twitter-kara-swisher-decode-podcast-full-transcript>>, acesso em 24.7.2019.

JONES, A. (Alex Jones). Página no Twitter. Sem localização, sem data. **Twitter**: realalexjones. Disponível em <<https://twitter.com/account/realalexjones>>, acesso em 19.6.2019.

KAYE, D. **Speech Police: The Global Struggle to Govern the Internet**. New York: Columbia Global Reports, 2019, 189 pp..

KANG, C. Nancy Pelosi Criticizes Facebook for Handling of Altered Videos. **The New York Times**, 29.5.2019, disponível em <<https://www.nytimes.com/2019/05/29/technology/facebook-pelosi-video.html>>, acesso em 20.6.2019.

KASTRENAKES, J. Twitter says people are tweeting more, but not longer, with 280-character limit. **The Verge**, 8.2.2018, disponível em <<https://www.theverge.com/2018/2/8/16990308/twitter-280-character-tweet-length>>, acesso em 22.7.2019.

KELLY, H. Drag queens lead "real names" protest at Facebook. **CNN Business**, 1.6.2015, disponível em <<https://money.cnn.com/2015/06/01/technology/facebook-protest-names/index.html>>, acesso em 23.7.2019.

KELLY, M. Twitter will now hide - but not remove - harmful tweets from public figures. **The Verge**, 27.6.2019, disponível em <<https://www.theverge.com/2019/6/27/18761132/twitter-donald-trump-rules-violation-tweet-hide-remove-political-figures>>, acesso em 22.7.2019.

KINGSLEY, P. New Zealand Massacre Highlights Global Reach of White Extremism. **The New York Times**, 15.3.2019, disponível em <<https://www.nytimes.com/2019/06/18/world/australia/new-zealand-terrorism-christchurch.html>>, acesso em 23.7.2019.

KLEIN, E. Mark Zuckerberg on Facebook's hardest year, and what comes next. **Vox**, 2.4.2018, disponível em <<https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge>>, acesso em 30.7.2019.



KLEIN, E. Mark Zuckerberg on Facebook's hardest year, and what comes next. **Vox**, 2.4.2018, disponível em <<https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge>>, acesso em 30.7.2019.

KLEINMAN, A. Why 'Kill George Zimmerman' Facebook Pages Still Exist. **The Huffington Post**, 17.7.2013, disponível em <[https://www.huffpostbrasil.com/2013/07/17/kill-george-zimmerman-facebook\\_n\\_3611633.html](https://www.huffpostbrasil.com/2013/07/17/kill-george-zimmerman-facebook_n_3611633.html)>, acesso em 24.7.2019.

KLEINMAN, Z. Fury over Facebook 'Napalm girl' censorship. **BBC News**, 9.9.2016, disponível em <<https://www.bbc.com/news/technology-37318031>>, acesso em 13.6.2019.

KLONICK, K. Facebook v. Sullivan. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, disponível em <<https://knightcolumbia.org/content/facebook-v-sullivan>>, acesso em 30.7.2019.

\_\_\_\_\_. The New Governors: The People, Rules, And Processes Governing Online Speech. **Harvard Law Review**, v. 131, n. 6, pp. 1.598/1.670, 2017.

\_\_\_\_\_. You'll Never Guess This One Crazy Thing Governs Online Speech. **Slate**, 24.8.2016, disponível em <<https://slate.com/technology/2016/08/free-speech-is-the-wrong-way-to-think-about-twitter-and-facebook.html>>, acesso em 19.6.2019.

KLONICK, K., KADRI, T. How to Make Facebook's 'Supreme Court' Work. **The New York Times**, 17.11.2018, disponível em <<https://www.nytimes.com/2018/11/17/opinion/facebook-supreme-court-speech.html>>, acesso em 30.7.2019.

KNIGHT FIRST AMENDMENT INSTITUTE. **Knight Institute v. Trump - Lawsuit Challenging President Trump's Blocking of Critics on Twitter**. Disponível em <<https://knightcolumbia.org/content/knight-institute-v-trump-lawsuit-challenging-president-trumps-blocking-critics-twitter>>, acesso em 23.7.2019.

KOEBLER, J. e COX, J. The Impossible Job: Inside Facebook's Struggle to Moderate Two Billion People. **Motherboard**, 23.8.2018, disponível em

<[https://motherboard.vice.com/en\\_us/article/xwk9zd/how-facebook-content-moderation-works](https://motherboard.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works)>, acesso em 22.7.2019.

KOSEFF, J. **The Twenty-Six Words that Created the Internet**. Ithaca: Cornell University Press, 2019, 313 pp.

LASSON, K. Holocaust Denial and the First Amendment: The Quest for Truth in a Free Society. **George Mason Law Review**, v. 6, n. 35, pp. 35/86, 1997.

LECHER, C. Twitter says it banned more than 200 Russia-linked propaganda accounts. **The Verge**, 28.8.2017, disponível em <<https://www.theverge.com/2017/9/28/16381134/twitter-russia-election-ban>>, acesso em 23.7.2019.

LEIGH, D., HARDING, L. **WikiLeaks: Inside Julian Assange's War on Secrecy**. London: Guardian Books, 2011, 352 pp.

LESSIG, L. The Law of the Horse: What Cyberlaw Might Teach. **Harvard Law Review**, v. 113, n. 2, pp. 501/546, 1999.

