PATRICIA KUZMENKO FURLAN

**FATORES DETERMINANTES PARA A ADOÇÃO DAS GOVERNANÇAS DE
DADOS E DE INFORMAÇÃO NO AMBIENTE *BIG DATA***

São Paulo

2018

PATRICIA KUZMENKO FURLAN

**FATORES DETERMINANTES PARA A ADOÇÃO DAS GOVERNANÇAS DE
DADOS E DE INFORMAÇÃO NO AMBIENTE *BIG DATA***

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Ciências.

São Paulo

2018

PATRICIA KUZMENKO FURLAN

**FATORES DETERMINANTES PARA A ADOÇÃO DAS GOVERNANÇAS DE
DADOS E DE INFORMAÇÃO NO AMBIENTE *BIG DATA***

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Ciências.

Área de concentração: Engenharia de
Produção

Orientador: Prof. Dr. Fernando José Barbin
Laurindo

São Paulo
2018

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catlogação-na-publicação

Furlan, Patricia

FATORES DETERMINANTES PARA A ADOÇÃO DAS GOVERNANÇAS DE DADOS E DE INFORMAÇÃO NO AMBIENTE BIG DATA / P. Furlan -- versão corr. -- São Paulo, 2018.

163 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Governança de dados 2.Governança de informação 3.Inteligência competitiva 4.Big data I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

Dedico esta tese à minha mãe e ao meu noivo.
Foram anos de muita evolução que ficarão registrados aqui para sempre, em cada
palavra desta pesquisa.

AGRADECIMENTOS

Ao Prof. Dr. Fernando José Barbin Laurindo, pela confiança, apoio e dedicação em minha jornada acadêmica para o desenvolvimento desta pesquisa. Que o término desta tese seja apenas o começo de nossos trabalhos.

Aos colegas e amigos do LETICIC. Em particular, agradeço e dedico este trabalho àqueles que ativamente estiveram presentes no desenvolvimento desta pesquisa e que contribuíram com análises, comentários, pontos de vistas distintos: Prof. Dra. Cláudia Mattos, Prof. Dr. Renato Moraes, Prof. Dr. Mauro Spínola, Prof. Dra. Kumiko Kissimoto, Prof. Dr. Cláudio de Oliveira, George, Olavo, Lídia, Célia, Macir, Michel, Felipe, Alexandre, Rafael, Leandro, André.

Aos colegas e amigos do Departamento de Engenharia de Produção, que são pessoas inspiradoras e que tornaram o ambiente de trabalho mais diverso e interessante. Agradeço também todos os envolvidos no PPGE/USP.

Agradeço a Prof. Dra. Michela Arnaboldi, do *Politecnico di Milano*, pela oportunidade concedida em fazer parte do quadro de pesquisadores do Departamento de Engenharia Gestional do POLIMI. O período dedicado ao doutorado sanduíche foi precioso para o desenvolvimento desta pesquisa e de muitas outras que virão.

Aos profissionais consultados e que dedicaram tempo, atenção. São muitos àqueles que estiveram envolvidos nesta pesquisa. Gostaria de agradecer profundamente cada um de vocês e também às empresas alvo de estudo. Foi um desafio enorme estudar a governança de dados ou de informação no ambiente *big data*.

À minha família, que sempre me motivou e apoiou nos estudos. Sem o apoio familiar, nada é possível. Agradeço, em especial, minha mãe e meu noivo pela companhia nesses anos de tese, pela compreensão dos momentos vividos para esta pesquisa tomar forma, pelo incentivo e por acreditar, sempre.

À CAPES pelo apoio financeiro para realização desta pesquisa, tanto no Brasil quanto no exterior para realizar o doutorado sanduíche.

“Tutte le verità sono facili da capire una volta che sono state rivelate.

Il difficile è scoprirle.”

Galileo Galilei

RESUMO

No ambiente *big data*, as organizações se preocupam em extrair valor dos dados e das informações com o intuito de obter vantagens competitivas. No entanto, são necessários esforços organizacionais com relação aos ativos de dados, incluindo a definição de responsabilidades com relação ao uso dos dados, a garantia da qualidade dos dados, dentre outros aspectos contemplados pelos modelos de governança de dados ou de informação. Deste modo, esta pesquisa investigou como as organizações podem adotar as governanças de dados ou de informação no ambiente *big data* e, para tanto, foram contemplados estudos de casos multi-setoriais para identificar os fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*. Foram investigados os elementos e os conteúdos dos modelos de governança de dados ou de informação e analisados os aspectos dos modelos com relação à inteligência de negócios e ao *big data analytics*. Notou-se que as ações organizacionais com relação à governança de dados ou de informação são pouco consolidadas, mas conhecidas pelas organizações. Além disto, os modelos de governança de dados ou de informação são adotados por organizações com diferentes níveis de capacidades analíticas. Tais modelos contemplam a definição dos objetivos estratégicos da governança e domínios como o gerenciamento da qualidade dos dados ou das informações, o gerenciamento dos dados (em especial meta-dados), a transformação da mentalidade organizacional com relação aos dados e as informações e necessitam de competências de colaboração e comunicação dos *stakeholders*. Foram identificados oito fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*, os quais contemplam práticas estruturais, relacionais e operacionais do modelo de governança: 1 - Organizações grandes, globais, difusas, com estruturas descentralizadas de negócios e portfolio complexo de produtos ou serviços; 2 – Apontar um *C-level*, definir gerentes na estrutura e determinar *data owners* e *data stewards*; 3 - Estabelecer comitê de dados ou outros meios para reunir a alta cúpula e os principais líderes da organização; 4 - Atuação do departamento de TI nas atividades de gerenciamento de dados ou de informação, viabilizando e executando atividades operacionais com relação aos dados e as informações dentre as bases

de dados e sistemas de informação; 5 - Atuar ativamente na transformação cultural da organização para *data-driven*; 6 - Promover a comunicação e a colaboração interna; desenvolver a comunicação com relação à eficácia das políticas e a necessidade de adequação dos *stakeholders*; 7 - Definir, gerenciar e controlar metadados; 8 - Definir os padrões, as exigências e o controle sobre a qualidade dos dados. A pesquisa oferece uma consolidação teórica relevante para o campo da governança de dados ou da informação, contemplando vasta lista de variáveis da literatura de inteligência competitiva, da literatura de governança de TI, governança de dados e governança de informação. Foi também possível expandir o modelo de governança de dados ou de informação englobando os domínios relativos à colaboração e comunicação, mudança cultural. Propõem-se uma expansão na conceituação geral dos termos governança de dados e governança de informação.

Palavras-chave: Governança de dados. Governança de informação. Inteligência competitiva. *Big data*.

ABSTRACT

In the big data environment, organizations are concerned with extracting value from data and information in order to acquire competitive advantage. However, organizational efforts are required to organize data assets, determine responsibilities with regard to the data assets, ensure data quality, and other aspects. Such activities are covered by data or information governance models. This research investigated how organizations can adopt data or information governance in the big data environment. Thus, it was conducted multi-sectoral case studies to identify determinants factors for the adopting of data or information governance in the big data environment. The research protocol encompassed elements and contents of the data or information governance models and those related to big data value extraction. It was noted that the organizational approaches regarding data or information governance are poorly consolidated, but are well known to organizations. In addition, data or information governance models are adopted by organizations with different levels of analytical capabilities. Those models include the definition of the strategic objectives, and domains like data or information quality management, data management (especially metadata), transformation of the organizational cultural in relation to the data and the information, and collaboration and communication among stakeholders. Eight determinants factor were identified for the adoption of data or information governance in the big data environment, including structural, relational and operational practices of the governance model: 1 - Large, global and diffuse organizations with decentralized business and complex portfolio of products or services; 2 - Define C-level, managers, data owners and data stewards; 3 - Establish a data committee or other means to bring together the top leaders of the organization; 4 – Engagement of the IT department on the data management activities, enabling and executing operational activities in relation to data and information among databases and information systems; 5 - Actively engage in the cultural transformation of the organization into data-driven; 6 - Promote communication and internal collaboration; develop communication on the effectiveness of policies and the need for stakeholder adequacy; 7 - Define, manage and control metadata; 8 - Define standards, requirements and control over data quality. This research provides a relevant theoretical consolidation to the field of data

or information governance, contemplating a vast list of research variables on the fields of competitive intelligence, IT governance, data and information governance literatures. It was also possible to expand the data or information governance model through the addition of domains such as collaboration, communication, and cultural transformation. The research also proposes an expansion in the general conceptualization of the terms data governance and information governance.

Keywords: Data governance. Information governance. Competitive intelligence. Big data.

LISTA DE FIGURAS

Figura 1 – Panorama geral da pesquisa. Fonte: elaborado pela autora.	22
Figura 2 – Capítulos da tese. Fonte: elaborado pela autora.	23
Figura 3 – Publicações sobre governança de dados e governança de informação (dados coletados em Agosto 2017). Fonte: ISI <i>Web of Science</i> e Scopus.....	29
Figura 4 – Estrutura do capítulo de revisão de literatura.....	32
Figura 5 – Visão holística da IC. Fonte: Adaptado de Cabral Netto (2011)	33
Figura 6 – Evolução das publicações em inteligência de negócios e <i>big data analytics</i> (dados retirados em Agosto de 2017). Fonte: ISI <i>Web of Science</i> ; <i>Scopus</i>	35
Figura 7 – Publicações sobre governança de dados ou de informação e a combinação deste campo com a inteligência de negócios e o <i>big data analytics</i> (dados de Agosto de 2017). Fonte: ISI <i>Web of Science</i> ; <i>Scopus</i>	35
Figura 8 – Relação de equivalência de termos entre autores que exploram os campos da BI, <i>analytics</i> , e <i>big data</i> . Fonte: elaborado pela autora.	40
Figura 9 – Fases da inteligência de negócios. Fonte: Bose (2009)	41
Figura 10 – Modelo para classificar as capacidades analíticas das organizações. Fonte: Adaptado de LaValle et al. (2011)	46
Figura 11 – Domínios da governança de TI. Fonte: Weill e Woodham (2002); Weill e Ross (2005); Khatri e Brown (2010).....	52
Figura 12 – Estrutura organizacional da governança de TI. Fonte: Rau (2004)	54
Figura 13 – Níveis dos elementos do modelo de governança. Fonte: elaborado pela autora.	58
Figura 14 – Modelos com domínios da governança de dados. Fonte: Fu et al. (2011); Khatri e Brown (2010).....	60
Figura 15 – Adaptação do modelo de Khatri e Brown sobre os domínios das governanças de dados e de informação. Fonte: Khatri e Brown (2010).....	63
Figura 16 – Domínios do modelo de governança de TI e do modelo de governança de dados ou de informação. Fonte: Khatri e Brown (2010)	64
Figura 17 – <i>Stakeholders</i> da estrutura de governança de dados e de informação. Fonte: elaborado pela autora.	68

Figura 18 – Esquema metodológico com as proposições de pesquisa. Fonte: elaborado pela autora.	75
Figura 19 – Citação dos artigos de Khatri e Brown (2010), Tallon (2013) e Tallon et al. (2014) para as bases Scopus e ISI <i>Web of Science</i> (Janeiro de 2018). Fonte: Scopus; ISI <i>Web of Science</i>	77
Figura 20 – Estrutura organizacional da empresa 1. Fonte: elaborado pela autora	86
Figura 21 – Estrutura organizacional para a empresa 2. Fonte: elaborado pela autora	89
Figura 22 – Frentes do roteiro de pesquisa. Fonte: elaborado pela autora.	91
Figura 23 – Contribuições ao modelo de Khatri e Brown (2010) para a governança de dados ou de informação. Fonte: elaborado pela autora	123
Figura 24 – Estrutura organizacional dos casos com relação à governança de dados ou de informação. Fonte: elaborado pela autora	126
Figura 25 – Definição adotada para as governanças de dados ou de informação. Fonte: elaborado pela autora	138

LISTA DE QUADROS

Quadro 1 – Principais definições da governança de dados e de informação.	27
Quadro 2 – Fatores viabilizadores e inibidores para adoção da governança de informação.	30
Quadro 3 – Caracterização de Chen et al. (2012) sobre a evolução da <i>BI&A</i>	42
Quadro 4 – Caracterização de Davenport (2013) sobre a <i>analytics</i>	43
Quadro 5 – Quadro teórico da revisão da literatura de inteligência competitiva, inteligência de negócios e <i>big data analytics</i>	47
Quadro 6 – Definição dos principais termos utilizados na sessão.	48
Quadro 7 – Resumo das definições sobre os níveis no modelo de governança.	59
Quadro 8 – Práticas da governanças de dados ou de informação.	66
Quadro 9 – Elementos e conteúdos dos modelos de governança de dados ou de informação.	67
Quadro 10 – Quadro teórico da revisão da literatura de governança de TI, governança de dados ou de informação.	72
Quadro 11 – Citações dos artigos que citam Khatri e Brown (2010) pela Scopus (Janeiro 2018).	78
Quadro 12 – Premissas para escolha dos casos organizacionais.	79
Quadro 13 – Definição de construtos e variáveis de pesquisa sobre capacidade analítica e extração de valor no ambiente <i>big data</i>	83
Quadro 14 – Definição de construtos e variáveis de pesquisa.	84
Quadro 15 – Empresas consultadas na pesquisa piloto.	85
Quadro 16 – Sub-práticas operacionais descartadas do roteiro de pesquisa.	90
Quadro 17 – Questões sobre os elementos e conteúdos da governança de dados ou de informação.	92
Quadro 18 – Questões com foco em BI.	92
Quadro 19 – Casos contemplados na pesquisa em campo.	94
Quadro 20 – Perfil das entrevistas realizadas nos estudos de caso.	95
Quadro 21 – Características comuns das organizações da pesquisa em campo.	114
Quadro 22 – Objetivos gerais identificados pelas estruturas de governança de dados ou de informação.	115

Quadro 23 – Estrutura organizacional e abrangência das estruturas de governança de dados ou de informação.....	116
Quadro 24 – Classificação das capacidades analíticas das organizações estudadas.	140
Quadro 25 – Análise das proposições de pesquisa.	142
Quadro 26 – Fatores determinantes para a adoção das governanças de dados ou de informação no ambiente <i>big data</i>	146

SUMÁRIO

1. Introdução	18
1.1 Objetivos	20
1.2 Estrutura da tese	21
2. Revisão de literatura	24
2.1 Inteligência competitiva e capacidades analíticas	33
2.1.1 Evolução das capacidades analíticas	36
2.2 Estrutura da governança de tecnologia da informação	48
2.3 Estrutura das governanças de dados e de informação	57
2.3.1 Domínios	59
2.3.2 Práticas e sub-práticas	65
3. Metodologia de pesquisa.....	73
3.1 Mecanismos para a definição do campo	79
3.2 Variáveis de Pesquisa e Pesquisa Piloto	83
3.2.1 Empresa 1	85
3.2.2 Empresa 2	88
3.3 Roteiro de pesquisa	90
3.4 Estudos de casos	93
3.4.1 Caso A	95
3.4.2 Caso B	98
3.4.3 Caso C	101
3.4.4 Caso D	103
3.4.5 Caso E	105
3.4.6 Caso F	107
3.4.7 Caso G	109
3.4.8 Caso H	112
4. Resultados e discussões	114
4.1 Objetivos gerais das estruturas de governança de dados ou de informação.....	114
4.2 Elementos e conteúdos da governança de dados ou de informação ...	126
4.3 Capacidades analíticas e a governança de dados ou de informação	136

4.4 Proposições de pesquisa	142
4.5 Fatores motivadores, viabilizadores e determinantes para a adoção das governanças de dados ou de informação no ambiente <i>big data</i>	143
5. Conclusões	148
Referências	155

1. Introdução

O aumento na geração e na disponibilidade de dados e de informação (emprega-se o termo “*big data*” para caracterizar o fenômeno) tem contribuído para as mudanças estratégicas das corporações, com a consequente necessidade de adaptação de suas estruturas operacionais. Além disto, o aprimoramento do uso dos dados tem se mostrado útil e necessário não apenas no contexto organizacional, mas também para a sociedade como geral, para os governos, dentre outras esferas. Deste modo, políticas e práticas de governança de dados ou de informações estão sendo foco das entidades público ou privadas para orientar o uso apropriado e mais responsável dos dados no ambiente *big data*.

Dentre os objetivos estratégicos das organizações compreendem-se as atividades de extração de valor dos dados ou das informações para obtenção de vantagens competitivas. Além disto, parte das organizações tem adotado estruturas de governança de dados ou de informação para delimitar as responsabilidades na organização com relação aos ativos de dados, definir políticas, práticas e procedimentos que contribuam para o melhor gerenciamento dos dados, aumento da qualidade dos dados, proteção dos *stakeholders* com relação às exigências legais e regulatórias, dentre outros aspectos com relação aos ativos de dados (ALHASSAN et al., 2016; KOLTAY, 2016).

As pesquisas sobre governança de dados ou de informação são emergentes na literatura de sistemas de informação (ALHASSAN et al., 2016). Apresentam, também, relevância crescente para os negócios e tem atraído o centro da atenção dos executivos (TALLON et al., 2014; KOLTAY, 2016). No entanto, Fu et al. (2011) mencionam que a construção e a delimitação de uma estrutura de governança eficaz são desafiadoras para a organização. Isto porque a simples adoção das tecnologias é insuficiente para governar os ativos de dados e de informação (HAIDER; HAIDER, 2013). Para tanto, devem ser delimitados princípios corretos e práticas adequadas para a organização. Caso contrário, a adoção de práticas excessivas pode culminar com burocracia adicional que prejudicam os retornos positivos da governança de dados ou de informação para a organização (OTTO, 2011).

Em geral, os modelos de governança determinam quais decisões devem ser tomadas por quem na organização para assegurar a gestão eficiente (KHATRI; BROWN, 2010; FU et al., 2011). No ambiente organizacional podem ser definidos modelos para a governança corporativa, governança da tecnologia da informação (TI), governança de dados ou de informação. De forma complementar, por “gestão” compreende-se as tomadas de decisão e aos processos de adoção das decisões estipuladas no âmbito da governança. A gestão se preocupa em "fazer e implementar decisões" (KHATRI; BROWN, 2010).