\_\_\_\_\_. What Things Regulate Speech: CDA 2.0 v.s. Filtering. **Jurimetrics**, v. 38, n. 4, pp. 629/670, 1998.

LEVIN, S., WONG, J., HARDING, L. Facebook backs down from 'napalm girl' censorship and reinstates photo. **The Guardian**, 9.9.2016, disponível em <<https://www.theguardian.com/technology/2016/sep/09/facebook-reinstates-napalm-girl-photo>>, acesso em 13.6.2019.

LIAO, S. Twitter permanently suspends Infowars and Alex Jones. **The Verge**, 6.9.2018, disponível em <<https://www.theverge.com/2018/9/6/17829188/twitter-permanently-suspension-infowars-alex-jones>>, acesso em 19.6.2019.

LIPMAN, D. To whom it may concern – an open letter to Facebook. **The Jerusalem Post**, 9.6.2016, disponível em <<https://www.jpost.com/Opinion/To-whom-it-may-concern-an-open-letter-to-Facebook-456388>>, acesso em 30.7.2019.

LIPMAN, J. Facebook: Holocaust denial pages can stay. **The Jewish Chronicle**, 28.7.2011, disponível em <<https://www.thejc.com/news/world/facebook-holocaust-denial-pages-can-stay-on-site-1.25833>>, acesso em 19.6.2019.

LIPTAK, A. Facebook says that it removed 1.5 million videos of the New Zealand mass shooting. **The Verge**, 17.3.2019, disponível em <<https://www.theverge.com/2019/3/17/18269453/facebook-new-zealand-attack-removed-1-5-million-videos-content-moderation>>, acesso em 23.7.2019.

LOUIS, E. We need to talk about Alex Jones. **CNN**, 2.5.2019, disponível em <<https://edition.cnn.com/2018/08/06/opinions/infowars-opinion-roundup/index.html>>, acesso em 19.6.2019.

LUHMANN, N. 'Was ist der Fall' und 'Was steckt dahinter?' Die zwei Soziologie und Gesellschaftstheorie. **Zeitschrift für Soziologie**, v. 22, n. 4, pp. 245//260, 1993.

\_\_\_\_\_. **Aufsätze und Reden**. Stuttgart: Philipp Reclam, 2001, 334 pp..

\_\_\_\_\_. **Ausdifferenzierung des Rechts: Beiträge zur Rechtssoziologie und Rechtstheorie**. Frankfurt: Suhrkamp, 1999, 458 pp..

\_\_\_\_\_. **Das Recht der Gesellschaft**. Frankfurt: Suhrkamp, 1995, 598 pp..

\_\_\_\_\_. Deconstruction as Second-Order Observing. **New Literary History**, v. 24, n. 4, pp. 763/782, 1993.

\_\_\_\_\_. **Der Begriff der Gesellschaft**. Disponível em <[http://www.hyperkommunikation.ch/literatur/luhmann\\_begriff\\_gesellschaft.htm](http://www.hyperkommunikation.ch/literatur/luhmann_begriff_gesellschaft.htm)>, acesso em 4.1.2020, n. p.

\_\_\_\_\_. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 1, 594 pp..

\_\_\_\_\_. **Die Gesellschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1998, v. 2, 1.164 pp..

- LUHMANN, N. **Die Politik der Gesellschaft**. Frankfurt: Suhrkamp, 2002, 444 pp..
- \_\_\_\_\_. **Die Realität der Massenmedien**. Opladen: Westdeutscher Verlag, 1996, 219 pp..
- \_\_\_\_\_. **Die Wissenschaft der Gesellschaft**. Frankfurt: Suhrkamp, 1992, 732 pp..
- \_\_\_\_\_. Differentiation of Society. **The Canadian Journal of Sociology**, v. 2, n. 1, pp. 29/53, 1977.
- \_\_\_\_\_. Einführende Bemerkungen zu einer Theorie symbolisch generalisierter Kommunikationsmedien. **Zeitschrift für Soziologie**, v. 3, n. 3, pp. 236/255, 1974.
- \_\_\_\_\_. **Einführung in die Theorie der Gesellschaft**. Heidelberg: Carl-Auer Verlag, 2005, 336 pp..
- \_\_\_\_\_. **Einführung in die Theorie der Gesellschaft**. Heidelberg: Carl Auer Verlag, 2005, 336 pp..
- \_\_\_\_\_. **Essays on Self-Reference**. New York: Columbia University Press, 1990, 245 pp..
- \_\_\_\_\_. Globalization or World Society: How to Conceive of Modern Society? **International Review of Sociology**, v. 7, n. 1, pp. 67/79, 1997.
- \_\_\_\_\_. **Legitimação pelo Procedimento**. Brasília: Ed. Universidade de Brasília, 1980, 202 pp..
- \_\_\_\_\_. **Legitimation durch Verfahren**. Frankfurt: Suhrkamp, 1983, 261 pp..
- \_\_\_\_\_. **Macht**. Stuttgart: Lucius & Lucius, 2003, 153 pp..
- \_\_\_\_\_. O Conceito de Sociedade, *in* NEVES, C. E. B.; SAMIOS, E. M. B. (Org.). **Niklas Luhmann: a nova teoria dos sistemas**. Porto Alegre: Editora da UFRGS; Goethe-Institut, pp. 75/91, 1997.

LUHMANN, N. Operational Closure and Structural Coupling: The Differentiation of the legal system. **Cardozo Law Review**, v. 13, pp. 1.419/1.441, 1991.

\_\_\_\_\_. Quod Omnes Tangit: Remarks on Jurgen Habermas's Legal Theory. **Cardozo Law Review**, v. 17, pp. 883/899, 1995.

\_\_\_\_\_. **Rechtssoziologie**. Opladen: Westdeutscher Verlag, 1983, 393 pp..

\_\_\_\_\_. **Soziale Systeme: Grundriß einer allgemeinen Theorie**. Frankfurt: Suhrkamp, 1991, 675 pp..

\_\_\_\_\_. **Soziologie des Risikos**. Berlin: Walter de Gruyter, 1991, 252 pp..

\_\_\_\_\_. **Soziologische Aufklärung 1**. Opladen: Westdeutscher Verlag, 1970, 268 pp..

\_\_\_\_\_. **Soziologische Aufklärung 2: Aufsätze zur Theorie der Gesellschaft**. Wiesbaden: Springer, 1975, 221 pp..

\_\_\_\_\_. **Soziologische Aufklärung 4: Beiträge zur funktionalen Differenzierung der Gesellschaft**. Opladen: Westdeutscher Verlag, 1987, 276 pp..

\_\_\_\_\_. **Soziologische Aufklärung 6: Die Soziologie und der Mensch**. Opladen: Westdeutscher Verlag, 1995, 275 pp..

\_\_\_\_\_. Tautology and Paradox in the Self-Descriptions of Modern Society. **Sociological Theory**, v. 6, n. 1, pp. 21/37, 1988.

\_\_\_\_\_. **Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität**. Stuttgart: Lucius & Lucius, 2014, 128 pp.

\_\_\_\_\_. Was ist Kommunikation? **Soziologische Aufklärung 6: Die Soziologie und der Mensch**. Opladen: Westdeutscher Verlag, pp. 113/124, 1995.

LUHMANN, N., BEHNKE, K. The Modernity of Science. **New German Critique**, v. 61, pp. 9/23, 1994.

LUI, K. Mark Zuckerberg Has Been Blasted Over Facebook's Censorship of the 'Napalm Girl' Photo. **Time**, 9.9.2016, disponível em <<https://time.com/4484905/facebook-censorship-napalm-girl-newspaper-editor-norway/>>, acesso em 13.6.2019.

MANESS, J. M. Library 2.0 Theory: Web 2.0 and its implications for Libraries. **Webology**, v. 3, n. 2, 2006 (n. p.), *apud* BEER, D., BURROWS, R. Sociology and, of and in Web 2.0: Some Initial Considerations. **Sociological Research Online**, v. 12, n. 5, 2007, disponível em <<http://www.socresonline.org.uk/12/5/17.html>> (n. p.).

METZ, C. A Fake Zuckerberg Video Challenges Facebook's Rules. **The New York Times**, 11.6.2019, disponível em <<https://www.nytimes.com/2019/06/11/technology/fake-zuckerberg-videofacebook.html>>, acesso em 20.6.2019.

MICROSOFT. **How does PhotoDNA technology work?** Disponível em <<https://www.microsoft.com/en-us/photodna>>, acesso em 23.7.2019.

MORRIS, D. YouTube Pulls Alex Jones Video Saying Student Anti-Gun Activists Were Actors. **Fortune**, 24.2.2018, disponível em <<http://fortune.com/2018/02/24/youtube-pulls-alex-jones-infowars-video/>>, acesso em 19.6.2019.

MURPHY, P. InfoWars' main YouTube channel is two strikes away from being banned. **CNN**, 24.2.2018, disponível em <<https://edition.cnn.com/2018/02/23/us/infowars-youtube-videos-trnd/index.html>>, acesso em 19.6.2019.

NASSEHI, A. Funktionale Analyse *in* JAHRAUS, O, NASSEHI, A., GRIZELK, M., SAAKE, I., KIRCHMEIER, C., MÜLLER, J. **Luhmann-Handbuch: Leben, Werk, Wirkung**. Stuttgart: Springer, 2012, pp. 83/84.

NEVES, M. Do pluralismo Jurídico à Miscelânea Social: o problema da falta de identidade da(s) esfera (s) de juridicidade da modernidade periférica e suas implicações na América Latina. **Revista Direito em Debate**, v. 4, n. 5, 2013.

NEWMAN, L. Facebook apologizes for removing pro-Israel post. **The Jerusalem Post**, 15.6.2016, disponível em <<https://www.jpost.com/Diaspora/Facebook-apologizes-for-removing-pro-Israel-post-456774>>, acesso em 30.7.2019.

NEWTON, C. Bodies in Seats. **The Verge**, 19.6.2019, disponível em <<https://www.theverge.com/2019/6/19/18681845/facebook-moderator-interviews-video-trauma-ptsd-cognizant-tampa>>, acesso em 23.7.2019.

\_\_\_\_\_. Facebook is deleting links to a viral attack on a Charlottesville victim. **The Verge**, 14.8.2017, disponível em <<https://www.theverge.com/2017/8/14/16147126/facebook-delete-viral-post-charlottesville-daily-stormer>>, acesso em 30.7.2019.

\_\_\_\_\_. The Trauma Floor. **The Verge**, 25.2.2019, disponível em <<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>>, acesso em 24.7.2019.

O'REILLY, T. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. **International Journal of Digital Economics**, n. 65, pp. 17/37, 2007.