Koronios et al. (2014) ressaltam que poucos trabalhos exploram os atributos estratégicos sobre *big data* e como obter sucesso neste contexto, embora as publicações sobre *big data* sejam crescentes (KORONIOS et al., 2014). É ainda incerto como tratar os dados e as informações no contexto *big data* (TALLON et al., 2014), como obter valor dos dados ou como certificar se a organização está usufruindo de todo o potencial dos dados (LAVALLE et al., 2011; BARTON; COURT, 2012). Embora novas tecnologias possam gerar e coletar dados em escala cada vez maiores, seu uso ainda é um desafio para as corporações (LAVALLE et al., 2011). Kooper et al. (2011) ressaltam que a maior parcela das organizações, senão todas, carecem de políticas de governança de dados e de informação, cujos esforços são necessários para a melhor compreensão do fenômeno, como preveem Otto (2011). Alhassan et al. (2016) conduziram pesquisas bibliométricas nas bases de dados científicas e identificam escassez de trabalhos que explicitamente estudam as atividades da governança de dados ou de informação. Assim, Tallon et al. (2014) e Alhassan et al. (2016) indicam a necessidade de pesquisas sob a forma de estudo de casos para contribuir com maiores detalhes sobre as governanças de dados ou de informação. Neste sentido, foi definida como questão geral desta pesquisa: **como as organizações podem adotar as governanças de dados ou de informação no ambiente *big data***. Para respondê-la, lançou-se mão de duas questões complementares:

1. Quais são os elementos e conteúdos previstos pela literatura para as governanças de dados ou de informação? Como eles determinam, se isto for possível, na adoção das governanças de dados ou de informação?
2. De que modo a capacidade analítica da organização pode ser um fator determinante para a adoção das governanças de dados ou de informação?

Neste momento introdutório, é importante ressaltar que a governança de dados e a governança de informação foram tratadas indistintamente ao longo desta pesquisa. Por isto será empregado o termo “governança de dados e/ou de informação”. Tal consideração foi defendida pelos Weber et al. (2009), Furlan e Laurindo (2017b), que mencionam que a maior parte da literatura de governança de dados ou de informação trata os termos indistintamente.

Embora a literatura tradicional de IS defina “dados” como os fluxos de fatores brutos e isolados, geralmente desorganizados, sem que haja a transmissão de um significado específico, e “informação” como conjunto de dados devidamente tratados de forma a serem providos de significado (BRAGANZA, 2004; LAURINDO, 2008), há autores na literatura *de big data analytics* que removem a barreira entre “dados” e “informação”, como Batra (2014), consolidando a definição dos dados ou das informação em um único nível. Para efeito desta tese, os dados e informações podem ser considerados:

"Dados ou informação podem contemplar fluxos de fatores brutos; isolados ou não; desorganizados ou não; podem ser estruturados, semi estruturados ou não estruturados; podem ter significado ou não; podem ser tratados ou não; podem contemplar ou não uma finalidade específica na organização" (BRAGANZA, 2004; LAURINDO, 2008).

1.1 Objetivos

Esta tese tem como objetivo investigar e propor fatores que determinam na adoção das governanças de dados ou de informação no ambiente *big data*. Tais fatores são aqueles que desempenham papel fundamental e imprescindível para que o modelo de governança de dados ou de informação possa ser estabelecido no ambiente organizacional e que possa cumprir com seus objetivos estratégicos. Para auxiliar na identificação dos fatores determinantes, foram definidos como objetivos intermediários:

- Estudar o campo teórico emergente (governança de dados ou de informação) e realizar consolidação teórica com variáveis das literaturas de inteligência competitiva e governança de TI, dados e de informação;
- Verificar se e como as práticas e as sub-práticas propostas pela literatura de governança de TI, de dados ou de informação podem constituir fatores determinantes para a adoção das governanças de dados ou de informação;
- Verificar se e como a capacidade analítica da organização pode contribuir (ou determinar) na adoção das governanças de dados e de informação no contexto *big data*.

Para se alcançar os objetivos propostos, a pesquisa contemplou bases teóricas das literaturas de inteligência competitiva, governança de TI, governança de dados e governança de informação. Quanto à classificação das capacidades analíticas da organização, foi empregado o modelo de LaValle et al (2011) que sugere três níveis de capacidades analíticas: aspiracionais, experientes e transformadas. Mais detalhes sobre o modelo serão apresentados na sessão 2.1.1 do capítulo de revisão de literatura.

As variáveis de pesquisa foram aplicadas em múltiplos estudos de casos em organizações globais que atuam no ambiente *big data* e que possuem estruturas de governança de dados ou de informação formalmente constituídas. Do ponto de vista acadêmico, almeja-se contribuir para a formação da literatura da governança de dados ou de informação e, do ponto de vista organizacional, contribuir para a adoção efetiva destas governanças. Além disto, almeja-se servir como referência para trabalhos futuros no campo da extração de valor dos dados ou das informações no ambiente *big data*, sob a vertente dos modelos de governança de dados ou de informação.

1.2 Estrutura da tese

A Figura 1 ilustra um panorama geral desta pesquisa. Através dela é possível observar que a questão de pesquisa orientou na identificação da lacuna teórica que

esta pesquisa almejou explorar. E, dado os objetivos da pesquisa, foi conduzida a revisão de literatura para delimitar um modelo teórico com proposições e variáveis de pesquisa.

Figura 1 – Panorama geral da pesquisa. Fonte: elaborado pela autora.

Com base na revisão de literatura, foi elaborado o roteiro de pesquisa para auxiliar na condução das pesquisas em campo. A metodologia contemplou uma pesquisa piloto para refinamento do roteiro de pesquisa, o qual foi então utilizado para os estudos de casos. As organizações foram selecionadas segundo os critérios: *i.* existência de estrutura formal de governança de dados ou de informação; *ii.* diferentes capacidades analíticas; *iii.* ambiente *big data*; *iv.* TI apresenta desempenho estratégico ou em transição, segundo o *Grid* estratégico de Porter e Millar (1984); *v.* alto conteúdo de informação no produto e alta intensidade de informação na cadeia de valor, segundo a Matriz de intensidade de informação de Porter e Millar (1985). Tais pesquisas em campo foram conduzidas dentre os anos 2017 e 2018. Mais detalhes serão fornecidos adiante no trabalho.

Para a compreensão da estrutura desta tese, a Figura 2 lista os principais temas abordados em cada capítulo desta pesquisa.

Figura 2 – Capítulos da tese. Fonte: elaborado pela autora.

2. Revisão de literatura

O termo "governança" corresponde às formas pelas quais uma organização pode assegurar que suas estratégias sejam estipuladas, monitoradas e alcançadas (RAU, 2004; WEBER et al., 2009), estando presente na literatura organizacional desde 1960 (EELLS, 1960). Neste contexto, a governança focaliza temas específicos, como a governança corporativa, a de tecnologia da informação, de dados ou de informação.

A governança corporativa corresponde ao conjunto de processos, políticas, leis e instituições que afetam no modo como a corporação é direcionada, administrada e controlada (KOOPEL et al., 2011). Esta governança objetiva estabelecer o modelo institucional e as políticas corporativas (WEBER et al., 2009). Inclui-se o relacionamento entre os *stakeholders* principais, *shareholders*, a alta gestão, o *board* de diretores, os funcionários, fornecedores, consumidores e toda a comunidade envolvida nos negócios da corporação (KOOPEL et al., 2011).

Já a governança de TI consiste num braço da governança corporativa que focaliza nos assuntos relativos à tecnologia na corporação (WEBER et al., 2009; KOOPEL et al., 2011; ASSIS, 2011). A governança de TI almeja estabelecer as políticas e as práticas sobre a tecnologia da informação na organização, de modo a contribuir com os negócios e assegurar que a tecnologia não seja impeditiva à implementação das estratégias organizacionais (WEBER et al., 2009). Ela se fundamenta em uma série de políticas e procedimentos racionais e hierárquicos atrelados aos sistemas da informação, sendo estes de administração, autorização, monitoramento e auditoria. Tais atividades são fundamentos racionais que podem limitar as ações empreendedoras, inovadoras, criativas e espontâneas com relação à criação de valor (KOOPEL et al., 2011). As governanças de dados e de informação emergem como um instrumento corporativo para agregar valor às governanças pré-estabelecidas.

O interesse acadêmico pela governança de TI começou nos anos 1990 sobre os temas relacionados ao gerenciamento de projetos de TI (SAMBAMURTY; ZMUD, 1999). Mais adiante, o interesse prático pela governança de TI foi marcado pelo Ato de Sarbanes-Oxley nos EUA, em 2002, e pela Basiléia II na União Européia, em 2004.

A governança de TI focaliza nos artefatos de infraestruturas físicas (*hardwares* e *softwares*), ao passo que as governanças de dados e de informação focalizam artefatos não-físicos ou de informação (HAIDER; HAIDER, 2013; TALLON et al., 2014). A característica não-física dos dados ou da informação possibilita sua rápida replicação em localidades distintas, diferentemente dos artefatos físicos. O valor da informação não necessariamente decresce com o tempo ou com seu uso, como acontece com ativos físicos, podendo até mesmo ter seu valor aumentado com o tempo ou maior uso (TALLON et al., 2014).

Para Kooper et al. (2011), as governanças de dados ou de informação almejam se encaixar nas limitações da governança de TI. Uma delas consiste na falta de preocupação com os dados, em seu processo de criação, de pesquisa, de consumo, de troca, de modo a criar valor para os negócios (KOOOPER et al., 2011). Como descrito por Kemp (2014), "não se pode definir leis nos dados, mas, sim, leis e obrigações em relação aos dados".

De um modo geral, as governanças de dados e de informação tem como objetivos:

- Definir quais dados ou informações a organização precisa (KOOOPER et al., 2001);
- Definir como a organização pode fazer uso dos dados e/ou das informações (KOOOPER et al., 2001);
- Definir quem na organização será responsável por fazer uso dos dados e das informações (KOOOPER et al., 2001);
- Estruturar e gerenciar os dados como ativos (WEBER et al., 2009; OTTO, 2011; TALLON, 2013);
- Maximizar o valor da informação para a organização de modo a assegurar que a informação seja confiável, segura e acessível para a tomada de decisão (TALLON et al., 2014);
- Proteger a informação de modo que seu valor para a organização não seja diminuído pelo uso da tecnologia, pelos erros humanos, pelo uso inapropriado ou pela perda do acesso no tempo (TALLON et al., 2014). Isto inclui a mitigação dos fatores que destroem ou que limitam o valor dos dados (WEBER et al., 2009; OTTO, 2011; TALLON, 2013);

- Promover o equilíbrio entre os riscos e os benefícios da retenção dos dados nos centros corporativos (WEBER et al., 2009; OTTO, 2011; TALLON, 2013);
- Promover a segurança da corporação (WEBER et al., 2009; OTTO, 2011; TALLON, 2013).

As principais definições sobre a governança de dados ou de informação foram listadas no Quadro 1. Por meio dele é possível notar conteúdos ora mais estratégicos e ora mais operacionais. Weber et al. (2009) definem a governança de dados sob a perspectiva da governança de TI, que fora proposta por Weill (2004): “a governança de dados delimita um *framework* para as tomadas de decisões e responsabilidades em relação aos dados”. Khatri e Brown (2010) introduzem o conceito de dados como ativos, cujos detalhes serão expostos adiante no trabalho. Kooper et al. (2011) introduzem uma série de responsabilidades operacionais para a governança de dados, como o estabelecimento de regras e procedimentos para a avaliação, criação, coleta, análise, distribuição, armazenamento, uso e controle da informação. Tallon et al. (2014) complementa a definição ao enfatizar a governança com base no ciclo de vida da informação. Para Tallon et al. (2014), a governança de informação corresponde “a um conjunto de competências ou de práticas para a criação, captura, monetização, armazenamento, uso, controle, acesso, arquivamento e descarte de informações ao longo de seu ciclo de vida” (TALLON et al., 2014).

Otto (2011) propõem uma definição baseada no gerenciamento de dados, que é parte da governança de dados ou de informação. Inclusive Haider e Haider (2013) definem a governança de dados através dos aspectos gerenciais, com ênfase na qualidade dos dados.

Considerando as principais contribuições para definir a governança de dados ou a governança de informação e partindo-se do pressuposto que ambas são tratadas indistintamente nesta tese, foi possível adotar a seguinte definição:

“A governança de dados ou de informação envolve o estabelecimento de regras e de procedimentos sobre quem detém os poderes de decisão e as responsabilidades nas tomadas de decisão na organização com relação aos ativos de dados (KHATRI; BROWN, 2010). Também estabelece um conjunto de competências ou práticas para a avaliação, criação, coleta, análise, distribuição, armazenamento, uso, controle, acesso, arquivamento e exclusão dos dados ou das informações ao longo de seus ciclos de vida” (KOOPER et al., 2011; KHATRI; BROWN, 2010; TALLON et al., 2014).

Quadro 1 - Principais definições da governança de dados e de informação.

Autores	Perspectiva	Definição do construto
Weber et al. (2009)	Governança de dados definida pela governança de TI	"A governança de dados especifica o <i>framework</i> para as tomadas de decisões e as responsabilidades para encorajar comportamentos desejáveis em relação aos dados" (WEBER et al., 2009).
Khatri; Brown (2010)	Dados como ativos com base na governança de TI	"A governança de dados se baseia em quem possui os direitos de decisão e a responsabilidade pela tomada de decisão na organização com relação aos ativos de dados" (KHATRI; BROWN, 2010).
Kooper et al. (2011)	Governança de informação a partir das limitações da governança de TI	"A governança de informação envolve o estabelecimento de regras e de procedimentos para a avaliação, criação, coleta, análise, distribuição, armazenamento, uso e controle da informação" (KOOPER et al., 2011).
Otto (2011)	Governança de dados a partir da governança de TI. Cita Weber et al. (2009) e Khatri e Brown (2010)	"A governança de dados contempla as funções de elaboração de políticas, execução e controle dos dados, onde o gerenciamento dos dados consiste em um sub-grupo. Este, por sua vez, volta-se à execução das políticas traçadas pela governança de dados" (OTTO, 2011).
Haider; Haider (2013)	Governança de dados sob a perspectiva da governança de TI. Cita Weill e Ross (2005)	"A governança de dados corresponde a uma disciplina de controle de qualidade para a avaliação, gerenciamento, uso, aprimoramento, monitoramento, manutenção e proteção da informação organizacional" (HAIDER; HAIDER, 2013).
Tallon (2013)	Governança de dados sob a ótica do custo, risco e retorno. Cita Weber et al (2009) e Khatri e Brown (2010)	"As práticas da governança de dados correspondem a políticas ou procedimentos organizacionais que descrevem como os dados podem ser gerenciados ao longo de seu ciclo de vida econômico. Tais práticas consistem num instrumento para a compreensão e mitigação dos fatores de riscos que podem prejudicar o uso dos dados" (TALLON, 2013).
Tallon et al. (2014)	Governança da informação com base na governança de TI. Cita Khatri e Brown (2010), Kooper et al. (2011), Otto (2011)	"A governança de informação pode ser definida como um conjunto de competências ou de práticas para a criação, captura, monetização, armazenamento, uso, controle, acesso, arquivamento e descarte de informações ao longo de seu ciclo de vida (TALLON et al., 2014).

Fonte: elaborado pela autora

Como pode ser visto pela definição adotada para a governança de dados e de informação, ela contempla duas vertentes: uma estratégica com foco no estabelecimento de uma estrutura organizacional para tomada de decisão, e a outra operacional com conteúdos como a avaliação, criação, coleta, análise, distribuição, armazenamento, uso, controle, acesso, arquivamento e exclusão dos dados ou das informações ao longo de seus ciclos de vida.

A maior parcela do conhecimento sobre a governança de dados ou de informação advém de estudos de caso e de modelos que oferecem percepções sobre os objetivos e os conteúdos da governança (TALLON et al., 2014). Tallon et al. (2014) indicam que a literatura de governança de TI pode contribuir para a formação da literatura de governança de dados e de informação, mas que devem ser contempladas suas especificidades. Não se tem claro como elas se complementam, sobrepõem-se ou divergem umas das outras (TALLON et al., 2014). Há autores que defendem que a governança de dados ou de informação corresponde a um subgrupo da governança de TI, como Haider e Haider (2013) e Tallon et al. (2014), e há aqueles que atrelam o sucesso da governança de TI com a existência das governanças de dados e de informação, como Haupt et al. (2015). Independentemente disto, é necessário garantir o alinhamento e a integração estratégica entre a governança de TI com a governança de dados ou informação (BARATA, 2015). Inclusive podem ser utilizados os atributos da governança de TI para as governanças de dados e de informação (TALLON et al., 2014).

A Figura 3 ilustra as curvas de publicação pela ISI *Web of Science* e *Scopus* sobre a governança de dados e governança de informação. Tais dados foram coletados em Agosto de 2017, onde foram empregadas as chaves de busca: <data governance> OR <information governance> em buscas isoladas nestas plataformas, contemplando todas as formas de publicação e idiomas. Furlan e Laurindo (2017b) notam que as principais publicações científicas no campo abordam áreas como a ciência da computação, medicina, saúde (através dos serviços de ciência da saúde e informática médica), negócios, gestão e ciência da informação.

As primeiras publicações no campo da governança de dados ou de informação originaram-se em 2003, cujo aumento aconteceu principalmente após 2008 (FURLAN; LAURINDO, 2017b). Antes disso, foi identificado o artigo de Brodhag (2000), o qual foi considerado um *outlier* pois focaliza na governança geral das entidades públicas (mais especificamente a agenda Rio 21). Em particular, o

trabalho de Goodhue et al. (1992), não contemplado dentre os artigos da amostra, aborda o papel estratégico dos dados para as organizações ainda nos anos 1990. Este serve de parâmetro sobre a relevância dos dados para as organizações antes do fenômeno *big data*.

Figura 3 – Publicações sobre governança de dados e governança de informação (dados coletados em Agosto 2017). Fonte: ISI Web of Science e Scopus.

Com base nos trabalhos de Sambamurthy e Zmud (1999), Kirsch (2004), Raghupathi (2007), Xue et al. (2008), Willson e Pollard (2009) e Xue et al. (2011), Tallon et al. (2014) listam os fatores teóricos que influenciam positivo ou negativamente no processo de adoção da governança de TI, os quais foram listados pelo Quadro 2.

Os fatores viabilizadores são aqueles que facilitam a adoção das políticas, das práticas e das sub-práticas da governança de dados ou de informação. Aqui se fala na etapa de implementação do modelo, cujos elementos facilitadores podem compreender aspectos estruturais, tecnológicos ou setoriais. Por outro lado, os

fatores inibidores são aqueles que prejudicam, atrasam ou impossibilitam a adoção das políticas, práticas e sub-práticas da governança de dados ou de informação.

Por fatores determinantes compreendem-se aqueles imprescindíveis para o estabelecimento do modelo de governança e que determinam o alcance dos objetivos pré-determinados das estruturas de governança de dados ou de informação. É objetivo deste trabalho o estudo dos fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*.

Quadro 2 – Fatores viabilizadores e inibidores para adoção da governança de informação.