OHLHEISER, A. Facebook backs down, will no longer censor the iconic 'Napalm Girl' war photo. **Washington Post**, 9.9.2016, disponível em <[https://www.washingtonpost.com/news/the-intersect/wp/2016/09/09/abusing-your-power-mark-zuckerberg-slammed-after-facebook-censors-vietnam-war-photo/?utm\\_term=.2176e89e5381](https://www.washingtonpost.com/news/the-intersect/wp/2016/09/09/abusing-your-power-mark-zuckerberg-slammed-after-facebook-censors-vietnam-war-photo/?utm_term=.2176e89e5381)>, acesso em 13.6.2019.

OSOFSKY, J. (Justin Osofsky). I want to share some thoughts about expression and community standards on Facebook. Sem localização, 9 de setembro de 2016. **Facebook**: josofsky. Disponível em <<https://www.facebook.com/josofsky/posts/10157347245570231>>, acesso em 13.6.2019.

PEARCE, M., PIERSON, D. Tech companies gave massive platforms to conspiracy theorists like Alex Jones. Is the crackdown finally here? **Los Angeles Times**, 7.8.2018, disponível

em <<https://www.latimes.com/nation/la-na-alex-jones-20180807-story.html>>, acesso em 20.6.2019.

PRESS ASSOCIATION. Twitter bans 270,000 accounts for 'promoting terrorism'. **The Guardian**, 5.4.2018, disponível em <<https://www.theguardian.com/technology/2018/apr/05/twitter-bans-270000-accounts-to-counter-terrorism-advocacy>>, acesso em 22.7.2019.

PRICE, R. Facebook just censored the prime minister of Norway for posting one of the most famous photos ever. **Business Insider**, 9.9.2016, disponível em <<https://www.businessinsider.com/facebook-censors-norway-prime-minister-erna-solberg-napalm-girl-vietnam-photo-aftenposten-2016-9>>, acesso em 19.6.2019.

PROJECT VIC. **About us**. Disponível em <<https://www.projectvic.org/about-2/>>, acesso em 23.7.2019.

READ, M. Facebook Is Going to Have a Supreme Court. Will It Work? **New York Magazine**, 30.1.2019, disponível em <<http://nymag.com/intelligencer/2019/01/facebook-new-oversight-board-is-a-supreme-court.html>>, acesso em 9.9.2019.

REPÚBLICA FRANCESA. Tribunal de Grande Instance de Paris. LICRA ET UEJF CONTRE YAHOO! Inc. et YAHOO FRANCE, disponível em <<http://www.lapres.net/ya2011.html>>, acesso em 8.9.2019.

ROETTIGERS, J. Facebook Bans Alex Jones, Milo Yiannopoulos, Other Far-Right Figures. **Variety**, 2.10.2019, disponível em <<https://variety.com/2019/digital/news/facebook-alex-jones-ban-1203203444/>>, acesso em 19.6.2019.

ROGERS, K. Twitter 'sorry' for suspending Guy Adams as NBC withdraws complaint. **The Guardian**, 31.7.2012, disponível em <<https://www.theguardian.com/technology/2012/jul/31/guy-adams-twitter-growing-pains>>, acesso em 30.7.2019.



ROOSE, K. Facebook and YouTube give Alex Jones a Wrist Slap. **The New York Times**, 27.7.2018, disponível em <<https://www.nytimes.com/2018/07/27/technology/alex-jones-facebook-youtube.html>>, acesso em 19.6.2019.

\_\_\_\_\_. Facebook Banned Infowars. Now What? **The New York Times**, 10.8.2018, disponível em <<https://www.nytimes.com/2018/08/10/technology/facebook-banned-infowars-now-what.html>>, acesso em 19.6.2019.

ROSE, F. Zuckerberg was right about how to handle Holocaust deniers. **The Washington Post**, 30.7.2018, disponível em <[https://www.washingtonpost.com/news/theworldpost/wp/2018/07/30/facebook/?noredirect=on&utm\\_term=.39567f50c7f5](https://www.washingtonpost.com/news/theworldpost/wp/2018/07/30/facebook/?noredirect=on&utm_term=.39567f50c7f5)>, acesso em 19.6.2019.

ROSEN, A., IHARA, I. Giving you more characters to express yourself. **Twitter**, 26.9.2017, disponível em <[https://blog.twitter.com/en\\_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html](https://blog.twitter.com/en_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html)>, acesso em 23.7.2019.

ROSEN, J. The Deciders: The Future of Privacy and Free Speech in the Age of Facebook and Google. **Fordham Law Review**, v. 80 n. 4, pp. 1.525/1.538, 2012.

ROSENBERG, E. Alex Jones Apologizes for Promoting ‘Pizzagate’ Hoax. **The New York Times**, 25.3.2017, disponível em <<https://www.nytimes.com/2017/03/25/business/alex-jones-pizzagate-apology-comet-ping-pong.html>>, acesso em 19.6.2019.

ROSENBERG, Y. Mark Zuckerberg Is Doubly Wrong About Holocaust Denial. **The Atlantic**, 19.7.2018, disponível em <<https://www.theatlantic.com/ideas/archive/2018/07/the-solution-to-holocaust-deniers-on-facebook/565562/>>, acesso em 19.6.2019.

ROTH, Y., HARVEY, D. How Twitter is fighting spam and malicious automation. **Twitter Blog**, 26.6.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html](https://blog.twitter.com/en_us/topics/company/2018/how-twitter-is-fighting-spam-and-malicious-automation.html)>, acesso em 30.7.2019.

SALES, B. Facebook says it will allow Holocaust denial. **The Jerusalem Post**, 4.5.2019, disponível em <<https://www.jpost.com/Diaspora/Antisemitism/Facebook-says-it-will-allow-Holocaust-denial-588651>>, acesso em 19.6.2019.

SALINAS, S. YouTube removes Alex Jones' page, following bans from Apple and Facebook. **CNBC**, 6.8.2018, disponível em <<https://www.cnbc.com/2018/08/06/youtube-removes-alex-jones-account-following-earlier-bans.html>>, acesso em 19.6.2019.

SANTOS, B. The Law of the Oppressed: The Construction and Reproduction of Legality in Pasargada. **Law & Society Review**, v. 12, n. 1, pp. 5/126, 1977.

SCHUMACHER, E. Facebook refuses to censor Holocaust denial as social media sites struggle with German laws. **Deutsche Welle**, 27.7.2018, disponível em <<https://www.dw.com/en/facebook-refuses-to-censor-holocaust-denial-as-social-media-sites-struggle-with-german-laws/a-44855519>>, acesso em 19.6.2019.

SCHNEIER, B. Why technologists need to get involved in public policy? 2018 (14m52s). Disponível em: <<https://www.youtube.com/watch?v=U2jn4pXDZn0>>, acesso em 12.1.2020.

SCOTT, M., ISAAC, M. Facebook Restores Iconic Vietnam War Photo It Censored for Nudity. **The New York Times**, 9.9.2016, disponível em <<https://www.nytimes.com/2016/09/10/technology/facebook-vietnam-war-photo-nudity.html>>, acesso em 13.6.2019.

SHABAN, H. Twitter reveals its daily active user numbers for the first time. **The Washington Post**, 7.2.2019, disponível em <[https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-firsttime/?utm\\_term=.84f6d67871a1](https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-firsttime/?utm_term=.84f6d67871a1)>, acesso em 23.7.2019.

SHAMSIAN, J. Facebook's head of policy it would allow 'denying the Holocaust' in the weeks before banning high-profile anti-Semitic conspiracy theorists. **Business Insider**, 3.5.2019, disponível em <<https://jewishinsider.com/2019/05/daily-kickoff-facebook-removes-inflammatory-figures-but-will-allow-holocaust-denial-sfrc-confirms-david-schenker-nyu-tel-aviv-singled-out/>>, acesso em 19.6.2019.

SIMS, A. Facebook to Reinstate Censored Image of ‘Napalm Girl’ after Mark Zuckerberg accused of ‘abusing power’. **The Independent**, 9.9.2016, disponível em <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-to-reinstate-censored-image-of-napalm-girl-after-mark-zuckerberg-accused-of-abusing-power-a7235021.html>>, acesso em 13.6.2019.

SOLBERG, E. Facebook had no right to edit history. **The Guardian**, 9.9.2016, disponível em <<https://www.theguardian.com/commentisfree/2016/sep/09/facebook-napalm-vietnamese-deleted-norway>>, acesso em 21.6.2019.

SPEYER, L. Sting Operation Reveals Facebook’s Double ‘Community Standards’ in Relation to Pro-Israel Posts. **The Algemeier**, 15.6.2016, disponível em <<https://www.algemeiner.com/2016/06/15/sting-operation-reveals-facebooks-double-community-standards-in-relation-to-pro-israel-posts/>>, acesso em 30.7.2019.

STACK, L. Trump Wants Your Tales of Social Media Censorship. And Your Contact Info. **The New York Times**, 15.5.2019, disponível em <<https://www.nytimes.com/2019/05/15/us/donald-trump-twitter-facebook-youtube.html>>, acesso em 19.6.2019.

STATISTA. **Hours of video uploaded to YouTube every minute 2007-2019**. 15.7.2019, disponível em <<https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>>, acesso em 23.7.2019.

STATT, N. Tim Cook says Apple didn’t coordinate Alex Jones ban with other tech companies. **The Verge**, 2.10.2018, disponível em <<https://www.theverge.com/2018/10/2/17930534/apple-ceo-tim-cook-alex-jones-infowars-ban-independent-decision>>, acesso em 19.6.2019.

\_\_\_\_\_. Twitter removes Trump’s video featuring music from The Dark Knight for copyright infringement. **The Verge**, 10.4.2019, disponível em <<https://www.theverge.com/2019/4/10/18303841/trump-twitter-the-dark-knight-rises-copyright-infringement-video-removed-hans-zimmer>>, acesso em 23.7.2019.

STECKLOW, S. Why Facebook is losing the war on hate speech in Myanmar. **Reuters**, 15.8.2018, disponível em <<https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>>, acesso em 30.7.2019.

STEWART, E. YouTube finally banned content from neo-Nazis, Holocaust deniers, and Sandy Hook skeptics. **Vox**, 5.6.2019, disponível em <<https://www.vox.com/recode/2019/6/5/18653666/youtube-bans-white-supremacists-holocaust-carlos-maza>>, acesso em 19.6.2019.

STF. Jobim se despede da 2ª Turma para assumir presidência do STF. Notícias STF, 4.5.2004, disponível em <<http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=62661>>, acesso em 12.8.2019.