	Conteúdos	Viabilizadores	Inibidores
1	Estratégia organizacional	Estratégia focada nos clientes e na inovação.	
2	Estratégia de TI	Estratégia de TI alinhada com cliente (<i>data analytics</i>) e com as estratégias de inovação.	
3	Estrutura organizacional e estrutura de TI	Organizações centralizadas.	
4	Complexidade produto-serviço		Complexidade de produtos e amplo <i>mix</i> dificulta a adoção da governança de informação.
5	Arquitetura de TI	Padronização da TI e alta integração de processos.	Sistemas de legado e baixa integração de processos.
6	Cultura de TI	Cultura de TI que promove o uso estratégico da informação.	Cultura de acumulador dificulta a adoção da governança.
7	Regulação setorial	Regulações estipulam requerimentos específicos para proteção ou retenção dos dados.	
8	Taxa de aumento dos dados e das informações	A aceleração na quantidade de dados leva as organizações a adotarem a governança.	

Fonte: Adaptado de Tallon et al. (2014)

A estratégia organizacional e, por consequente, a estratégia de TI voltadas ao cliente, à inovação e à cultura *data-driven* podem viabilizar a adoção das governanças de dados ou de informação. Neste contexto, os autores relatam maior predisposição do ambiente para o consenso de políticas e medidas com relação aos dados. Além disto, a cultura de TI baseada na informação e no conhecimento de TI podem contribuir positivamente para a adoção do modelo de governança de dados ou de informação.

Tallon et al. (2014) indicam que a capacidade de obter informações sobre os dados é estabelecida pela capacidade de serem criados mecanismos adequados de governança, contemplando estruturas e políticas para uso e proteção dos dados. A tendência no aumento do volume de dados corresponde a um incentivo para a adoção da governança de dados ou de informação, sendo os projetos de *analytics* um dos principais fatores que culminam com a maior manipulação de dados na organização. As ideias que permeiam este parágrafo levam ao questionamento se a capacidade analítica mais avançada de uma organização pode corresponder a um fator determinante para o processo de adoção das governanças de dados ou de informação. Tal questionamento foi reforçado mais adiante por autores da inteligência competitiva, embasando a primeira proposição de pesquisa.

Ainda com base no Quadro 2, Tallon (2013) indica que a estrutura centralizada de uma organização viabiliza a adoção da governança de dados ou de informação, pois ela estabelece um ambiente organizacional com menos conflito e maior aderência às políticas definidas. Em particular, a centralização de TI pode diminuir a criação de políticas específicas dentro das unidades de negócios (WEBER et al., 2009; TALLON, 2013).

A padronização da TI e a integração dos processos facilitam na adoção das práticas de governança de informação. Em contrapartida, a existência e o uso de sistemas de legado levam as organizações a manterem múltiplas bases de dados e com sistemas desconexos, prejudicando na formulação de uma única identidade para a governança de dados ou de informação (TALLON, 2013). A cultura de acumuladores também foi notada como um inibidor para a adoção da governança de informação, pois grande volume de dados incompletos e desqualificados podem permanecer na organização por tempo indeterminado (TALLON, 2013; TALLON et al., 2014).

A oferta complexa de produtos e serviços foram notados por Tallon et al. (2014) como limitadores para a adoção das governanças de dados ou de informação. Isto porque a heterogeneidade deste ambiente contempla múltiplos requisitos que podem prejudicar o estabelecimento comum das políticas e das práticas de governança. Por complexidade dos negócios compreende-se o portfolio de produtos ou de serviços, quão diverso eles são uns dos outros e quão específicos são os esforços para desenvolvimento, produção, comercialização dos produtos. O tamanho da empresa também se relaciona à complexidade organizacional, pois se

aumenta a quantidade de fornecedores, produtos intermediários, matérias-primas, produtos acabados, dentre outros itens. Além disto, Tallon (2013) e Tallon et al. (2014) citam a regulação setorial como um influenciador para o aumento do volume de dados e uma das principais forças para o estabelecimento da governança de informação.

No contexto com *big data*, a afirmação de que "quanto mais, melhor" pode acarretar em dificuldades organizacionais, pois os custos para armazenamento dos dados dobram a cada ano com o aumento do volume de dados (TALLON et al., 2014). Assim como gastos com energia, manutenção e *software* crescem com o aumento dos dados armazenados nos *data centers* (TALLON, 2013), indo na contramão do barateamento dos itens de *hardware*. Tallon (2013) cita que o custo total de manutenção dos *data centers* pode variar de 5 a 7 vezes com relação ao custo de aquisição de *hardware*. Assim, a instituição de políticas, diretrizes, normas e procedimentos para regular e controlar os ativos de dados numa organização almeja definir aquilo que é bom, pelo período de tempo necessário, maximizando os benefícios para a corporação. Neste esforço, incluem-se as práticas de governança de dados e de informação.

Para contribuir para a apresentação dos conteúdos da revisão de literatura, a Figura 4 ilustra a estrutura do Capítulo 2.

Figura 4 – Estrutura do capítulo de revisão de literatura.

2.1 Inteligência competitiva e capacidades analíticas

"A dificuldade não está mais na geração de informações, mas, sim, na etapa de análise de conteúdo para orientar ações futuras de forma diferenciada" (McAFEE; BRYNJOLFSSON, 2012)

A literatura de Inteligência Competitiva, IC, *competitive intelligence*, em inglês, introduz e delimita o ambiente das governanças de dados e de informação. Corresponde a uma disciplina emergente após os anos 1980 e mais antiga ao tópico da governança de dados ou de informação (DU TOIT, 2003). As governanças de dados ou de informação surgem principalmente após os anos 2005 (FURLAN; LAURINDO, 2017b).

A inteligência competitiva corresponde a um campo estratégico para a organização, contemplando as etapas de coleta dos sinais do mercado no ambiente competitivo, as análises dos dados e a disseminação dos conteúdos para fundamentar tomada de decisões (CHEN et al., 2012). A IC monitora o ambiente externo da empresa para eticamente adquirir informações relevantes para o processo de tomada das decisões e manter-se um passo à frente da concorrência (TEO; CHOO, 2001; CHEN et al., 2002).

Figura 5 - Visão holística da IC. Fonte: Adaptado de Cabral Netto (2011)

Ser capaz de obter informações sobre os dados e usá-los nas tomadas de decisões tornou-se cada vez mais importante, como enfatizou uma recente edição especial da *MIS Quarterly* sobre questões de transformação de grandes dados e análises em negócios em rede (LISMONT et al., 2017). A Figura 5 representa o ciclo holístico da inteligência competitiva proposto por Cabral Netto (2011), o qual é composto pelos campos de estratégica competitiva da corporação; o monitoramento do ambiente externo; o *business Intelligence* com seus conteúdos de análise de dados e de informação para auxiliar as tomadas de decisão; a tomada de decisão; e a gestão do conhecimento para disseminar o aprendizado e o conhecimento organizacional.

O termo inteligência competitiva difere do tópico inteligência de negócios, esta abreviada por BI, pois no inglês é *Business Intelligence*. Como pode ser visto pelo ciclo holístico da IC proposto por Cabral Netto (2007), a inteligência de negócios corresponde a um campo da IC voltado para a captura e análise dos dados coletados no ambiente competitivo. Ela consiste num conjunto de sistemas de informação e de tecnologias que objetivam fornecer informações relevantes e conhecimento aos tomadores de decisão. Tais subsídios são oriundos das análises de dados que provém de diferentes fontes (SUN et al., 2015a).

Numa escala maior, a inteligência competitiva contém atributos de processos e de produto que, no caso, é a inteligência (BERNHARDT, 1994; BOSE, 2008). Refere-se a um processo quando há transformação dos dados ou das informações em inteligência (KAHANER, 1996) e, produto, quando se tem inteligência para fornecer vantagens competitivas e propiciar melhores desempenhos organizacionais (TEO; CHOO, 2001).

A primeira publicação sobre BI na Scopus aconteceu em 1975, embora o termo tenha sido estabelecido em 1958 por cientistas da IBM (CHEN et al., 2012; SUN et al., 2015a). O tema se desenvolveu principalmente nos anos 1990, onde as publicações sobre BI permeiam o tempo mesmo com a inserção de outros termos como o “*big data analytics*” e o “*business analytics*” (LIM et al., 2013).

O *big data analytics* se refere a um fenômeno emergente, cujas publicações se iniciam em 2010 e apresentam elevado crescimento nos anos subsequentes (vide Figura 6). O *big data analytics* engloba o uso de tecnologias emergentes para manipular e extrair valor de *big data* (SUN et al., 2015a), com o intuito de coletar, organizar e analisar os dados para a descoberta de padrões, conhecimento, inteligência ou outras informações (SUN et al., 2015a; SUN et al., 2015b). O

business analytics está atrelado aos principais componentes analíticos da BI. E, embora tenha surgido em meados dos anos 1950 (DAVENPORT, 2013), passa a ser utilizado com mais frequência no contexto do *big data* (DAVENPORT, 2006; CHEN et al., 2012).

Figura 6 – Evolução das publicações em inteligência de negócios e *big data analytics* (dados retirados em Agosto de 2017). Fonte: ISI Web of Science; Scopus

Observando a evolução das publicações em BI ou *big data analytics* no campo da governança de dados ou de informação, a Figura 7 ilustra amostras de publicação explorando: BI e governança de dados ou governança de informação; *big data analytics* e governança de dados ou governança de informação; e total das publicações em governança de dados ou de informação.

Figura 7– Publicações sobre governança de dados ou de informação e a combinação deste campo com a inteligência de negócios e o *big data analytics* (dados de Agosto de 2017). Fonte: ISI Web of Science; Scopus.

Observando os retornos, foi possível identificar que o *h-index* das publicações em *big data analytics* da ISI *Web of Science* é de 30, ao passo que a de BI é 44. No entanto, dividindo-se a soma total das citações de um determinado tema, com o total publicado neste mesmo campo, somente para os retornos da ISI *Web of Science*, tal divisão para o *big data analytics* é de 2,966; para BI 3,3318 e, para governança de dados ou de informação, um valor superior, equivalente a 4,669. Isto traz a luz que as publicações em governança de dados ou de informação possuem maior quantidade de citações por artigo publicado. Ainda na ISI *Web of Science*, a pesquisa do campo governança de dados ou governança de informação apresentou *h-index* de 15, com total de publicações de 251 e soma das citações de 1172. Este parâmetro calculado é ainda maior na amostra referente à BI e governança de dados ou de informação (4,909). Esta amostra, em específico, conferiu 132 itens publicados na ISI *Web of Science*, com *h-index* de 13 e soma das citações igual a 648.

Com vistas ao parâmetro calculado (razão da soma total de citações com o total publicado), as publicações no campo inteligência de negócios e governança de dados ou de informação apresentam média de citação por item publicado acima da média, dentre demais campos isolados ou combinados pesquisados. Isto retrata a relevância da atual pesquisa, ao abordar análises sobre a inteligência de negócios e a governança de dados ou de informação.

Na sessão seguinte será apresentada a evolução das competências analíticas das organizações, a qual fundamenta parte das variáveis de pesquisa desta tese. Posteriormente, a revisão focaliza o campo da governança, contemplando com mais detalhes a governança de dados e de informação.

2.1.1 Evolução das capacidades analíticas

As capacidades analíticas organizacionais são eminentes dos investimentos em infraestrutura de TI aliado ao desenvolvimento de competências organizacionais, pessoais e gerenciais. Esta composição permite que a organização incorpore os resultados gerados pela *analytics* ao processo decisório da organização (ROSS et al., 2013).

No âmbito tecnológico, focalizam-se na geração, manipulação e análise de dados ou informações. Além das fontes externas da corporação de dados e de informações, eles podem ser gerados internamente na organização como sub-produto de processos de negócios, de operações ou inter-departamentais (BEKOOZ, 2017). No âmbito das fontes externas, a IoT (em inglês, *Internet of Things*) decisivamente tem contribuído para o aumento na geração de dados sobre indivíduos. Além disto, há mercado de compra e venda de dados para o enriquecimento dos dados, podendo contemplar dados sobre consumidores, fornecedores, dentre outros (BEKOOZ, 2017).

O termo "datatificando" ilustra o fenômeno de digitalização do mundo físico a partir da "Internet das Coisas", a IoT ou a "Web das Coisas" (em inglês, *Web of Things*, WoT) (BATRA, 2014; MANYIKA et al., 2015). A IoT aproxima os mundos físico e virtual, contribuindo para a gestão sob a ótica preditiva (PORTER; HEPPELMANN, 2014; SUN et al., 2014). Ela se fundamenta na conectividade e na inteligência dos produtos (PORTER; HEPPELMANN, 2014), cujas tecnologias podem incluir RFID, *wireless*, aparelhos móveis e sensores (XU et al., 2014).

O termo "*big data*" se refere ao alto volume de dados virtuais complexos, diversos, heterogêneos, provenientes de múltiplas e autônomas fontes, com controles distribuídos e descentralizados (McAFEE; BRYNJOLFSSON, 2012; WU et al., 2014; ZHANG, X. et al., 2014; FURLAN, LAURINDO, 2016). Este fenômeno engloba a geração de dados estruturados, semi-estruturados e não estruturados em alto volume, em alta velocidade e em múltiplas fontes (RUSSON, 2011; McAfee e BRYNJOLFSSON, 2012; CHEN et al., 2012), podendo ser considerado a "mina de ouro do século 21" (SUN et al., 2015b).

Uma definição prática e bastante disseminada sobre o termo "*big data*" corresponde ao *framework* dos 3 V's (volume, variedade e velocidade):

- Volume: Este V retrata a quantidade de dados gerados no mundo, sendo cada vez maior devido ao maior acesso da população à *internet* e aos dispositivos móveis;
- Variedade: os dados podem assumir diferentes formas e serem provenientes de distintas fontes (META GROUP, 2001; GANDOMI; HAIDER, 2015; NEVES et al. 2016). O *big data* pode contemplar dados estruturados que são aqueles que se repetem em um determinado padrão; e dados semi-estruturados ou

não estruturados que fogem da padronização, compreendendo recursos textuais, imagens, áudio, dentre outros sinais (IBM, 2014);

- Velocidade: este V contempla as diferentes taxas que os fluxos de dados podem atingir na entrada ou na saída dos sistemas.

Complementarmente aos V's acima listados, constam na literatura demais características, como:

- Veracidade - característica proposta pela IBM que se remete à qualidade do dado, ou seja, quão real e verdadeiro são seus conteúdos e sua origem. Refere-se à confiabilidade do dado, contemplando aspectos de integridade e disponibilidade. As organizações devem assegurar que os dados e suas análises sejam corretas (NEVES et al., 2016; GANDOMI; HAIDER, 2015);
- Variabilidade: a variabilidade consiste numa característica introduzida pela SAS que se refere à variação na geração no fluxo dos dados. Já a complexidade está atrelada à origem do dado ou da informação, pois estes podem advir de múltiplas fontes e podem requerer, portanto, tecnologias para conectar com demais fontes; limpá-los; transformá-los; dentre outras necessidades (GANDOMI; HAIDER, 2015);
- Valor: esta característica foi introduzida pela Oracle e se refere à qualidade e ao potencial de uso dos dados para se obter benefícios e então diferenciação no mercado competitivo (GANDOMI; HAIDER, 2015).

As publicações sobre *big data* emergem principalmente a partir de 2012, onde os principais artigos sugeridos pelo ISI *Web of Science* incluem: Kambatla et al. (2014), Wu et al. (2014), Zhang, J. et al. (2014), Varian (2014), Zhang, X. et al. (2014) e Kourtessis et al. (2014). Uma das razões pelo significativo aumento nas publicações deste tópico consiste em sua multidisciplinariedade, englobando diferentes áreas do conhecimento, como a computação, matemática, biologia, saúde, medicina, negócios (SUN et al., 2015a; FURLAN; LAURINDO, 2015). Furlan e Laurindo (2017a) observam que, em específico para o campo do *big data analytics*, existem grupos epistemológicos distintos na comunidade científica, explorando:

- Mineração dos dados (em inglês, *data mining*);
- Análise dos dados e a geração de conhecimento;

- Gestão, estratégia e negócios;
- Influência da tecnologia no comportamento humano e nas alterações sociais;
- *Internet das Coisas*.

McAfee e Brynjolfsson (2012) retratam este fenômeno como revolucionário do ponto de vista da gestão, contribuindo para a alteração das estruturas organizacionais e da dinâmica mercadológica (BOSE, 2008; SUN et al., 2015b). No entanto, a importância dos dados para os negócios tem sido abordada na literatura desde a década de 80, quando Jenster (1987) trouxe à luz a relevância dos dados e das informações para o planejamento estratégico das organizações. Fazer uso de *big data* e "monetizá-los" corresponde a um desafio para as corporações (RANSBOTHAM et al., 2015). Estima-se que 99% dos dados da IoT não foram utilizados para gerar valor para o ano de 2015 (MANYIKA et al., 2015).

No contexto *big data*, novos temas organizacionais tornam-se imprescindíveis para governar e utilizar os ativos de dados em crescimento de volume e relevância estratégica. O trato com informações externas e livres é particularmente desafiador do ponto de vista governamental (KOOOPER et al., 2011).

No ambiente *big data*, almeja-se a compreensão dos fatos com maior granularidade; encontrar padrões antes não perceptíveis; criar produtos e serviços (DAVENPORT et al., 2012; KEMP, 2014). O não uso dos ativos de dados pode ser considerado uma desvantagem competitiva (RIBARSKY; WANG; DOU, 2014). A característica dos contextos que manipulam *big data* é a previsibilidade, em tempo real, dos acontecimentos no presente e no breve espaço futuro de tempo, permitindo antever futuras tomadas de decisão (LAVALLE et al., 2011). Estabelece-se um outro nível de rapidez no acesso aos dados à medida que são criados, cuja resposta possivelmente deve contemplar desdobramentos futuros antes mesmo deles terem acontecido (LAVALLE et al., 2011). A investigação de fatos passados deixa de ser oportuna e garantia de obtenção de vantagens competitivas.

O *big data* foi eminente da expansão das redes computacionais e pelo avanço nos campos da geração, coleta, armazenamento, visualização e análise de dados (CHEN et al., 2012; SUN et al., 2014; ZHANG, X. et al., 2014; CHEN et al., 2014). As fontes geradoras são diversas e incluem, por exemplo, os cliques na *web*, os conteúdos das mídias sociais (*blogs, tweets, facebook*), os vídeos, as vozes de *call*

centers, até mesmo dados genômicos de banco de dados de pesquisas biológicas e da área da medicina (DAVENPORT et al., 2012).

No início dos anos 2000, Chen et al. (2002) identificaram que grandes corporações apresentavam unidades de inteligência competitiva formais e bem organizadas, almejando contribuir com as tomadas de decisões gerenciais sobre assuntos relevantes de negócios. Antes do ambiente *big data*, os dados de mercado e de clientes, por exemplo, restringiam-se aos departamentos de *marketing*. No contexto com *big data*, ocorre a descentralização da posse dos dados, podendo os vários departamentos da organização manipular e colaborar com demais áreas para agregar valor às informações (RANSBOTHAM et al., 2015).