SWENEY, M. Twitter shares fall amid concern over falling user numbers. **The Guardian**, 7.2.2019, disponível em <<https://www.theguardian.com/technology/2019/feb/07/twitter-shares-fall-amid-concern-over-falling-user-numbers>>, acesso em 23.7.2019.

SWISHER, K. Mark Zuckerberg clarifies: ‘I personally find Holocaust denial deeply offensive, and I absolutely didn’t intend to defend the intent of people who deny that’. **Vox**, 18.7.2018, disponível em <<https://www.vox.com/2018/7/18/17588116/mark-zuckerberg-clarifies-holocaust-denial-offensive>>, acesso em 19.6.2019.

\_\_\_\_\_. Zuckerberg: The Recode interview. **Vox**, 8.10.2018, disponível em <<https://www.vox.com/2018/7/18/17575156/mark-zuckerberg-interview-facebook-recode-kara-swisher>>, acesso em 19.6.2019.

TAN, J. Twitter cracks down on right-wing media personality Alex Jones. **CNBC**, 15.8.2018, disponível em <<https://www.cnbc.com/2018/08/15/twitter-suspends-alex-jones.html>>, acesso em 19.6.2019.

TARABAY, J., GRAHAM-MCLAY, C. Could the Christchurch Attacks Have Been Prevented? **The New York Times**, 18.6.2019, disponível em <<https://www.nytimes.com/>>

[2019/06/18/world/australia/new-zealand-terrorism-christchurch.html](https://www.2019/06/18/world/australia/new-zealand-terrorism-christchurch.html)>, acesso em 23.7.2019.

TASKUS. Outsourcing Glossary. Disponível em <<https://www.taskus.com/glossary/content-moderation/>>, acesso em 22.7.2019.

TEUBNER, G. Codes of Conduct multinationalaler Unternehmen: Unternehmensverfassung jenseits von Corporate Governance und gesetzlicher Mitbestimmung, in HOLLAND, A. *et al* (ed.) **Arbeitnehmermitwirkung in einer sich globalisierenden Arbeitswelt**. Berlin: Bwv Berliner-Wissenschaft, pp. 109/117, 2005 (n. p.).

\_\_\_\_\_. Constitutional Drift: On the Spontaneous Co-evolution of Social ‘Ideas’ and Legal ‘Structures’, in DOWDLE, M., WILKINSON, M. (Ed.) **Constitutionalism beyond Liberalism**. Cambridge: Cambridge University Press, pp. 79/95, 2017 (n. p.).

\_\_\_\_\_. Corporate Codes in the Varieties of Capitalism: How their Enforcement Depends Upon the Difference Among Production Regimes, in BECKERS, A. (ed.). **Enforcing Corporate Social Responsibility. Indiana Journal of International Law**, v. 24, pp. 81/97, 2017 (n. p.).

\_\_\_\_\_. Die zwei Gesichter des Janus: Rechtspluralismus in der Spätmoderne, in SCHIMDT, E., WEYERS, H. (ed.) **Liber Amicorum Josef Esser**. Heidelberg: C. F. Müller, pp. 191/214, 1995.

\_\_\_\_\_. Fragmented Foundations: Societal Constitutionalism Beyond the Nation State, in DOBNER, P., LOUGHLIN, M. (ed.). **The Twilight of Constitutionalism?** Oxford: Oxford University Press, pp. 327/341, 2010.

\_\_\_\_\_. Globale Bukowina: Zur Emergenz eines transnationalen Rechtspluralismus. **Rechtshistorisches Journal**, v. 15, pp. 255/290, 1996.

\_\_\_\_\_. Globale Zivilverfassungen: Alternativen zur staatszentrierten Verfassungstheorie. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 63, pp. 1/28, 2003.

TEUBNER, G. Horizontal Effects of Constitutional Rights in the Internet: A Legal Case on the Digital Constitution. **Italian Law Journal**, v. 3, n. 1, pp. 193/205, 2017.

\_\_\_\_\_. **Le droit: Un Système Autopoïétique**. Paris: Presses Universitaires de France, 1993, 296 pp..

\_\_\_\_\_. Privatregimes: Neo-Spontanes Recht und duale Sozialverfassungen in der Weltgesellschaft, in SIMON, D., WEISS, M. (Ed.) **Zur Autonomie des Individuums: Liber Amicorum Spiros Simitis**. Baden-Baden: Nomos, 2000, pp. 437/453 (n. p.).

\_\_\_\_\_. Quod omnes tangit: Transnationale Verfassungen ohne Demokratie? **Der Staat**, v. 57, n. 2, pp. 171/194, 2018.

\_\_\_\_\_. Societal Constitutionalism: Nine Variations on a Theme by David Sciulli, in BLOKKER, P., THORNHILL, C. **Sociological Constitutionalism**. Cambridge: Cambridge University Press, pp. 313/340, 2016.

\_\_\_\_\_. The Corporate Codes of Multinationals: Company Constitutions Beyond Corporate Governance and Co-Determination, in NICKEL, R. (ed.). **Conflict of Laws and Laws of Conflict in Europe and Beyond: Patterns of Supranational and Transnational Juridification**. Oxford: Intersentia, 2009 (n. p.).

\_\_\_\_\_. Transnationaler Verfassungspluralismus: Neun Variationen über ein Thema von David Sciulli. **Zeitschrift für ausländisches öffentliches Recht und Völkerrecht**, v. 76, pp. 1/25, 2016.