A centralização dos dados nos departamentos de *marketing* foi resultado da concentração dos sistemas de BI nestas áreas da organização (CHEN et al., 2002; RANSBOTHAM et al., 2015), cujo acesso às informações se baseavam em relatórios impressos e outras formas de informação escrita (CHEN et al., 2002). Os acessos às bases de dados e aos *softwares* de *business intelligence* limitavam-se ao público gerencial (MIMS, 2015). No contexto com *big data*, os dados extrapolam o armazenamento departamental, conferindo maior necessidade pela adoção de políticas de governança (DAVENPORT et al., 2012).

Figura 8 - Relação de equivalência de termos entre autores que exploram os campos da BI, *analytics*, e *big data*. Fonte: elaborado pela autora.

Ainda na revisão de literatura dos campos da inteligência competitiva, é possível visualizar possíveis relações entre tópicos explorados por relevantes autores no campo da inteligência de negócios e *big data analytics*. Algumas variações são encontradas para designar fenômenos correlatos, como pode ser visto pela Figura 8. Já Bose (2009) classifica a BI em duas etapas, como pode ser observado pela Figura 9:

- Operacional: esta fase da inteligência de negócios se configurou antes do contexto *big data*. Foi demarcada pelo intenso uso dos dados estruturados, com foco na operação e no aumento da eficiência;
- Analítica: esta fase emerge com o ambiente *big data* e se estabelece pelo uso intensivo de dados estruturados e não estruturados, com foco em técnicas analíticas avançadas para criação de valor, minimização de riscos, e otimização.

Figura 9 - Fases da inteligência de negócios. Fonte: Bose (2009)

Na fase da BI analítica, Bose (2009) reforça o uso da expressão "*advanced analytics*" (AA). Este termo se refere ao uso de ferramentas para adquirir e analisar massivas quantidades de dados, de modo a prever possíveis resultados futuros e respostas (DAVENPORT, 2006). São geradas análises avançadas com fundamentação descritiva, preditiva e a prescritiva (SUN et al., 2015a). Isto acontece, no entanto, após as etapas de busca e de preparo dos dados, utilizando a mineração de dados e a integração dos dados (BOSE, 2009).

Com outra denominação, McAfee e Brynjolfsson (2012) utilizam as expressões "*analytics*" e "*big data*" para, respectivamente, referirem-se à BI operacional e à BI analítica de Bose (2009). Para McAfee e Brynjolfsson (2012), a fase *analytics* fomentou o desenvolvimento de técnicas mais rigorosas para a tomada de decisão, que se estabeleceram na fase seguinte, a *big data*, contribuindo para a simplificação dos ambientes de negócios à medida que os torna mais poderosos.

Chen et al. (2012) utilizam a expressão "*Business Intelligence and Analytics (BI&A)*", definindo-na em três momentos: 1.0, 2.0 e 3.0. Com relação aos demais autores, a BI&A 1.0 se refere à BI operacional de Bose (2009) e a *Analytics* de McAfee e Brynjolfsson (2012), ao passo que a BI&A 2.0 e 3.0 referem-se a BI analítica e a *big data* de McAfee e Brynjolfsson (2012).

Chen et al. (2012) aproximam os campos de inteligência de negócios e *analytics* pela expressão BI&A. Esta se refere as técnicas, tecnologias, práticas e aplicações que analisam dados críticos para os negócios, e que ajudam as organizações a melhor compreender seus negócios e o mercado (CHEN et al., 2012; LIM et al., 2013). O Quadro 3 lista as principais características propostas por Chen et al. (2012) para a BI&A 1.0, 2.0 e 3.0.

Quadro 3 - Caracterização de Chen et al. (2012) sobre a evolução da BI&A.

BI&A	Momento	Principais Características
1.0	1990 - presente (tecnologias e aplicações atualmente adotadas pelas indústrias podem ser consideradas BI&A 1.0)	Dependência de tecnologias para coleta de dados, extração e análises; uso de dados estruturados, coletados através de diferentes sistemas de legado; as técnicas analíticas se popularizaram em 1990 e se baseiam em métodos estatísticos dos anos 1970; muitas das tecnologias incorporadas nas plataformas comerciais de BI.
2.0	~2000 - presente (<i>internet</i> e mídias sociais)	Conteúdos não estruturados obtidos pela <i>internet</i> (<i>web intelligence</i> e <i>web analytics analytics</i>); mídias sociais (<i>social media analytics</i>); volume significativo de <i>feedbacks</i> e opiniões; as tecnologias devem ser integradas aos sistemas de mineração de texto
3.0	~ 2010 - presente (aplicativos móveis e sensores)	Conteúdo de dispositivos móveis (celulares, <i>tablets</i>), de sensores diversos, como RFID (<i>mobile analytics</i> ; <i>Internet das coisas</i>).

Fonte: Chen et al. (2012)

O Quadro 4 lista as principais características propostas por Davenport (2013), onde é possível notar sua classificação proposta para a *analytics* nas fases 1.0, 2.0 e 3.0.

Quadro 4 - Caracterização de Davenport (2013) sobre a *analytics*.

<i>Analytics</i>	Momento	Principais Características
1.0	Era do <i>business intelligence</i> (~1950 - ~2005)	Pela primeira vez os dados de processos, produção, vendas, interação com clientes foram agregados e analisados; Novas tecnologias computacionais; Era da <i>enterprise data warehouse</i> (capturam informações) e de <i>softwares</i> de BI (reportavam dados); Novas competências (habilidade de gestão de dados); Diminuição da gestão com base na intuição; Muito tempo envolvido no preparo dos dados para análise; Análises demoradas (semanas ou meses); Análises baseadas em fatos passados e sem explicações ou predições (futuro); Sistemas de informação eram construídos de forma personalizada para grandes empresas cuja larga escala justificavam o investimento. Em outros momentos os mesmos sistemas eram comercializados por fornecedores externos em formatos mais genéricos; Relatórios em branco e preto.
2.0	Era do <i>big data</i> (~2005)	Iniciada com o estabelecimento de companhias de <i>internet</i> e de redes sociais (Google, eBay, etc); Os dados não são gerados apenas por sistemas internos das empresas (como na 1.0); Dados provêm de múltiplas fontes, como internet, sensores, iniciativas públicas; Necessidade de ferramentas mais poderosas e tecnologias inovadoras; Companhias necessitam de novas competências; <i>Big data</i> não conseguem ser processados por único processador - processados com servidores paralelos, utilizando Hadoop; Novas base de dados (NoSQL) para manipular dados não estruturados; Ambiente de computação em nuvem; Relatórios coloridos e com visual distinto; Necessidade de cientista de dados (habilidades computacionais e analíticas)
3.0	Oferta de dados enriquecidos	A competição não se restringe mais no uso de dados para a tomada de decisão mas sim em seu uso para a criação de produtos e serviços mais valiosos; Organizações terão que integrar grande e pequenos volumes de dados, de fontes internas e externas, em formatos estruturados e não estruturados, de modo a ganhar novas percepções proveniente de modelos preditivos e prescritivos; Uso de todas as tecnologias desenvolvidas na <i>Analytics</i> 1.0 e 2.0; O desafio é adaptar os processos operacionais, de desenvolvimento de produto e de tomada de decisão para tirar vantagem dos benefícios dessa era.

Fonte: Davenport (2013)

Para o contexto do *big data*, LaValle et al. (2011), McAfee e Brynjolfsson (2012), Ross et al. (2013) e Sun et al. (2015a) fornecem recomendações quanto as boas práticas para as organizações extraírem valor dos dados ou das informações:

1. Definir uma estrutura para tomadas de decisões, onde estão alocadas corretamente as informações e as responsabilidades (McAFEE; BRYNJOLFSSON, 2012)
2. Tornar a cultura organizacional baseada em evidências (*data-driven*) (McAFEE; BRYNJOLFSSON, 2012; ROSS et al., 2013);

3. Fomentar a comunicação e a interação entre *stakeholders* da organização (LAVALLE et al., 2011; McAfee; BRYNJOLFSSON, 2012);
4. Definir os propósitos de negócios preliminarmente à coleta de dados e análises (LAVALLE et al., 2011);
5. Utilizar os resultados analíticos nas decisões estratégicas e de negócios (LAVALLE et al., 2011; McAfee; BRYNJOLFSSON, 2012; ROSS et al., 2013);
6. Realizar investimentos em infraestrutura de TI (ROSS et al., 2013);
7. Fomentar o alinhamento entre a estratégia de negócios e a *analytics* (LAVALLE et al., 2011);
8. Utilizar o conhecimento (e a experiência) organizacional de manipular os dados armazenados nos sistemas operacionais (convencionais) para operar e extrair valor de *big data* (ROSS et al., 2013);
9. Assegurar que os resultados analíticos sejam compreensíveis aos usuários (LAVALLE et al., 2011);
10. Manter competências tecnológicas e gerenciais existentes – a adição de novas ferramentas ou competências não necessariamente requer a extinção daquelas usuais (LAVALLE et al., 2011);
11. Utilizar tecnologias emergentes para manipular *big data* (McAfee; BRYNJOLFSSON, 2012; SUN et al., 2015a).

No ambiente *big data*, McAfee e Brynjolfsson (2012), Lavallo et al. (2011) citam que os desafios tecnológicos são significativos, mas que os gerenciais e culturais são ainda maiores. McAfee e Brynjolfsson (2012) apontam que as organizações não conseguirão fazer a transição para o ambiente *big data* se elas não conseguirem gerenciar as mudanças de forma eficiente. Dentre os desafios apontados pelos autores, compreendem-se o processo de tomada de decisões e a cultura organizacional. No campo das tomadas de decisões, deve-se estabelecer uma estrutura organizacional para que os tomadores de decisão possuam em mãos os dados certos.

A cultura organizacional baseada em evidências (*data-driven*) indica a predisposição da organização para se voltar aos seus ativos de dados. Neste caso, McAfee e Brynjolfsson (2012), Ross et al. (2013) citam a necessidade da cultura *data-driven* para embasar as tomadas de decisão com base em fatos reais. No campo do *big*

data, McAfee e Brynjolfsson (2012) concluem que quanto mais as organizações se consideram *data-driven*, maior é o desempenho financeiro com relação às aquelas menos *data-driven*.

LaValle et al. (2011), McAfee e Brynjolfsson (2012) determinam a necessidade da comunicação, interação e colaboração entre *stakeholders* da organização. A liderança deve criar um ambiente de cooperação entre funções para troca de *expertise* (McAFEE; BRYNJOLFSSON, 2012). Além disto, promover o alinhamento entre negócios e *analytics* é fundamental para extrair valor dos dados.

LaValle et al. (2011) enfatizam a necessidade de os objetivos dos negócios serem definidos preliminarmente à coleta dos dados e definir o que se almeja responder com a *analytics*. Os dados e as informações devem ser coletados de acordo com as perguntas de negócios e não o oposto. Deve-se também assegurar que os resultados alcançados pela *analytics* sejam utilizados na prática organizacional para fundamentar a tomada de decisões (LAVALLE et al., 2011; McAFEE; BRYNJOLFSSON, 2012; ROSS et al., 2013).

Ross et al. (2013) complementam com o aporte financeiro que a organização realiza sobre as ferramentas tecnológicas. Os investimentos em TI correspondem a um elemento relevante para a definição da competência analítica da organização, para uso e extração de valor de *big data*. No entanto, tal investimento deve vir acompanhado das mudanças culturais e nas competências organizacionais para permitir o maior uso dos ativos de dados.

No âmbito das competências e práticas, Sun et al. (2015a) indicam o uso de tecnologias emergentes para manipular, armazenar, utilizar *big data*, como é o caso de Hadoop (um *software* livre para armazenamento de dados estruturados, não estruturados ou semi estruturados), Apache SparkTM (para processamento de *big data*). Em específico, LaValle et al. (2011) mencionam que adicionar competências para manipular *big data* não necessariamente requer que as antigas sejam eliminadas. Ross et al. (2013) salientam que as organizações devem inicialmente aprender como utilizar os dados existentes em seus sistemas operacionais, para quem então possam explorar valor de *big data*. Além disto, seus resultados devem ser compreensíveis aos usuários.

Para LaValle et al. (2011) e Ross et al. (2013), a criação de valor de *big data* acontece em casos organizacionais com maior capacidade analítica, o que é alcançado pela combinação de elementos culturais, gerenciais e tecnológicos. A

Figura 10 retrata os níveis de capacidade analítica das organizações segundo LaValle et al. (2011). Este modelo foi utilizado como referência para avaliar a capacidade analítica da corporação.

Figura 10 – Modelo para classificar as capacidades analíticas das organizações. Fonte: Adaptado de LaValle et al. (2011)

As competências de BI podem existir em organizações voltadas a serviços ou manufaturas. No caso de organizações do setor financeiro, as competências em BI estão distribuídas dentre os vários segmentos de negócios. Para o caso de bancos, alguns exemplos de sistemas de BI correspondem ao CRM – *Customer Relationship Management*, PM – *Performance Management*, RM – *Risk Management*, ALM – *Asset and Liability Management* (UBIPARIPOVIC; DURKOVIC, 2011).

Ubiparipovic e Durkovic (2011) citam alguns exemplos da aplicação de BI em bancos. Para as soluções de BI para ALM (Gerenciamento de Ativos e Passivos), Ubiparipovic e Durkovic (2011) citam que as soluções devem gerar um conjunto

completo de relatórios internos a partir de balanços, análise de liquidez fluxo de caixa, estrutura da taxa de juros, dentre outros elementos. Além disto, Ubiparipovic e Durkovic (2011) definem o RM (Gerenciamento de Riscos) como um processo no qual um banco metodologicamente gerencia todas as fases de processamento de risco (identificação, análise, medição, controle e relatórios) que representam uma ameaça aos seus objetivos e atividades empresariais individuais. O nível de risco alcançado não deve colocar em risco operações seguras e estáveis do banco. Alguns dos riscos enfrentados pelos bancos incluem riscos de crédito, riscos de mercado, riscos de taxa de juros, riscos cambiais estrangeiros, riscos de liquidez, riscos operacionais, riscos de reputação, etc (UBIPARIPOVIC; DURKOVIC, 2011). Com vistas à revisão de literatura realizada sobre inteligência competitiva, BI e *big data analytics*, foi possível constituir o Quadro 5 com o resumo teórico apresentado.

Quadro 5 - Quadro teórico da revisão da literatura de inteligência competitiva, inteligência de negócios e *big data analytics*.

Conteúdo	Descrição	Teoria
Extração de valor	A inteligência competitiva delimita o ambiente de extração de valor dos dados, como proposto por Cabral Netto (2007). A extração de valor dos dados e uso para tomada de decisões é previsto por diversos autores do campo da inteligência competitiva, inteligência de negócios, <i>big data analytics</i> .	Teo e Choo (2001); Chen et al. (2002); Davenport (2006); Cabral Netto (2007); Bose (2008); LaValle et al. (2011); McAfee e Brynjolfsson (2012); Davenport (2013); Ross et al. (2013); Koronios et al. (2014); Sun et al. (2015a); Sun et al (2015b); Ransbotham et al. (2015); Lismont et al (2017); Berkooz (2017)
Recomendações das boas práticas	Recomendações das boas práticas para as organizações extraírem valor dos dados no ambiente <i>big data</i> .	LaValle et al. (2011), McAfee e Brynjolfsson (2012); Ross et al. (2013) e Sun et al. (2015a)
Capacidades analíticas	Desenvolvimento das capacidades analíticas da organização e o aumento da extração de valor nos dados. Permite-se definir a capacidade analítica das organizações com base nos retornos da <i>analytics</i> para a organização.	LaValle et al. (2011); Ross et al. (2013); Lismont et al. (2017)

Fonte: elaborado pela autora

Além disto, o Quadro 6 lista os principais conceitos discutidos no âmbito da inteligência competitiva.

Quadro 6 – Definição dos principais termos utilizados na sessão.

Termo	Definição
Inteligência competitiva (IC)	Engloba as atividades internas e externas à empresa que visam o monitoramento do ambiente competitivo e a geração de valor para a empresa; inclui as atividades de estratégia competitiva; tomada de decisão; monitoramento do ambiente externo; <i>business Intelligence</i> ; e a gestão do conhecimento (CABRAL NETTO, 2011).
<i>Business intelligence</i> (BI)	Engloba ferramentas de <i>software</i> manipulam dados para gerar informações e/ou conhecimento para a empresa; conjunto de sistemas de informação e de tecnologias que objetivam fornecer informações relevantes e conhecimento ao tomadores de decisão.
<i>Big data analytics</i>	Estabelece-se sobre as tecnologias emergentes para lidar com <i>big data</i> (SUN et al., 2015a); contempla o processo de coleta, organização e análise de <i>big data</i> para a descoberta de padrões, conhecimento, inteligência ou outras informações, dentro do contexto do <i>big data</i> (SUN et al., 2015a; SUN et al., 2015b).
<i>Business Analytics</i>	Corresponde a um braço da BI centrado em dados, e que utiliza uma série de tecnologias e de métodos avançados para a coleta, extração e análise de dados (LIM et al., 2013).
<i>Big data</i>	Contempla o fluxo de fatores brutos ou não, isolados ou não, podendo contemplar elementos constituintes da informação (BRAGANZA, 2004; ZINS, 2007); podem ser virtuais, complexos, diversos, heterogêneos, estruturados, semi-estruturados ou não-estruturados, provenientes de múltiplas e autônomas fontes, com controles distribuídos e descentralizados (McAFEE; BRYNJOLFSSON, 2012; WU et al., 2014; ZHANG, X. et al., 2014; FURLAN, LAURINDO, 2016).
Dados e informação	Podem contemplar fluxos de fatores brutos; isolados ou não; desorganizados ou não; podem ser estruturados, semi estruturados ou não estruturados; podem ter significado ou não; podem ser tratados ou não; podem contemplar ou não uma finalidade específica na organização" (BRAGANZA, 2004; LAURINDO, 2008).

Fonte: elaborado pela autora

2.2 Estrutura da governança de tecnologia da informação

A governança é comumente instituída através de um *framework* hierárquico baseado em princípios, políticas, responsabilidades e procedimentos que devem assegurar determinado nível de controle no âmbito organizacional (KOOPER et al., 2001).

Kooiman (2003) identifica outros dois tipos de governança: a "co-governança" ("*co-governance*") e a "governança própria" ("*self-governance*"). Segundo o autor, o elemento principal e comum na co-governança é a autonomia e a identidade das partes envolvidas - elas possuem algo em comum para perseguir. Ela consiste numa governança horizontal, isenta de um ator central ou ditatorial. Neste contexto,

os atores se comunicam e colaboram uns com os outros, resultando em maiores níveis de sucesso na troca das informações e na governança.