\_\_\_\_\_. Transnationales Recht: Legitimation durch horizontale Grundrechtswirkung in VIELLECHNER, L. **Transnationalisierung des Rechts**. Velbrück: Weilerswist, prefácio, 2013 (n. p.).

\_\_\_\_\_. Unitas Multiplex: Das Konzernrecht in der neuen Dezentralität der Unternehmensgruppen. **Zeitschrift für Unternehmens- und Gesellschaftsrecht**, v. 20, n. 2, pp. 189/217, 1991.

TEUBNER, G., KORTH, P. Zwei Arten des Rechtspluralismus: Normkollisionen in der doppelten Fragmentierung der Weltgesellschaft, *in* KÖTTER, M., SCHUPPERT, G. (Ed.), **Normative Pluralität ordnen**. Baden-Baden: Nomos, 2009.

THE GUARDIAN. How Facebook handles Holocaust denial. **The Guardian**, 24.5.2017, disponível em <<https://www.theguardian.com/news/gallery/2017/may/24/how-facebook-handles-holocaust-denial>>, acesso em 19.6.2019.

THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION. Disponível em <<https://santaclaraprinciples.org/>>, acesso em 30.7.2019.

\_\_\_\_\_. An Open Letter to Mark Zuckerberg: The World's Freedom of Expression is in Your Hands. Disponível em <<https://santaclaraprinciples.org/open-letter/>>, acesso em 29.7.2019.

TRUMP, D. (Donald Trump). Sem título. Sem localização, 3 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124447302544965634>>, acesso em 19.6.2019.

\_\_\_\_\_. (Donald Trump). Sem título. Sem localização, 3 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124454393749934080>>, acesso em 19.6.2019.

\_\_\_\_\_. (Donald Trump). Sem título. Sem localização, 4 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124645001533771776>>, acesso em 19.6.2019.

\_\_\_\_\_. (Donald Trump). Sem título. Sem localização, 4 de maio de 2019. **Twitter:** realdonaldtrump. Disponível em <<https://twitter.com/realDonaldTrump/status/1124743267873116160>>, acesso em 19.6.2019.

TUMBLR. **O que é o Tumblr?** Disponível em <[https://www.tumblr.com/login?language=pt\\_BR](https://www.tumblr.com/login?language=pt_BR)>, acesso em 23.7.2019.

TWITTER PUBLIC POLICY. Evolving our Twitter Transparency Report: expanded data and insights. **Twitter Blog**, 12.12.2018, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2018/evolving-our-twitter-transparency-report.html](https://blog.twitter.com/en_us/topics/company/2018/evolving-our-twitter-transparency-report.html)>, acesso em 29.7.2019.

\_\_\_\_\_. **Update: Russian interference in the 2016 US presidential election**. 28.9.2017, disponível em <[https://blog.twitter.com/official/en\\_us/topics/company/2017/Update-Russian-Interference-in-2016--Election-Bots-and-Misinformati-on.html](https://blog.twitter.com/official/en_us/topics/company/2017/Update-Russian-Interference-in-2016--Election-Bots-and-Misinformati-on.html)>, acesso em 23.7.2019.

TWITTER SAFETY (Twitter). Sem título. Sem localização, 1 de novembro de 2017. **Twitter: TwitterSafety**. Disponível em <<https://twitter.com/TwitterSafety/status/925827487677870080>>, acesso em 30.7.2019.

\_\_\_\_\_. (Twitter). Sem título. Sem localização, 2 de abril de 2019. **Twitter: TwitterSafety**. Disponível em <<https://twitter.com/TwitterSafety/status/1113139073303089152>>, acesso em 29.7.2019.

\_\_\_\_\_. (Twitter). Sem título. Sem localização, 27 de março de 2018. **Twitter: TwitterSafety**. Disponível em <<https://twitter.com/TwitterSafety/status/978768701481656320>>, acesso em 30.7.2019.

\_\_\_\_\_. Defining public interest on Twitter. **Twitter Blog**, 27.6.2019, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2019/publicinterest.html](https://blog.twitter.com/en_us/topics/company/2019/publicinterest.html)>, acesso em 23.7.2019.

\_\_\_\_\_. Enforcing new rules to reduce hateful conduct and abusive behavior. **Twitter Blog**, 18.12.2017, disponível em <[https://blog.twitter.com/en\\_us/topics/company/2017/safetypoliciesdec2017.html](https://blog.twitter.com/en_us/topics/company/2017/safetypoliciesdec2017.html)>, acesso em 29.7.2019.

TWITTER. **Diretrizes e políticas gerais – Nossas opções de medidas corretivas**. Disponível em <<https://help.twitter.com/pt/rules-and-policies/enforcement-options>>, acesso em 29.7.2019.



TWITTER. **Regras do Twitter – Política de Mídia Sensível**. Disponível em <<https://help.twitter.com/pt/rules-and-policies/media-policy>>, acesso em 23.7.2019.

\_\_\_\_\_. **Sobre**. Disponível em <<https://about.twitter.com/pt.html>>, acesso em 23.7.2019.

\_\_\_\_\_. **Termos de Serviço**. Disponível em <<https://twitter.com/pt/tos>>, acesso em 23.7.2019.

\_\_\_\_\_. Twitter Rules enforcement. **Twitter Transparency Report**, 2018, disponível em <<https://transparency.twitter.com/en/twitter-rules-enforcement.html>>, acesso em 30.7.2019.