A governança própria (ou "*self-governance*") se refere à capacidade de entidades sociais se auto governarem. Geralmente possuem acordos informais, regras auto-aplicáveis e códigos de conduta semi-formalizados (KOOIMAN, 2003). Em geral, as comunidades de *internet* se organizam sob a forma de governança própria (KOOOPER et al., 2011).

A governança de TI se preocupa em formular políticas, controlar, gerenciar e monitorar artefatos físicos de tecnologia na organização (RAU, 2004; TALLON et al., 2014). Ela almeja definir quais decisões no âmbito da TI devem ser tomadas na organização, por quem, como, e através de quais processos (WEILL; WOODHAM, 2002; WEILL; ROSS, 2005). Pode-se dizer que a governança de TI trata da "administrabilidade" da área de TI na organização (LAURINDO, 2008), a qual engloba não apenas as tecnologias da BI e do *big data analytics* na organização, mas todo o artefato de *hardware* e *software*.

A governança de TI pode ser vista como um processo ou como uma estrutura: a primeira, como processo, indica um conjunto de sistemas, processos e procedimentos em vigor na organização para garantir a eficiência e a eficácia dos recursos da TI; e a segunda, como estrutura, apresenta-se como um arcabouço para a tomada de decisão nas organizações sobre as questões da TI, com base no estabelecimento de padrões de autoridades, matriz de responsabilidades, direitos de decisão (ASSIS, 2011). A adoção da governança de TI almeja garantir os resultados esperados pela tecnologia na organização, contribuindo para estender os benefícios da tecnologia na organização e aprimorar a percepção da TI na (SCHWARZ; HIRSCHHEIM, 2003).

Weill e Ross (2005) apontam seis estruturas típicas para a tomada de decisão com relação a TI, contemplando organizações altamente centralizadas para aquelas altamente descentralizadas. Como definem os autores, em uma "monarquia de negócios", os executivos seniores ou um grupo de executivos, em geral o CIO tomam as decisões com relação a TI para toda a organização. Trata-se da característica mais centralizada. Em uma "monarquia de TI", tais decisões são tomadas por executivos de TI ou um grupo de executivos de TI. Já em um "sistema federativo", executivos do *C-level* e representantes dos negócios colaboram com o departamento de TI. Em um "duopólio de TI", as decisões são tomadas por

executivos de TI e líderes de negócios representando as unidades de negócios. Em um “sistema feudal”, as unidades de negócios e os líderes de processos tomam decisões separadas no contexto de sua unidade e das necessidades de seus processos. Por fim, o sistema mais descentralizado comentado por Weill e Ross (2005) corresponde a “anarquia”, onde cada usuário ou um pequeno grupo estabelece sua própria agenda com relação à TI.

A definição dos princípios e das práticas da governança de TI são específicos para cada caso organizacional. Para tanto, deve-se considerar parâmetros como o papel predominante da TI na organização ou o grau de deliberação da governança na organização (WEILL; WOODHAM, 2002). Modelos organizacionais foram desenvolvidos para a prática dos negócios, tais como o COBIT (*Control Objectives for Information and Related Technology*) e o ITIL (*Information Technology Infrastructure Library*) (LAURINDO, 2008). Em geral, tais modelos focalizam o gerenciamento dos riscos e a entrega de valor (ITGI, 2003).

Para o ITGI (2003), as áreas viabilizadoras para a adoção da governança de TI são:

- Alinhamento estratégico entre TI e negócio: como definido no título, esta área trata do alinhamento estratégico entre TI e negócio e foi também discutida por Laurindo (2008) como um parâmetro relacionado à eficácia. Esta objetiva o alcance de resultados desejáveis para a organização, almejando "fazer as coisas certas" (LAURINDO, 2008);
- Gerenciamento dos recursos: esta área proposta pelo ITGI (2003) se relaciona aos investimentos em tecnologia na organização, estando relacionado ao conceito de eficiência. Por eficiência, entende-se "fazer bem as coisas" (LAURINDO, 2008).
- Medição do desempenho: esta área proposta pelo ITGI (2003) também está relacionada à eficiência, pois objetiva o monitoramento dos processos internos relativos à implementação das estratégias e projetos. Ela consiste numa área viabilizadora pois almeja o monitoramento do desempenho das políticas e das práticas da governança.

Complementarmente, a academia propõem *frameworks* mais robustos, compostos por domínios ou áreas de decisão, competindo à eles uma função ou competência no exercício da governança. Tais modelos são caracterizados por dimensões, que

podem ser múltiplas para um único domínio (FU et al., 2011). A Figura 11 ilustra os domínios da governança de TI sugeridos por Weill e Woodham (2002), Weill e Ross (2005) e Khatri e Brown (2010). Com relação ao modelo apresentado para a governança de TI, os domínios são:

- Princípios de TI: correspondem as afirmações generalistas de como a TI deve ser utilizada na empresa (WEILL, 2004). Este domínio captura a essência do direcionamento da TI e como ela será utilizada (WEILL; WOODHAM, 2002);
- Arquitetura de TI: corresponde a um conjunto de opções técnicas que orientam a organização para satisfazer as necessidades dos negócios. A arquitetura compreende um conjunto de políticas e regras que governam o uso da TI e definem os caminhos dos negócios (WEILL, 2004). Incluem-se neste domínio: os padrões da tecnologia, o uso dos dados, o desenho de novos aplicativos e a mudança de processos para utilizar novas tecnologias (WEILL; WOODHAM, 2002);
- Infra-estrutura de TI: este domínio aborda como construir os fundamentos da TI na corporação. Suas competências correspondem a um fator crítico na velocidade com que novas iniciativas de negócios podem ser implementadas (WEILL; WOODHAM, 2002);
- Aplicações negócios: este domínio especifica os requerimentos de negócios para as aplicações de TI adquiridas ou desenvolvidas internamente na organização (WEILL, 2004; NFUKA, 2012);
- Investimentos em TI e a priorização: este domínio cobre o processo de tomada de decisão dos investimentos em TI. Isto inclui a priorização de onde alocar os investimentos e suas respectivas justificativas (WEILL; WOODHAM, 2002; WEILL, 2004).

Como forma de executar os objetivos de cada domínio, práticas estruturais, operacionais e relacionais são necessárias. Elas possuem as seguintes características:

- Estruturais: referem-se ao planejamento estratégico da governança de TI; estabelecem a estrutura organizacional da TI na organização, incluindo

comitês, *boards*, dentro outros atores; definem papéis e responsabilidades sobre a TI;

- Procedurais ou operacionais: referem-se à execução das estratégias da governança de TI, às tomadas de decisão, ao controle de custos da área de TI, à alocação de recursos tecnológicos na organização. Contempla as questões de infraestrutura, arquitetura, investimentos em TI, definição das prioridades (TALLON et al, 2014);
- Relacionais: prática que estabelecem o envolvimento da governança de TI com os negócios, que compartilha conhecimentos de TI, troca de ideias, comunicação, gestão de conflitos, colaboração (TALLON et al, 2014).

Figura 11 - Domínios da governança de TI. Fonte: Weill e Woodham (2002); Weill e Ross (2005); Khatri e Brown (2010)

Na literatura, os mecanismos da governança podem ser tratados como práticas da governança, como no trabalho de Simonsson et al. (2010) ou se referirem aos veículos para a implementação da governança, como no trabalho de Weill e Woodham (2002). Na visão deles, os mecanismos da governança são elementos que compõem as práticas ou as sub-práticas, que podem ser específicos de um domínio ou cobrir múltiplos deles (WEILL; WOODHAM, 2002).

Conforme os trabalhos de Weill e Woodham (2002), Rau (2004), Weill e Ross (2005), é comum as organizações adotarem três mecanismos distintos para contribuir para o aumento da eficácia na governança de TI:

- Estrutura de tomada de decisão: os comitês organizacionais correspondem ao principal mecanismo para a governança de TI (WEILL; ROSS, 2005). O comitê executivo visa obter a visão holística dos negócios e da TI na organização; os conselhos de TI devem envolver os gerentes seniores para obter valor aos negócios através da TI; o comitê de arquitetura deve identificar as tecnologias estratégicas e os padrões (WEILL; WOODHAM, 2002);
- Processos de alinhamento: correspondem às técnicas de alinhamento para a dispersão de conteúdos e envolvimento de *stakeholders* nas decisões da governança e de sua implementação (WEILL; ROSS, 2005);
- Comunicações formais: correspondem aos meios de comunicação, que podem incluir: a comunicação geral; a instituição de comitês formais; a comunicação regular da alta gestão. Rau (2004) discute que simplesmente anunciar a necessidade de adoção da governança é insuficiente para garantir a adoção efetiva da governança. O autor sugere que internamente deve ser difundida a comunicação através de reuniões ou *workshops* para o ganho de comprometimento dos indivíduos (RAU, 2004).

Weill e Ross (2005) discutem que sem uma estrutura formal (de TI, no caso, podendo ser de dados ou de informação), os gerentes são levados a resolver os problemas cada um por si e de forma isolada. Além disto, sem a comunicação dos princípios de governança, as ações dos indivíduos na organização podem ser, e frequentemente são, opostas umas às outras (WEILL; ROSS, 2005).

Weill e Ross (2005) apontam que uma grande barreira para a governança efetiva da TI é a falta de compreensão de quais decisões estão sendo tomadas, quais processos estão sendo implementados e quais os resultados são esperados. Weill e Ross (2005) propõem que maiores níveis de comunicação em geral significam maiores níveis de eficácia da governança de TI. A comunicação pode acontecer por diferentes formas, sendo as reuniões de comitês uma delas.

Figura 12 - Estrutura organizacional da governança de TI. Fonte: Rau (2004)

A Figura 12 ilustra a estrutura da governança de TI proposta por Rau (2004). Pelo modelo proposto, é possível observar que o Conselho da Governança de Tecnologia da Informação se reporta ao *board* de diretores, satisfazendo o primeiro critério. Este conselho tem como objetivo desenvolver a estratégia de TI e promover o financiamento dos projetos de tecnologia. Contudo, embora menos frequente, é possível encontrar organizações que possuem duas entidades correspondentes ao Conselho da Governança de TI aqui apresentado. Nestas organizações, uma entidade focaliza os aspectos estratégicos da governança e, a outra, no gerenciamento e no controle da governança de TI na organização (RAU, 2004).

O Conselho da Governança de TI é responsável pelas atividades que permeiam todas as áreas e funções do negócio, que incluem: o fomento de políticas; o controle orçamentário; a medição e a divulgação do desempenho da TI na organização. Ele é suportado por uma série de grupos de trabalho que gerenciam, priorizam, controlam e implementam atividades, sendo eles:

- CIO (Chief Information Officer): Esta entidade é central para a governança, sendo conhecida como o "chefe da TI", ou aquele que dirige as funções de TI. Sua principais atividades incluem: atender as demandas por serviços de

TI eficiente e eficazmente; promover o relacionamento com clientes internos e externos de TI; manter-se atualizado sobre as inovações e tecnologias emergentes que são apropriadas para a organização;

- Escritório de projetos: esta equipe deve desenvolver e aplicar os padrões e os procedimentos relativos aos projetos e programas de TI, bem como monitorar a aderência destes. Reportam-se à eles os gerentes de projetos;
- Customer services: estes agentes comumente são conhecidos por executivos de contas. Contribuem para que o CIO alcance maiores níveis no atendimento aos clientes;
- Equipes de priorização da área do usuário: responsáveis por priorizar e gerenciar o extensivo número de pedidos por serviços de TI pelas áreas de negócios;
- User program managers: O conselho da governança de TI denota um gerente dessa categoria para comandar determinado projeto de TI na organização. Tais gerentes atuam em conjunto com gerentes de projetos, líderes de projetos, dentre outros *stakeholders*;
- Auditores externos / internos: os auditores externos são contadores ou consultores que trabalham com o *board* de diretores para assegurar as compilações financeiras, os relatórios e a integridade dos processos de controles na organização. Os auditores internos possuem o papel na organização de investigar, analisar e reportar à organização se os recursos financeiros estão sendo utilizados de forma eficiente e eficaz.

Como "Fatores Críticos de Sucesso" (FCS) para a adoção da governança de TI, foi possível identificar os trabalhos de Nfuka e Rusu (2011) e Kurti et al. (2014). Apesar das pesquisas contemplando FCS's ainda serem bastante restritas no campo da governança de TI (NFUKA; RUSU, 2011).

Para Rockart (1979), "os Fatores Críticos de Sucesso correspondem ao número limitado de áreas em que os resultados, se satisfatórios, asseguram o desempenho competitivo de sucesso para a organização (ROCKART, 1979, p. 85 *apud* ESTEVES, 2004)". Tais fatores podem compreender aspectos estratégicos, gerenciais ou operacionais da organização (OLSZAK; ZIEMBA, 2012) sob a forma de eventos, condições, circunstâncias, atividades, características ou variáveis com

relevância estratégica (JENSTER, 1987; LEIDECKER; BRUNO, 1984). Os FCS's podem também ser caracterizados como um conjunto de fatores "considerados críticos por gerentes para o sucesso em um determinado contexto" (DADASHZADEH, 1989 *apud* ESTEVES (2004).

O método científico sobre os FCS's foi proposto por Rockart (1979) no final da década de 70 e corresponde a uma das primeiras tentativas de unir o uso da TI aos objetivos e estratégia das empresas (LAURINDO, 2008; CARVALHO, LAURINDO, 2012). No entanto, este método transcende a área de TI e pode ser aplicado às práticas de gestão e de planejamento estratégico em geral (LAURINDO, 2008).

Para o caso da governança de TI, Nfuka e Rusu (2011) e Kurti et al. (2014) propuseram Fatores Críticos de Sucesso (FCS's):

1. Liderança em TI para compreender os objetivos do negócio e as contribuições que a TI pode trazer para a gestão;
2. Envolvimento e suporte dos gerentes *seniores*;
3. Encorajar e apoiar a comunicação e a parceria entre TI e negócio;
4. Engajar *stakeholders* principais;
5. Definir e alinhar as estratégias de TI com as da corporação; implementá-las em *top-down* na organização;
6. Consolidar as estruturas de TI para definir responsabilidade e capacidade de resposta;
7. Consolidar, padronizar e gerenciar a infraestrutura de TI para otimizar custos e o fluxo da informação pela organização;
8. Conscientizar os indivíduos da organização sobre a governança de TI e treiná-los para o uso ótimo da TI;
9. Atrair, desenvolver e reter profissionais competitivos de TI;
10. Consolidar a medição do desempenho e *benchmarks* para rastreamento e demonstração do sucesso;
11. Consolidar, comunicar e fazer cumprir a política e as práticas que propiciem a aquisição e o uso da TI na organização com custo efetivo.

A lista sugerida de FCS's para adoção da governança de TI contempla fatores com diferentes temas e especificidades. Em particular o método "*top-down*" pode acarretar em resistência interna e erros na implementação, caso a organização não tenha conhecimento da real necessidade de adoção da governança de TI.

2.3 Estrutura das governanças de dados e de informação

Por governança de dados ou de informação podem ser contempladas as entidades organizacionais, práticas, políticas, padrões, tomadas de decisão, responsabilidades e métodos com relação aos dados (KOLTAY, 2016). Como menciona Koltay (2016), a governança de dados opera segundo os princípios da integridade, da transparência e da responsabilidade com relação aos dados. Consequentemente, a adoção da governança de dados ou de informação pode diminuir conflitos operacionais, encorajar a adoção de práticas comuns na organização com relação aos dados, reduzir custos, aumentar a efetividade através da coordenação dos esforços e permitir a transparência dos processos (KOLTAY, 2016).

Koronios et al. (2014) e LaValle et al. (2011) ressaltam que a extração efetiva de valor de dados ou de informação no ambiente *big data* requer a combinação de elementos relativos as pessoas, processos e tecnologias. Tais elementos serão apresentados ao longo desta sessão, pois são contemplados dentre as sub-práticas dos modelos previstos pela literatura. Este tripé também foi discutido por Kooiman (2003), que particularmente defende o papel dos *stakeholders* na governança como elemento central para o bom funcionamento.

Koltay (2016) discute que os princípios tradicionais da governança de dados podem ser aplicados no ambiente *big data*. Aliás, o autor menciona que um dos grandes erros dos ambientes *big data* é subestimar os benefícios da governança de dados para a organização.

Os modelos de governança de dados ou de informação caracterizam-se através de domínios, os quais possuem dimensões. Práticas, sub-práticas e mecanismos podem ser lançados para operacionalizar ou viabilizar a implementação dos objetivos definidos para a governança de dados ou de informação. A Figura 13 ilustra uma pirâmide retratando tais atributos do modelo.

Os domínios também podem ser chamados como áreas de decisão e contemplam a visão estratégica da governança, definindo os temas críticos tratados pela governança. Como pode ser visto na sessão anterior, o modelo da governança de TI também se caracteriza através de domínios, onde foram apontados cinco domínios no caso.

Figura 13 – Níveis dos elementos do modelo de governança. Fonte: elaborado pela autora.

As dimensões são elementos ou conteúdos específicos que são empregados para auxiliar a caracterização de um determinado domínio. Os níveis relativos aos domínios e suas dimensões focalizam na idealização dos princípios do modelo de governança, ao passo que os níveis relativos às práticas, sub-práticas e aos mecanismos focalizam na implementação efetiva do modelo de governança.

As práticas podem ser consideradas estruturais, relacionais ou operacionais, como também apresentado na sessão anterior sobre a governança de TI. Já as sub-práticas são itens com maior granularidade que constituem as práticas. Já os mecanismos constituem em elementos estruturais ou veículos que auxiliam na adoção das práticas ou sub-práticas. Weill e Woodham (2002) e Weill e Ross (2005) mencionam que as organizações geralmente adotam três mecanismos, sendo eles: as estruturas de tomada de decisão, os processos de alinhamento e as comunicações formais. O Quadro 7 resume tais níveis.

Quadro 7 – Resumo das definições sobre os níveis no modelo de governança.

Campo	Definição
Domínios	Domínios podem ser também as áreas de decisão da estrutura da governança de dados ou da governança de informação. Khatri e Brown (2010) e Fu et al. (2011) definem cinco domínios para as governanças de dados ou de informação.
Dimensões	São sub-grupos de atividades de cada domínio da estrutura da governança. As dimensões podem ser múltiplas para um único domínio (FU et al., 2011).
Práticas	São as formas pelas quais as organizações executam a governança de dados ou de informação (TALLON, 2013). Elas podem ser estruturais, relacionais ou operacionais.
Sub-práticas	São sub-grupos das práticas; especificam atividades com maior nível de granularidade; são elementos ou conteúdos relacionados à execução das práticas da governança.
Mecanismos	Elementos estruturais ou veículos que auxiliam na adoção das práticas ou sub-práticas, sendo geralmente as estruturas de tomada de decisão, os processos de alinhamento e comunicações formais (WEILL; WOODHAM, 2002; WEILL; ROSS, 2005)

Fonte: elaborado pela autora

A seguir consta a revisão de literatura das governanças de dados e de informação, a qual se encontra estruturada nas seguintes frentes:

- **2.3.1 Domínios:** retrata os modelos de governança de dados e de informação previstos pela literatura, enfatizando seus domínios;
- **2.3.2 Práticas e sub-práticas:** focaliza nas práticas e nas sub-práticas previstas pela literatura.

2.3.1 Domínios

As governanças de dados ou de informação podem ser caracterizadas por domínios ou áreas de decisão, cabendo a eles uma função ou competência no exercício da governança. Com base na revisão de literatura, foram identificados os modelos de Khatri e Brown (2010) e de Fu et al. (2011), como pode ser visto pela Figura 14. A diferença entre eles reside nos domínios "ciclo de vida dos dados", previsto pelo primeiro, e o de "disponibilidade dos dados", previsto pelo segundo.

Figura 14 - Modelos com domínios da governança de dados. Fonte: Fu et al. (2011); Khatri e Brown (2010)

Os "princípios dos dados" contemplam o topo do modelo de governança, em termos de domínios. Tais princípios devem orientar os demais e fornecer uma visão clara dos objetivos da governança, promover o alinhamento entre os ativos de dados e de informações com relação aos negócios, além de promover o uso apropriado dos dados e das informações na organização (KHATRI; BROWN, 2010; FU et al., 2011). Para Khatri e Brown (2010), este domínio almeja responder:

- Quais são os potenciais usos dos ativos dos dados e das informações para o negócio?
- Quais são os mecanismos contínuos para a comunicação dos usos e resultados na organização?
- Quais são os comportamentos desejáveis para uso dos dados como ativos?
- Como são as oportunidades para compartilhamento e reuso de dados?
- Como o ambiente regulatório influencia no uso dos dados nos negócios?

Os "princípios dos dados" demarcam os requerimentos para uso dos dados, que definem os padrões de qualidade em relação aos ativos de dados na organização. A qualidade se refere à capacidade de servir apropriadamente os propósitos dos dados e das informações em um contexto específico, determinando um domínio no modelo de governança (FU et al., 2011). Assim, cabe ao modelo de governança definir uma estrutura identificando quem na organização possui o direito de tomar decisões quanto aos padrões de qualidade dos dados, quais aspectos de qualidade

devem ser considerados e como assegurar que tais padrões sejam seguidos na organização (FU et al., 2011).

A gestão da qualidade dos dados foca no planejamento, no uso e no descarte dos dados. Ela aborda tanto questões técnicas quanto organizacionais, referindo-se à prontidão ou à compatibilidade dos dados em servir os propósitos de um dado contexto (WEBER et al., 2009; FU et al., 2011). Otto et al. (2011) indicam que a qualidade dos dados corresponde a um campo explorado por pesquisas empíricas, experiências práticas ou são resultantes de conclusões teóricas. O consenso sobre o termo qualidade dos dados corresponde à compatibilidade do dado ao uso, cuja avaliação depende do contexto no qual o dado é usado.

No âmbito do domínio qualidade, Khatri e Brown (2010) colocam como questões relevantes para serem tratadas:

- Quais são os padrões para a qualidade dos dados com relação à precisão, atualização, exaustividade e credibilidade?
- Qual é o programa para estabelecer e comunicar a qualidade dos dados?
- Como a qualidade dos dados e os programas serão avaliados?

As dimensões sugeridas pela literatura sobre o domínio relativo à qualidade são:

- Acurácia dos dados (*data accuracy*): atrelada à corretude e consistência dos dados (FU et al., 2011);
- Completude dos dados (*data completeness*): assegura que os dados requeridos são registrados sem a perda de valor (FU et al., 2011);
- Integridade dos dados (*data integrity*): remete-se à coerência, relevância, atualidade e veracidade dos dados. Este item adquiriu mais relevância pois grande parte dos conteúdos provém de fontes externas. Dados públicos ou das mídias sociais não necessariamente são precisos e confiáveis, especialmente quando múltiplas fontes de informação são combinadas (IBM, 2014).

Assegurar a qualidade dos dados é fundamental para as atividades analíticas organizacionais, pois ela determina no valor dos ativos de dados e influencia no tempo para preparo dos dados, que podem incluir ações de correções destes ativos (BEKOOZ, 2017). A baixa qualidade dos ativos dos dados influencia no tempo

necessário para preparo dos dados em projetos de BI, onde Berkooz menciona que 80% do tempo de projeto é alocado somente nesta atividade. Além disto, a baixa qualidade dos dados pode atrasar e encarecer os projetos de TI e acarretar impacto negativo no lucro das corporações (FU et al., 2011; BEKOOZ, 2017). No entanto, Otto (2011) enfatiza a necessidade de mais estudos para trazer constatações práticas de como implementar a governança e garantir a maximização da qualidade dos dados (OTTO, 2011).

É por meio do domínio "qualidade dos dados" que se pode diminuir os custos de base instalada, focalizar em conteúdos de valor, maximizar as análises, garantir a confiabilidade para as tomadas de decisões. Segundo Bekooz (2017), o modo prático para controlar a qualidade dos dados é mensurá-la no momento da criação do dado, pois sua verificação posterior é custosa e bastante trabalhosa. No entanto, isto requer a mudança no comportamento daqueles que criam os dados, assim como quais ferramentas de TI são utilizadas.

Como definem Khatri e Brown (2010), Fu et al (2011), o domínio qualidade, por sua vez, é pré-requisito para definir o domínio meta-dados, o qual se refere em como os dados podem ser definidos e interpretados. Fu et al. (2011) tratam os meta-dados como aqueles que descrevem dados. Os meta-dados descrevem as unidades essenciais do negócio de uma organização, incluindo os fornecedores, os clientes, os produtos, os empregados e os ativos (OTTO et al, 2011).

Como mencionam Otto et al. (2011), a qualidade dos meta-dados se tornou uma questão fundamental para as organizações, pois é através deles que se pode criar uma fonte única de informação para gerar relatórios completos e que possam cobrir o universo completo da organização ou de uma unidade de negócios. Além disto, os autores indicam que há crescente preocupação em atender as questões regulatórias e jurídicas, cujos meta-dados contribuem para tal atribuição.

Assegurar a qualidade dos meta-dados os torna atualizados, consistentes e completos. Otto et al. (2011) indicam que tal atividade pode ser alcançada através da adoção das estruturas de governança de dados ou do gerenciamento de meta-dados. Segundo Khatri e Brown (2010), são questões pertinentes ao domínio meta-dados no modelo da governança de dados:

- Qual é o programa para documentar a semântica dos dados?

- Como os dados podem ser consistentemente definidos e modelados para serem interpretados?
- Qual é o plano para manter diferentes tipos de meta-dados atualizados?

A Figura 15 demarca a influência dos domínios no modelo de governança de dados e de informação. O acesso aos dados e a sua autorização constituem em um outro domínio do modelo de governança. Corresponde ao controle de acesso dos usuários a dados privados ou sensíveis com base nos privilégios de usuários. Para Khatri e Brown (2010), este domínio almeja responder:

- Qual é o valor de negócio do dado?
- Como será conduzida a análise de risco em uma base contínua?
- Como os resultados da avaliação serão integrados com os esforços de monitoramento de *compliance*?
- Como a consciência da segurança e a educação podem ser disseminadas?

Figura 15 - Adaptação do modelo de Khatri e Brown sobre os domínios das governanças de dados e de informação. Fonte: Khatri e Brown (2010)

Por fim, o domínio "ciclo de vida do dado ou da informação" corresponde a base do modelo, fazendo referência à definição, produção, retenção e descarte dos dados (KHATRI; BROWN, 2010; TALLON et al. 2014). Para Khatri e Brown (2010), este domínio deve contemplar os procedimentos e as práticas para mensurar o valor dos ativos de dados e de informação ao longo do tempo. Assim, cabe ao domínio definir quais dados devem ser capturados, por quais motivos, por quem e por quanto tempo retê-los nos centros corporativos (TALLON, 2013).

O ciclo de vida descreve as fases de crescimento e de decréscimo do valor do dado ou da informação no tempo, culminando com o momento de descarte quando deixa de ser relevante para a corporação. No entanto, a informação contém uma série de características que são únicas e que dificultam sua avaliação no tempo (KOOOPER et al., 2011).

O valor dos dados pode ter súbito aumento por razões imprevisíveis e distintas do ambiente competitivo. Nisto se aplicam os casos em que os dados ou os arquivos são armazenados para regulamentar conformidades legais das corporações, sendo também uma função do *big data* (ADMA, 2013). Tallon (2013) cita um caso de uma instituição financeira autuada no valor de US\$ 1,5 bilhões, em 2005, pelo fato de não possuir algumas informações específicas que removeria tal penalização.

Os dados e as informações devem ser tratados como ativos organizacionais, cujo valor deve ser incluído no valor financeiro da empresa, por exemplo (OTTO, 2011; BARATA; PRADO, 2015). A omissão deste valor no balanço financeiro é como se os mesmos não possuíssem valor, o que é incorreto (TALLON, 2013). No entanto, Leming (2015) ressalta que é incomum tratar os dados e as informações no âmbito do gerenciamento de ativos da organização (LEMING, 2015).

Em continuidade aos modelos de governança, Fu et al. (2011) propõem o domínio "disponibilidade dos dados" ao invés do domínio "ciclo de vida dos ativos de dados ou das informações". Segundo os autores, este domínio deve contemplar quais dados a organização possui à disposição e por quais formas podem ser acessados pelos usuários. Este domínio, no entanto, não focaliza a necessidade de constituir patamares de valor dos dados e das informações ao longo do tempo, como propõem o ciclo de vida dos dados/informações de Khatri e Brown (2010).

Figura 16 - Domínios do modelo de governança de TI e do modelo de governança de dados ou de informação. Fonte: Khatri e Brown (2010)

Khatri e Brown (2010) também traçam um contraponto entre os modelos de governança de TI com o de governança de dados ou de informação. Na Figura 16 é possível visualizar os esquemas propostos, que não necessariamente refletem o mesmo objeto entre os modelos previstos.

O funcionamento dos modelos de governança requer a estipulação dos princípios que regem os demais domínios. Como a governança de TI trata de artefatos físicos da tecnologia, estes são englobados dentre os domínios de infra-estrutura e de arquitetura da TI. Os domínios de aplicação de negócios, de investimentos em TI e priorização contemplam aspectos relativos aos benefícios, aos retornos financeiros e aos investimentos que devem nortear as aplicações em TI. Em contrapartida, o modelo para as governanças de dados e de informação constitui-se para tratar ativos não-físicos ou artefatos de informação. Isto justifica os domínios peculiares desta governança, como o ciclo de vida do dado ou da informação; meta-dados; acesso aos dados e qualidade dos dados. Os custos e as questões financeiras são particularmente tratados no domínio relativo ao ciclo de vida do dado ou da informação.

Na sessão seguinte serão apresentados os elementos e conteúdos referentes às práticas e sub-práticas das governanças de dados e de informação.

2.3.2 Práticas e sub-práticas

As práticas e as sub-práticas do modelo de governança de dados ou de informação são as formas pelas quais a organização executa os princípios da governança (TALLON, 2013). Assim, elas devem contemplar políticas e procedimentos que descrevem como os dados ou as informações devem ser gerenciados ao longo do ciclo de vida econômico (TALLON, 2013). Para tanto, esta pesquisa utiliza como base o modelo de Khatri e Brown (2010) pois este contempla o domínio ciclo de vida do dado ou da informação.

Assim como na literatura de governança de TI, as práticas da governança de dados ou de informação podem ser classificadas em estruturais, operacionais e relacionais. Elas diferem, no entanto, por focalizarem na gestão e na proteção dos

artefatos de informação ao invés de serem aplicadas sobre os artefatos de TI (TALLON et al., 2014). Demais características incluem:

- **Estruturais:** práticas que identificam os principais tomadores de decisão, que podem ser da TI ou não, e suas respectivas funções e responsabilidades quanto à posse dos dados, análise de valor e gerenciamento dos custos. Essas práticas também podem especificar os comitês que competem ao *compliance*, com políticas internas ou regras legais sobre a retenção dos dados. Exemplo: definir a função de um conselho corporativo para definir políticas e padrões no uso e na proteção dos dados (TALLON, 2013);
- **Procedurais ou operacionais:** correspondem as formas pelas quais as organizações executam a governança de dados ou de informação, contemplando práticas aplicadas em todos os estágios do ciclo de vida da informação (TALLON, 2013; TALLON et al., 2014). Estas englobam uma série de ações, podendo ser: migração dos dados, retenção dos dados, *e-discovery*, direitos de acesso, procedimentos de *back-up* ou de recuperação, armazenamento estratificado, dentre outras opções (TALLON, 2013);
- **Relacionais:** fomentam a educação, a comunicação e os relacionamentos na organização. Contemplam práticas que descrevem os elos formais entre CIO, gerentes de negócios e usuários em termos do compartilhamento de conhecimento, análise do valor, educação, treinamento e planejamento estratégico (TALLON, 2013).

O Quadro 8 propõem um resumo das práticas das governanças de dados e de informação, segundo Tallon et al. (2014).

Quadro 8 - Práticas da governanças de dados ou de informação.

Práticas	Descrição
Estruturais	Práticas relacionadas à definição do local de tomada das decisões em TI ou da posse de dados. Designam funções para os principais tomadores de decisão; englobam práticas de reporte à estrutura de TI; uso de comitês ou outras estruturas para definir políticas ou grupos de monitoramento.
Relacionais	Edução; compartilhamento de conhecimento; resolução de conflitos.
Operacionais ou Procedurais	Práticas que refletem o amadurecimento da informação em função de critérios; envolvem a valoração da informação; acesso de usuários; retenção; arquivamento; destruição; gerenciamento de custos; migração de dados entre camadas de armazenamento; recuperação de custos; práticas de <i>e-discovery</i> .

Fonte: Tallon et al. (2014)

Quadro 9 – Elementos e conteúdos dos modelos de governança de dados ou de informação.

Práticas	Elementos e conteúdos governanças de dados e de informação
Estruturais	1. Criar direitos e responsabilidades de " <i>data ownership</i> " (donos de dados) e " <i>data stewardship</i> " (administradores de dados) (MALAKAR, 2007; ROSENBAUM, 2010; TALLON, 2013; TALLON et al., 2014; THOMPSON et al., 2015; KOLTAY, 2016);
	2. Estabelecer comitês de direção ou outros meios para avaliar o valor dos dados e seu custo (RAU, 2004; WEILL; ROSS, 2005; MALAKAR, 2007; WEBER et al., 2009; KHATRI, BROWN, 2010; OTTO, 2011; TALLON, 2013; TALLON et al., 2014);
	3. Envolver o usuário na formulação de políticas e avaliação (TALLON, 2013);
Relacionais	4. Contemplar cultura baseada em evidências ou <i>data-driven</i> (TALLON et al., 2014);
	5. Promover a colaboração interna para desenhar a governança de informação (TALLON et al., 2013);
	6. Promover a supervisão compartilhada para a formulação de políticas e avaliação (TALLON et al., 2014);
	7. Desenvolver a comunicação com relação à eficácia das políticas e necessidades de usuários (TALLON, 2013; TALLON et al., 2014)
	8. Educar os usuários e gerentes de outras áreas (diferentes da TI) com relação ao armazenamento apropriado de dados e de informações, e seus custos envolvidos (TALLON, 2013; TALLON et al., 2013; TALLON et al., 2014);
	9. Instituir processos de gestão de mudanças (KOLTAY, 2016)
Operacionais	10. Eliminar práticas que governam em demasia as informações (TALLON et al., 2013)
	11. Permitir exceções nas políticas globais para atender as necessidades de políticas locais (TALLON et al., 2013)
	12. Contribuir para que usuários coloquem um valor financeiro nos dados (TALLON et al., 2013)
	13. Estabelecer procedimentos de cobrança retroativa para recuperar os custos operacionais e comunicar aos donos dos dados e usuários (TALLON, 2013; TALLON et al., 2014)
	14. Planejar e disponibilizar capacidade de armazenamento (TALLON, 2013)
	15. Avaliar segundo o ciclo de vida da informação (KHATRI, BROWN, 2010)
	16. Decretar políticas de retenção de dados, de possível auto-destruição ou arquivamento dos dados (TALLON, 2013; TALLON et al., 2014)
	17. Assegurar a qualidade dos ativos de dados; contemplar atividades de gerenciamento da qualidade dos dados (WEBER et al., 2009; KHATRI, BROWN, 2010; FU et al., 2011; KOOPER et al., 2011; OTTO, 2011; OTTO et al., 2011; TALLON et al., 2014)
	18. Estabelecer e gerenciar meta-dados (KHATRI, BROWN, 2010; FU et al., 2011; OTTO et al., 2011; KOLTAY, 2016)
	19. Classificar a informação de acordo com seu valor e alocar em infraestruturas estratificadas de armazenamento (TALLON et al., 2014)
	20. Migrar os dados e as informações entre as camadas de armazenamento com base em critérios definidos (TALLON, 2013; TALLON et al., 2014)
	21. Estabelecer e monitorar os direitos de acessos (KHATRI, BROWN, 2010; TALLON, 2013; TALLON et al., 2014)

Fonte: elaborado pela autora

O Quadro 9 lista as contribuições de Rau (2004), Weill e Ross (2005), Malakar (2007), Weber et al. (2009), Rosenbaum (2010), Khatri e Brown (2010), Fu et al. (2011), Otto (2011), Otto et al. (2011), Tallon (2013), Tallon et al. (2013), Tallon et al. (2014), Thompson et al. (2015) e Koltay (2016) com relação aos elementos e conteúdos que compõem a governança de dados ou de informação. As

organizações podem combinar diferentes práticas em função de suas necessidades e circunstâncias (TALLON, 2013).

A governança de dados ou de informação contempla esforços e atividades dentre múltiplos níveis hierárquicos, tanto *top-down* como *cross* na organização (THOMPSON et al., 2015). A Figura 17 ilustra a estrutura prevista pela literatura para a governança de dados ou de informação, com base em Malakar (2007), Marcan (2008), Otto (2011), Tallon (2013) e Thompson et al. (2015).

Figura 17 – Stakeholders da estrutura de governança de dados e de informação. Fonte: elaborado pela autora.

No alto escalão executivo, a literatura propõe como opções: o *Chief Financial Officer* (CFO), o *Chief Executive Officer* (CEO), o *Chief Technology Officer* (CTO) ou o *Chief Operations Officer* (COO) (MALAKAR, 2007; MARCAN, 2008). Estes são responsáveis pela direção estratégica da organização e pelo suporte financeiro (TALLON, 2013; TALLON et al., 2014; THOMPSON et al., 2015).