UT, N. The Terror of War. **Time**, The Most Influential Images of All Time. Disponível em <<http://100photos.time.com/photos/nick-ut-terror-war>>, acesso em 13.6.2019.

VILLAS BÔAS FILHO, O. V. **Teoria dos Sistemas e o Direito Brasileiro**. São Paulo: Saraiva, 2009, 436 pp.

VINCENT, J. Alex Jones hit with 30-day Facebook suspension for bullying and hate speech. **The Verge**, 27.7.2018, disponível em <<https://www.theverge.com/2018/7/27/17621334/alex-jones-facebook-suspension-30-day>>, acesso em 20.6.2019.

WARREN, R. Hotels at centre of rape allegations promoted on TripAdvisor. **The Guardian**, 5.3.2019, disponível em <<https://www.theguardian.com/world/2019/mar/05/hotels-at-centre-of-allegations-promoted-on-tripadvisor>>, acesso em 23.7.2019.

WATERSON, J. YouTube blocks history teachers uploading archive videos of Hitler. **The Guardian**, 6.6.2019, disponível em <<https://www.theguardian.com/technology/2019/jun/06/youtube-blocks-history-teachers-uploading-archive-videos-of-hitler>>, acesso em 23.7.2019.

WEMPLE, E. Facebook admits ‘mistake’ in blocking Fox News Radio commentator. **The Washington Post**, 29.6.2013, disponível em <<https://www.washingtonpost.com/blogs/>

[erik-wemple/wp/2013/06/29/facebook-admits-mistake-in-blocking-fox-news-radio-commentator/?noredirect=on&utm\\_term=.61936934d549](http://erik-wemple/wp/2013/06/29/facebook-admits-mistake-in-blocking-fox-news-radio-commentator/?noredirect=on&utm_term=.61936934d549)>, acesso em 30.7.2019.

WHITE HOUSE. Sem título. Sem localização, 15 de maio de 2019. **Twitter:** WhiteHouse. Disponível em <<https://twitter.com/WhiteHouse/status/1128765001223663617>>, acesso em 19.6.2019.

WHITNEY, H. Search Engines, Social Media, and the Editorial Analogy. **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2018, disponível em <<https://knightcolumbia.org/content/search-engines-social-media-and-editorial-analogy>>.

WIKILEAKS. **What is Wikileaks**. Disponível em <<https://wikileaks.org/What-is-WikiLeaks.html>>, acesso em 16.6.2019.

WIKIPEDIA. **About**. Disponível em <<https://en.wikipedia.org/wiki/Wikipedia:About>>, acesso em 23.7.2019.

WILLIAMSON, E. How Alex Jones and Infowars Helped a Florida Man Torment Sandy Hook Families. **The New York Times**, 29.3.2019, disponível em <<https://www.nytimes.com/2019/03/29/us/politics/alex-jones-infowars-sandy-hook.html>>, acesso em 19.6.2019.

WU, A. Strike you're out! Or maybe not? **YouTube Official Blog**, 2.7.2010, disponível em <<https://youtube.googleblog.com/2010/07/strike-youre-out-or-maybe-not.html>>, acesso em 30.7.2019.

\_\_\_\_\_. Is the First Amendment Obsolete? **Knight First Amendment Institute at Columbia University**, Emerging Threats, 2017, disponível em <<https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>>, acesso em 13.6.2019.

YOUTUBE. **Diretrizes da Comunidade**. Disponível em <<https://www.youtube.com/intl/pt-BR/yt/about/policies/#community-guidelines/>>, acesso em 23.7.2019.

YOUTUBE. Making our strikes system clear and consistent. **YouTube Creator Blog**, 19.2.2019, disponível em <<https://youtube-creators.googleblog.com/2019/02/making-our-strikes-system-clear-and.html>>, acesso em 29.7.2019.

\_\_\_\_\_. Our ongoing work to tackle hate. **YouTube Official Blog**, 5.6.2019, disponível em <<https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>>, acesso em 22.7.2019.

\_\_\_\_\_. **Sobre**. Disponível em <<https://www.youtube.com/intl/pt-BR/yt/about/>>, acesso em 23.7.2019.

\_\_\_\_\_. **YouTube in Numbers**. Disponível em <<https://www.youtube.com/intl/en-GB/yt/about/press/>>, acesso em 23.7.2019.

ZADROZNY, B. Vaccine misinformation and Infowars: Researchers wary of Facebook's embrace of 'Groups'. **NBC News**, 16.2.2019, disponível em <<https://www.nbcnews.com/tech/tech-news/vaccine-misinformation-infowars-researchers-wary-facebook-s-embrace-groups-n972286>>, acesso em 19.6.2019.

ZARSKY, T. Social Justice, Social Norms and the Governance of Social Media. **Pace Law Review**, v. 35, n. 1, pp. 154/191, 2014.

ZUCKERBERG, M. (Mark Zuckerberg). Sem título. Sem local, 24 de abril de 2018. **Facebook**: zuck. Disponível em <<https://www.facebook.com/zuck/posts/10104874769784071>>, acesso em 30.7.2019.

\_\_\_\_\_. **A Blueprint for Content Governance and Enforcement**. 15.11.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>>, acesso em 30.7.2019.

\_\_\_\_\_. **Preparing for Elections**. 13.9.2018, disponível em <<https://www.facebook.com/notes/mark-zuckerberg/preparing-for-elections/10156300047606634/>>, acesso em 23.7.2019.