No entanto, Demarquet (2016) estabelece foco sobre o CFO. Segundo a referência, o CFO tem atuado na transformação do negócio, migrando de um papel com foco na contabilidade para o de um agente digital. No ambiente *big data*, o CFO torna-se

ainda mais estratégico e atua como um parceiro comercial para apoiar e executar programas de transformação chave dentro da organização.

O CFO também é convidado a oferecer as "melhores funções de finanças" para melhor atender às necessidades do negócio. Assim, Demarquet (2016) identifica responsabilidades como: redução de custos, identificação e eliminação de gargalos, promoção de maior agilidade nos processos internos, aperfeiçoamento das aplicações de tecnologia da informação financeira para serem mais eficientes.

Demais *stakeholders* da estrutura contemplam o *Chief Data Steward* (CDS), *Chief Data Officer* (CDO), *Global Head of Data*, *Chief Information Quality Officer*, *Chief Data and Analytics Officer* (CDAO) (MARCAN, 2008; BERKOOZ, 2017). Este nível corresponde a um cargo de gestão, sendo responsável por implementar, liderar, gerenciar, centralizar, controlar e coordenar as decisões sobre a governança de dados ou de informação na organização (MARCAN, 2008). O CDO tem assumido cada vez mais destaque nas corporações com as responsabilidades sobre a governança de dados ou de informação. Já o CDAO possui responsabilidades sobre dados e *analytics* (BERKOOZ, 2017).

O comitê de dados corresponde ao corpo central nas decisões da governança. É liderado pelo nível relativo ao CDO e deve assegurar a visibilidade do comitê na organização e o engajamento dos gerentes seniores nas causas defendidas (MALAKAR, 2007). Cabe ao comitê especificar os princípios no uso dos dados, conciliar os diferentes interesses e demandas entre os *stewards* (OTTO, 2011). Trata-se de um canal de comunicação entre os *stewards*, as unidades de negócios e a área de TI (WEBER et al., 2009).

Os *stewards* de negócios podem também ser chamados de *stewards* de dados ("*business data stewards*", em inglês, ou "*data owners*", donos de dados) são recrutados de um departamento ou de uma unidade de negócios relevante, cabendo à eles especificar os requerimentos de negócios e de qualidade relativos aos dados (OTTO, 2011; KHATRI; BROWN, 2010).

Os *stewards* técnicos focalizam nos aspectos técnicos relacionados aos dados e devem possuir competências de TI (THOMPSON et al., 2015). Além disto, fornecem as definições para a padronização dos dados, detalham as fontes, os sistemas e o fluxo dos dados entre sistemas (WEBER et al., 2009). Assim, majoritariamente estão alocados nos departamentos de TI (THOMPSON et al., 2015).

Por "*steward*" compreendem-se as funções de administradores e, por "*owner*", de posse. Cabe ao "*chief steward*" aplicar padrões e colocar em prática as decisões de qualquer conselho ou comitê de governança. Geralmente são gerentes seniores na organização (THOMPSON et al., 2015). Complementarmente, o termo "*data stewardship*" pode ser visto como um conjunto de métodos de gestão de dados, englobando os procedimentos de aquisição, armazenamento, agregação, identificação, de liberação nos acessos aos dados, dentre outros (ROSENBAUM, 2010).

Não se pode confundir, no entanto, este cargo com o papel do *Chief Information Officer*, CIO, que está relacionado aos ativos tecnológicos e contemplado na estrutura de governança de TI. Como mencionado anteriormente na sessão destinada à governança de TI, o CIO corresponde à entidade central para a governança de TI, sendo os chefes de TI. Este *stakeholder* possui perfil mais reativo, focalizando em operações de eficiência, consolidando as tecnologias, e atuando nas questões de diminuição de custos. Em contrapartida, o CDO deve possuir perfil pró-ativo e contemplar conhecimentos sobre tecnologia, gestão e possuir competências diplomáticas para lidar com conflitos, pois estes são impreterivelmente eminentes (MALAKAR, 2007).

Os desafios são significativos no que tange à implementação das governanças de dados e de informação. Em grande parte das organizações, elas restringem-se a porções discretas da informação na corporação, não abrangendo todo o conteúdo organizacional (DONALDSON; WALKER, 2004). Tallon et al. (2013) observam algumas lições aprendidas em um caso organizacional e propõem fatores relativos ao processo de adoção das governanças de dados e de informação. São eles:

- Eliminar práticas que governam em demasia as informações: A alta gestão da empresa acreditava que proteger os dados poderia minimizar os riscos e contribuir para o sucesso. No entanto, notaram que a governança em demasia (*over-governance*) contribuiu para maiores custos e riscos. A empresa modificou a abordagem na tentativa de melhor equilibrar a necessidade de proteção dos dados com a maior acessibilidade destes recursos para a tomada de decisão;
- Educar os usuários sobre os riscos e os custos relativos aos dados: A empresa desenvolveu programas de educação de usuários sobre a

necessidade da governança da informação ao enfatizar: a necessidade de cada indivíduo responsabilizar-se; ser pró-ativo; trabalhar com as políticas e não ao redor delas;

- Promover a colaboração interna para desenhar a governança de informação: As regras da governança de dados/informação na empresa foram co-criadas com representantes de negócios que se preocuparam em identificar o que é aceitável e apropriado para cada usuário em sua área de negócio;
- Permitir exceções nas políticas globais para atender as necessidades de políticas locais: É improvável que um único *framework* da governança de dados/informação seja capaz de atender às políticas de regulação de diferentes países;
- Contribuir para que usuários adicionem um valor financeiro nos dados: Os usuários na Intel foram designados para alocar um valor para seus dados. O valor foi utilizado para avaliar o risco financeiro de se perder o dado e para identificar as etapas necessárias para se reduzir os riscos a níveis mais toleráveis. A classificação também foi utilizada para assegurar que os dados fossem armazenados em camadas adequadas em função dos níveis de serviço para seus requerimentos de negócios e custos relativos.

Com vista aos conteúdos apresentados com relação à governança de TI, governança de dados ou de informação, foi possível elaborar o Quadro 10 com o resumo teórico apresentado.

Quadro 10 – Quadro teórico da revisão da literatura de governança de TI, governança de dados ou de informação.

Conteúdo	Descrição	Teoria
Fatores viabilizadores e inibidores	Contemplam fatores antecedentes que podem viabilizar ou inibir o processo de adoção das governanças de dados ou de informação com base no referencial teórico da governança de TI.	Sambamurthy e Zmud (1999); Kirsch (2004); Raghupathi (2007); Xue et al. (2008); Willson e Pollard (2009); Xue et al. (2011); Tallon et al. (2013); Tallon et al. (2014)
Modelo de governança de dados ou de informação	Definem o modelo de governança de dados ou de informação e discutem sobre os domínios do modelo de governança de dados ou de informação	Weber et al. (2009); Khatri e Brown (2010); Kooper et al. (2011); Fu et al. (2011); Otto (2011); Otto et al (2011); Haider e Haider (2013); Tallon (2013); Tallon et al. (2014); Berkooz (2017)
Práticas estruturais	Práticas estruturais da governança de TI, governança de dados ou de informação. Tais práticas definem a estrutura organizacional para governança, os princípios e os objetivos estratégicos.	Weill (2004); Weill e Ross (2005); Weill e Woodham (2002); Malakar (2007); Khatri e Brown (2010); Rosenbaum (2010); Fu et al. (2011); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014); Thompson et al., 2015; Demarquet (2016)
Práticas relacionais	Práticas relacionais da governança de dados ou de informação	Rau (2004); Weill e Ross (2005); Khatri e Brown (2010); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014); Koltay (2016)
Práticas operacionais	Práticas operacionais da governança de dados ou de informação	Khatri e Brown (2010); Kooper et al. (2011); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014)

Fonte: elaborado pela autora

3. Metodologia de pesquisa

As pesquisas científicas podem contemplar abordagens qualitativas ou quantitativas, a depender dos objetivos e das questões de pesquisa. Tais abordagens podem assumir paradigmas interpretativo ou positivista, cujas principais características são (ARNABOLDI, 2017):

- **Paradigma interpretativo:** foco no relativismo e na compreensão, compreendendo-se as pesquisas qualitativas. Em geral, almejam a compreensão mais profunda de um fenômeno, podendo lançar-se mão de estudos de casos, pesquisação, observação em geral do pesquisador;
- **Paradigma positivista:** este paradigma focaliza a generalização e a explicação causal, contemplando as pesquisas quantitativas. Podem ser testadas hipóteses com o objetivo de estabelecer, explicar ou correlacionar variáveis. Podem ser empregados *survey*, experimentos, observação estruturada.

A presente pesquisa trata do paradigma interpretativo, cuja questão de pesquisa almeja identificar como as organizações podem adotar as governanças de dados ou de informação no ambiente *big data*. Os métodos de pesquisa que utilizam deste paradigma abordam um fenômeno contemporâneo, onde se lança à campo para melhor compreender o fenômeno e agregar percepções do pesquisador na coleta de informações. A pesquisação corresponde a um método particular onde o pesquisador vivencia e ativamente participa da execução do fenômeno. Já estudos de casos, podendo ser únicos ou múltiplos, corresponde a um método de pesquisa onde o pesquisador interage com o fenômeno de forma passiva, coletando uma série de insumos para melhor estruturar e compreender fenômenos emergentes.

Os estudos de casos são empregados quando as fronteiras entre o fenômeno e o contexto não são claramente definidas. Eles permitem o acesso a múltiplas fontes de evidências (YIN, 1984; MIGUEL, 2007; ARNABOLDI, 2017) através de entrevistas com profissionais pré-selecionados, observação do pesquisador no campo de estudo, dentre outros insumos (EISENHARDT, 1989).

Os estudos de casos são utilizados para auxiliar na compreensão da dinâmica de um contexto, podendo ser múltiplos ou únicos (EISENHARDT, 1989). Para a

aplicação dos estudos de casos, necessariamente o fenômeno deve ser contemporâneo e estar inserido num contexto real de vida (YIN, 1984; MIGUEL, 2007; ARNABOLDI, 2017). A definição de quantos casos empregar na pesquisa está em linha às perguntas de pesquisa: a partir do momento que são atingidos os objetivos de pesquisa, não há mais necessidade de utilizar mais casos organizacionais.

Foram considerados múltiplos estudos de caso na presente pesquisa com vistas aos objetivos definidos. Necessariamente devem ser contempladas organizações com estrutura formal de governança de dados ou de informação, que contemplem ambiente *big data*, que apresentem diferentes capacidades analíticas segundo o modelo de LaValle et al (2011), que possuam alto conteúdo de informação no produto, alta intensidade de informações na cadeia de valor e cuja TI assuma papel estratégico na organização ou em transição.

Lançou-se mão de uma pesquisa piloto com o intuito de aprimorar o roteiro de pesquisa. Para a escolha das organizações, foram utilizados os mesmos requisitos que para os estudos de casos. Para a pesquisa piloto, foram consideradas duas organizações do setor financeiro (empresas 1 e 2).

Tanto a pesquisa piloto quanto os estudos de casos foram conduzidas dentre os anos 2017 e 2018. Foram contempladas organizações globais dos setores financeiro, farmacêutico e de bens de consumo. Os estudos de casos foram conduzidos presencialmente dentre os escritórios das organizações em suas sedes corporativas ou subsidiárias, nas seguintes localidades: Milão (Itália), Arese (Itália), Turim (Itália), Zug (Suíça), Basileia (Suíça), São Paulo (Brasil).

As etapas metodológicas desta pesquisa foram elaboradas com base nos trabalhos de Miguel (2007) e Yin (2009) e ilustradas pela Figura 1. Como apresentado anteriormente, a questão geral desta pesquisa foi definida: *“como as organizações podem adotar as governanças de dados ou de informação no ambiente big data?”*

Para responder à tal questionamento, foram definidas duas perguntas como desdobramentos desta questão geral: (1) De que modo a capacidade analítica da organização pode ser um fator determinante para a adoção das governanças de dados ou de informação? (2) Quais são os elementos e conteúdos previstos pela literatura para as governanças de dados ou de informação? Como eles determinam, se isto for possível, na adoção das governanças de dados ou de informação? Tais perguntas foram retratadas pela Figura 18, onde é possível também notar as

proposições de pesquisa definidas ao término da revisão de literatura e que auxiliam na pesquisa em campo.

As proposições de pesquisa consistem em fundamentos teóricos que almejam direcionar a pesquisa em campo, podendo ser validadas ou refutadas ao término da pesquisa. A retórica que embasa cada proposição é necessária para garantir o maior alinhamento entre as perguntas de pesquisa e a pesquisa em campo a ser realizada.

Figura 18 – Esquema metodológico com as proposições de pesquisa. Fonte: elaborado pela autora.

As três proposições de pesquisa desenvolvidas com base na revisão de literatura de inteligência competitiva, governança de TI, governança de dados e governança de informação são:

Proposição 1: Capacidades analíticas transformadas correspondem a um fator determinante para a adoção das governanças de dados e de informação (SAMBAMURTHY; ZMUD, 1999; LAVALLE et al., 2011; ROSS et al., 2013; TALLON et al., 2014; SUN et al. 2015a, LISMONT et al., 2017)

As características com relação às organizações com capacidades analíticas transformadas foram previstas por LaValle et al. (2011). Neste quesito, tratam-se de organizações com capacidades analíticas mais desenvolvidas.

LaValle et al. (2011), Ross et al. (2013) e Sun et al. (2015a) discutem que organizações com capacidades analíticas mais desenvolvidas possuem também maiores habilidades para extrair valor de *big data*. Ora, organizações com capacidade analítica pouco ou médio desenvolvidas (nível aspiracional ou experiente do modelo de LaValle et al. (2011)), que possuem menor conhecimento sobre a manipulação e a extração de valor de dados, conseguem adotar estruturas de governança de dados ou de informação? A discussão proposta pela proposição 1 é relevante para melhor compreender o fenômeno relativo à governança de dados ou de informação nas organizações e sua possível relação com as capacidades analíticas.

A proposição 1 permite investigar se capacidades analíticas transformadas da organização correspondem a um fator determinante para a adoção das governanças de dados ou de informação. Para tanto, necessariamente devem ser contempladas organizações com diferentes capacidades analíticas para análise comparativa.

Proposição 2: As ações organizacionais para governar os dados ou as informações no contexto do *big data* ainda são pouco consolidadas e conhecidas no ambiente organizacional (LAVALLE et al., 2011; DAVENPORT; BARTH; BEAN, 2012; BARTON; COURT, 2012; TALLON et al., 2014; KORONIOS et al., 2014; RANSBOTHAM et al., 2015)

Com base na revisão teórica, foram evidenciadas diversas variáveis de pesquisa para a extração de valor dos dados no ambiente *big data*, além de diferentes parâmetros para caracterizar os modelos de governança de dados ou de informação, incluindo sub-práticas estruturais, relacionais e operacionais.

A segunda proposição sugere baixa consolidação das ações organizacionais com relação às práticas, sub-práticas e dos procedimentos relativos às governanças de dados e de informação no ambiente *big data*. Além disto, ela sugere que as ações para governar dados ou as informações no ambiente *big data* também são pouco conhecidas pelas organizações.

Assim, esta proposição de pesquisa impõem a necessidade de considerar múltiplos casos na pesquisa para observar a convergência ou não dos elementos e conteúdos das governanças de dados ou de informação no ambiente *big data*.

Proposição 3: Apesar das particularidades organizacionais, existem práticas, sub-práticas e fatores comuns que frequentemente são adotados para viabilizar a adoção das governanças de dados ou de informação (WEILL; WOODHAM, 2002; KHATRI; BROWN, 2010, TALLON, 2013; TALLON et al., 2014)

Já a terceira proposição de pesquisa sugere a existência de elementos e fatores comuns dentre as organizações que formalmente adotaram as estruturas de governança de dados ou de informação no ambiente *big data*. Caso validada esta proposição, podem ser sugeridos os fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*.

A consolidação teórica é fundamental para garantir a coleta sistemática e efetiva das informações na pesquisa em campo (EISENHARDT, 1989). Para tanto, foram consideradas as literaturas de inteligência competitiva e governança, cujas buscas por artigos científicos foram realizadas nas plataformas científicas indexadas ao portal de periódicos da CAPES (ISI *Web of Science* e *Scopus*) no período de 2014 a 2018. As principais palavras-chave empregadas nas buscas por artigos foram: "*business intelligence*"; "*business analytics*"; "*big data*"; "*big data analytics*"; "*analytics*"; "*data governance*"; "*information governance*"; "*information technology governance*"; "*IT governance*". Múltiplas buscas foram realizadas nas bases de dados empregando de forma isolada ou combinada as palavras-chave mencionadas.

Figura 19 - Citação dos artigos de Khatri e Brown (2010), Tallon (2013) e Tallon et al. (2014) para as bases Scopus e ISI *Web of Science* (Janeiro de 2018). Fonte: Scopus; ISI *Web of Science*

No campo da governança de dados e de informação, os principais artigos encontrados foram os de Khatri e Brown (2010), Tallon (2013) e Tallon et al. (2014). A Figura 19 ilustra as curvas de citação destes artigos para Janeiro de 2018. A partir dela, é possível notar rápido crescimento no número de citações na comunidade científica, influenciando demais trabalhos.

O modelo de Khatri e Brown (2010) tem sido adotado para representar a governança de dados, como no trabalho de Alhassan et al. (2016). Segundo a Scopus (para Janeiro de 2018), o artigo de Khatri e Brown (2010) foi citado por 112 documentos, os quais apresentam um total de 466 citações. Na ISI *Web of Science* (Janeiro de 2018), foram notados 52 documentos que referenciam Khatri e Brown (2010) e apresentam um total de 211 citações. O Quadro 11 lista os 15 artigos mais citados pela Scopus que possuem como base Khatri e Brown (2010).

Quadro 11– Citações dos artigos que citam Khatri e Brown (2010) pela Scopus (Janeiro 2018).

	Título	Autores referenciados nesta pesquisa	Periódico	Citações
1	<i>On the governance of information: Introducing a new concept ...</i>	Kooper et al. (2011)	<i>International Journal of Information Management</i>	59
2	<i>Corporate governance of big data: Perspectives on value, ris...</i>	Tallon (2013)	<i>Computer Journal of Management</i>	40
3	<i>The information artifact in IT governance: Toward a theory o...</i>	Tallon et al. (2013)	<i>Information Systems</i>	36
4	<i>DEMOS: A description model for data-as-a-service</i>			28
5	<i>Data governance in predictive toxicology: A review</i>	Fu et al. (2011)	<i>Journal of Cheminformatics</i>	20
6	<i>A survey of big data management: Taxonomy and state-of-the-a...</i>			19
7	<i>Organizing Data Governance: Findings from the telecommunicat...</i>	Otto (2011)	<i>Communications of the Association for Information Systems</i>	19
8	<i>Product data quality in supply chains: The case of beiersdor...</i>		<i>Electronic Markets</i>	18
9	<i>Sustaining accessibility of information through digital pres...</i>		<i>Journal of Information Science</i>	17
10	<i>Toward a functional reference model for master data quality ...</i>		<i>Information Systems and e-Business Management</i>	13

Fonte: elaborado pela autora

Com relação aos demais artigos, Tallon (2013) apresentou 59 citações dentre periódicos de alto impacto. E Tallon et al. (2014) apresentaram 53 citações dentre periódicos relevantes no campo da TI, como o *MIS Quarterly* (5.384), *Journal of Management Information Systems* (3.025), *European Journal of Information Systems* (2.895), para ambas as plataformas.

3.1 Mecanismos para a definição do campo

Eisenhardt (1989) aponta que o processo de seleção dos casos organizacionais para pesquisa é fundamental para o processo de construção de teoria. Os mecanismos para seleção de casos foram aplicados tanto para a pesquisa piloto quanto para os estudos de casos. Foram contempladas organizações dos setores financeiro, farmacêutico e de bens de consumo. As premissas utilizadas para a escolha dos casos encontram-se listadas pelo Quadro 12.

Quadro 12 – Premissas para escolha dos casos organizacionais.

Premissa	Descrição
1	<ul style="list-style-type: none"> As organizações devem possuir estruturas formais de governança de dados ou governança de informação para atender ao corte epistemológico desta pesquisa. Dentre os possíveis entrevistados compreendem <i>stakeholders</i> com as seguintes posições: <i>chief data officer</i>, diretor de meta-dados, gerente de meta-dados, diretor ou gerente de governança de dados, diretor ou gerente de governança de informação, <i>data steward</i>, gerente ou analista de qualidade dos dados;
2	<ul style="list-style-type: none"> As organizações devem contemplar o ambiente <i>big data</i> em seus negócios e possuir alguma competência analítica desenvolvida;
3	<ul style="list-style-type: none"> As organizações devem ser classificadas dentre os quadrantes transição ou estratégico com relação ao <i>Grid</i> Estratégico de McFarlan (1984). Nestas organizações, a TI assume posição estratégica para os negócios ou está orientada a tornar-se estratégica no breve espaço de tempo;
4	<ul style="list-style-type: none"> As organizações devem apresentar alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985);
5	<ul style="list-style-type: none"> As organizações que compõem a pesquisa devem apresentar diferentes níveis de capacidades analíticas.

Fonte: elaborado pela autora.

A primeira premissa para escolha de casos corresponde a existência formal da estrutura de governança de dados ou de informação. Para identificar possíveis organizações com tal estrutura, lançou-se mão da ferramenta LinkedInTM. Esta ferramenta foi utilizada por ser fonte (na maioria das vezes) de informações atualizadas e confiáveis de profissionais que atuam no ambiente corporativo a nível global. Nesta ferramenta foram conduzidas pesquisas por *stakeholders* de governança de dados ou de informação, onde foram empregados os seguintes termos nas buscas:

- *Chief data officer,*
- Diretor de governança de dados (*data governance director; head of data governance*),
- Diretor de governança de informação (*information governance director; head of information governance*),
- Diretor de meta-dados (*director of master data; head of master data*),
- Diretor de qualidade dos dados (*director of data quality; head of data quality*),
- Gerente de meta-dados (*master data manager*),
- Gerente de qualidade dos dados (*data quality manager*),
- Gerente de dados (*data manager*),
- Analista de meta dados (*master data analyst*),
- Analista de qualidade dos dados (*data quality analyst*);
- *Business data stewards, data stewards;*
- *Data owner.*

A existência de *stakeholders* para a governança de dados ou de informação é prevista pela referencial teórico e imprescindível para o corte epistemológico desta pesquisa. As considerações com relação aos cargos previstos foram previstas por Malakar (2007), Marcan (2008), Weber et al. (2009), Khatri e Brown (2010), Rosenbaum (2010), Otto (2011), Tallon et al. (2014) e Berkooz (2017).

A adoção das estruturas de governança de dados ou de informação corresponde a um fenômeno emergente nas organizações, que tem acontecido principalmente após os anos 2015. Em geral tais estruturas estão alocadas nos escritórios sede das organizações ou em subsidiárias com operações relevantes para o negócio.

Após o atendimento da primeira premissa, lançou-se mão para a identificação do ambiente *big data* dentre os negócios destas organizações.

A terceira premissa para escolha de organizações envolveu o *Grid* Estratégico de McFarlan (1984), o qual sugere o papel da TI para a organização. Nesta pesquisa, foram contempladas organizações dentre os quadrantes “estratégico ou em transição”. Nas organizações classificadas como “estratégicas” no *Grid* de McFarlan, a TI exerce significativa influência na estratégia geral da empresa. As aplicações atuais e futuras da TI são vistas como estratégicas para o negócio. A falha dos sistemas por um minuto ou mais acomete em perdas imediatas de negócios. Novos sistemas prometem grandes mudanças nos processos e nos serviços, além de reduzirem significativamente as desvantagens em relação aos competidores em termos de custos, serviços ou desempenho dos processos. Exemplo: Instituições financeiras (bancos, seguradoras, dentre outras), empresas de telecomunicações (LAURINDO, 2008).

Nas organizações classificadas em “transição” no *Grid* Estratégico de McFarlan (1984), a TI migra de uma situação mais discreta para uma de maior destaque na estratégia da empresa devido às novas aplicações de TI. Os novos sistemas prometem grandes mudanças nos processos e nos serviços e prometem reduções de custos; a TI constitui mais de 50% dos gastos com capita e mais que 15% das despesas da empresa (LAURINDO, 2008).

A quarta premissa para escolha das organizações utilizou a Matriz de Intensidade de Informação de Porter e Millar (1985). Em particular, foram objeto de investigação as organizações com alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor. Por último, a quinta premissa para escolha dos casos estabelece que organizações com diferentes capacidades analíticas devem ser empregadas na pesquisa para análise comparativa.

Após a identificação dos casos que atendiam as premissas, foram contatados os possíveis entrevistados através do correio eletrônico do próprio LinkedIn e através de *e-mail* corporativo. As pesquisas em campo foram conduzidas presencialmente dentre os anos 2017 e 2018, as quais foram registradas, transcritas, confirmadas e analisadas posteriormente.

Os casos contemplados nesta pesquisa compreendem organizações dos setores financeiro, farmacêutico e de bens de consumo. O setor farmacêutico é tecnologicamente sofisticado para o desenvolvimento de novos tratamentos e

medicamentos para o futuro da humanidade, sendo submetido a intensas pressões de regulação e de qualidade dos seus produtos.

Este setor é também bastante influente no desenvolvimento científico. Dentre as publicações científicas no âmbito das governanças de dados e de informação, o tema saúde acomete na maior parcela dos artigos. Segundo buscas nas plataformas *ISI Web of Science* e *Scopus* (em 2016), as áreas de pesquisa em: *health care sciences services; information science library science; medical informatics* acometem 60% das publicações. Kemp (2014) sugere que o setor de saúde tem adotado intensamente o *big data* e possivelmente será o setor que trará maior impacto na vida das pessoas.

As organizações do setor financeiro contam com o mais alto uso de TI no globo, além de estarem investindo rigorosamente em técnicas de *big data* para alcance de vantagens competitivas (KEMP, 2014). Movimentam-se em busca de melhores estruturas de governança e gestão de dados, com o acirramento de regulações e políticas globais de proteção dos dados, sigilo de informações pessoais. No caso europeu, a GDPR – *General Data Protection Regulation* – promulgada em 2016, tem levado principalmente organizações de saúde e do setor financeiro a aperfeiçoar sua gestão para atender aos novos requisitos de negócio.

Os bancos modernos devem responder aos inúmeros desafios, como a competição acirrada, as fusões e aquisições, o desenvolvimento de novos produtos e a segmentação de mercado. Os bancos registram alto volume de dados diariamente e com as mais diferentes características. Segundo Ubiparipovic e Durkovic (2011), tornou-se claro há um tempo que os bancos possuem grande quantidade de dados mas pouca informação. Possuem menos ainda conhecimento em muitos aspectos de suas operações. No entanto, os bancos de dados transacionais são enormes.

As organizações do setor de bens de consumo fornecem bens para o atendimento das necessidades da população em geral, do indivíduo ou das famílias. Este setor fornece bens com diferentes níveis de durabilidade e para o atendimento de distintas necessidades, podem-se compreender distintos produtos, como os alimentícios, automóveis, computadores, etc.

3.2 Variáveis de Pesquisa e Pesquisa Piloto

As variáveis de pesquisa correspondem aos insumos teóricos resultantes da revisão de literatura que são levados a campo para orientar a coleta de informações. Como apresentam os Quadros 13 e 14, foram identificadas 18 variáveis para o primeiro quadro e 39 variáveis para o segundo quadro. Com vistas ao alto volume de variáveis, em particular do Quadro 14, foi conduzida uma pesquisa piloto para validá-las e refinar o roteiro de pesquisa.

As variáveis apontadas pelo Quadro 13 remetem-se ao ambiente no qual há extração de valor dos dados, cujo resumo das vertentes teóricas foi listado pelo Quadro 5. Além disto, as variáveis para a avaliação das capacidades analíticas da organização foram listadas conforme modelo proposto por LaValle et al. (2011) ilustrado pela Figura 10.

Quadro 13 - Definição de construtos e variáveis de pesquisa sobre capacidade analítica e extração de valor no ambiente *big data*.

Construto	Variáveis	Autores
Extração de valor dos dados no ambiente <i>big data</i>	1. Estrutura para tomada de decisão; 2. Cultura <i>data-driven</i> ; 3. Comunicação e interação dos <i>stakeholders</i> ; 4. Definir os propósitos de negócio antes da coleta dos dados; 5. Utilizar os resultados analíticos nas decisões estratégicas do negócio; 6. Investimentos em TI e tecnologias para manipular <i>big data</i> ; 7. Alinhamento estratégico entre negócios e <i>analytics</i> ; 8. Utilizar conhecimento pré-existente da organização; 9. Resultados compreensíveis; 10. Competências tecnológicas e gerenciais;	Teo e Choo (2001); Chen et al. (2002); Davenport (2006); Cabral Netto (2007); Bose (2008); LaValle et al. (2011); McAfee e Brynjolfsson (2012); Davenport (2013); Ross et al. (2013); Koronios et al. (2014); Sun et al. (2015a); Sun et al. (2015b); Ransbotham et al. (2015); Lismont et al. (2017); Berkooz (2017)
Capacidade analítica da organização	11. Investimentos específicos em BI; 12. Experiência organizacional com projetos de BI; 13. Competências desenvolvidas em termos de pessoas, processos e ferramentas; 14. Uso da BI para obter vantagens competitivas; 15. Resultados da BI para a organização; 16. Interesse em expandir o uso da BI; 17. BI focaliza a experiência do cliente; 18. BI focaliza a automação de processos existentes.	LaValle et al. (2011), McAfee e Brynjolfsson (2012); Ross et al. (2013); Sun et al. (2015a); Lismont et al. (2017)

Fonte: elaborado pela autora

Quadro 14 – Definição de construtos e variáveis de pesquisa.

Construto	Variáveis	Autores
Fatores viabilizadores ou inibidores	1. Estratégia organizacional; 2. Estratégia de TI; 3. Estrutura organizacional (centralizada ou descentralizada); 4. Estrutura da TI; 5. Complexidade do portfolio de produtos e serviços; 6. Arquitetura de TI; 7. Cultura da TI; 8. Regulação setorial; 9. Taxa no aumento dos dados e das informações; 10. Abrangência da estrutura de governança de dados ou de informação na organização;	Sambamurthy e Zmud (1999); Kirsch (2004); Raghupathi (2007); Xue et al. (2008); Willson e Pollard (2009); Xue et al. (2011); Tallon et al. (2013); Tallon et al. (2014)
Modelo de governança de dados ou de informação	11. Princípio dos dados; 12. Qualidade dos dados; 13. Meta-dados; 14. Acesso aos dados; 15. Ciclo de vida dos dados; 16. Autorização dos dados; 17. Disponibilidade dos dados;	Weber et al. (2009); Khatri e Brown (2010); Kooper et al. (2011); Fu et al. (2011); Otto (2011); Otto et al. (2011); Haider e Haider (2013); Tallon (2013); Tallon et al. (2014); Berkooz (2017)
Práticas estruturais	18. Estrutura organizacional (<i>Chief Data Officer, data owner, data steward</i> , comitê de dados); 19. Envolvimento e suporte da alta gestão;	Weill (2004); Weill e Ross (2005); Weill e Woodham (2002); Malakar (2007); Khatri e Brown (2010); Rosenbaum (2010); Fu et al. (2011); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014); Thompson et al., 2015; Demarquet (2016)
Práticas relacionais	20. Cultura <i>data-driven</i> ; 21. Engajamento dos <i>stakeholders</i> ; 22. Comunicação e colaboração; 23. Supervisão compartilhada; 24. Educação e treinamento; 25. Gestão de mudanças;	Rau (2004); Weill e Ross (2005); Khatri e Brown (2010); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014); Koltay (2016)
Práticas operacionais	26. Burocracia; 27. Exceções globais; 28. Ciclo de vida das informações; 29. Direito de acesso; 30. Gerenciamento de meta-dados; 31. Assegurar a qualidade dos dados e das informações; 32. Medição do desempenho; 33. Otimizar custos; 34. Padronizar o fluxo dos dados e das informações; 35. Procedimentos de cobrança retroativa; 36. Capacidade de armazenamento; 37. Políticas de retenção de dados, auto-destruição, arquivamento de dados; 38. Classificar a informação de acordo com seu valor e alocar em infraestruturas estratificadas de armazenamento; 39. Migrar os dados entre as camadas de armazenamento.	Khatri e Brown (2010); Kooper et al. (2011); Nfuka e Rusu (2011); Tallon (2013); Kurti et al. (2014); Tallon et al. (2014)

Fonte: elaborado pela autora

A pesquisa piloto foi desenvolvida em duas organizações do setor financeiro que atendiam integralmente as premissas para escolha de casos definida nesta pesquisa. Na ocasião da pesquisa piloto, as organizações foram denominadas por empresa 1 e 2. O Quadro 15 resume as características destas empresas.

Quadro 15 – Empresas consultadas na pesquisa piloto.

Empresa	Características
1	Banco de investimentos norte-americano com 225 anos de história. Responsável por mais de 10% dos ativos no mundo. Presente em mais de 100 países. Ajuda seus clientes nos maiores desafios econômicos, onde focar para crescimento, como usar seus dados, quais funções terceirizar. Seus clientes são gerentes de ativos, donos de ativos, seguradoras e instituições oficiais. Por competências, a organização por ser dividida em: gerenciamento de ativos, pesquisa de investimentos e <i>trading</i> , serviços de investimento, dados e <i>analytics</i> .
2	Instituição financeira de um conglomerado norte-americano multinacional. A empresa fornece empréstimos comerciais e operações de <i>leasing</i> , além de suportar diversas outras operações financeiras do conglomerado. Em 2014, a empresa operava com 35.000 funcionários em mais de 40 países, com total de ativos avaliado em US\$499 bilhões.

Fonte: elaborado pela autora

As entrevistas foram conduzidas com base nas variáveis apontadas pelo Quadro 14.

A pesquisa piloto foi realizada no início do ano de 2017.

3.2.1 Empresa 1

A empresa 1 empregada na pesquisa piloto corresponde a um banco de investimentos norte-americano com 225 anos de história. A organização é responsável por mais de 10% dos ativos no mundo e está presente em mais de 100 países. A corporação possui cerca de 35 mil funcionários no mundo e faturamento de mais de US\$ 10 bilhões no ano de 2016.

A empresa 1 ajuda seus clientes nos maiores desafios econômicos. Ela está estruturada em três divisões principais, oferecendo serviços globais de gestão de ativos; serviços de consultoria de investimento financeiro e serviços de pesquisa de investimentos para investidores. Seus clientes são gerentes de ativos, donos de ativos, seguradoras e instituições oficiais.

Trata-se de uma organização voltada aos serviços financeiros, com portfólio complexo de serviços, podendo ser classificada no quadrante estratégico do *Grid* de MacFarlan. A TI assume perspectiva estratégica para a organização. Além disto, a organização contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985). A organização se considera *data-driven*, com capacidades analíticas transformadas, segundo o modelo de LaValle et al. (2011). O comentário abaixo de um dos profissionais consultados na pesquisa traz à tona a relevância dos dados para a organização consultada.

“Nos tornamos fornecedores de dados e consumidores de dados. Possuímos informações valiosas sobre grande parte da riqueza no mundo”, comenta o data governance officer da empresa 1.

Na pesquisa piloto, foram consultados o líder de governança de dados e o gerente de governança de dados, ambos situado no escritório de Nova Iorque, Estados Unidos. A estrutura de governança de dados da empresa 1 focaliza no gerenciamento e na promoção de altos índices de qualidade nos dados, pois a organização já se considera *data-driven*. A equipe de governança de dados consultada responde ao *Chief Data Officer* da organização e assim cabe a ela mapear os fluxos dos dados entre as aplicações, trazendo à luz a origem dos dados e seu destino. Em alguns casos, tais mapeamentos são requisitados pelos reguladores dos serviços financeiros.

Figura 20 – Estrutura organizacional da empresa 1. Fonte: elaborado pela autora

Como pode ser observado pela Figura 20, o *Chief Data Officer* da empresa possui quatro “*Data Governance Officers*”, um para cada unidade de negócios do banco. No nível do CDO são definidos os padrões, as políticas, ferramentas e plataformas para toda a organização. Já o trabalho operacional é conduzido pelo “*data governance officers*”, que contemplam os trabalhos de gerenciamento de dados através da catalogação. No caso desta organização, os “*data governance officers*” se consideram *data stewards* por estarem nas unidades de negócios.

A estrutura de governança de dados está sob responsabilidade do *Chief Operation Officer* (COO) que também acumula a função de *Chief Risk Officer* (CRO) da organização.

Complementarmente às atividades da estrutura de governança, o líder indicou que também são conduzidos múltiplos e pequenos projetos para positivamente impactar os negócios e ajudá-los. Assim, as unidades de negócios não os reconhecem como fomentadores de burocracia. Como mencionado pelo líder de governança de dados:

“Algumas das perguntas dos líderes de negócios são: quem está usando meus dados? Para onde vão os dados gerados pelos meus negócios? Qual aplicação utiliza meus dados? Existe muita preocupação das equipes de negócios com relação a segurança e a qualidade dos dados. Eles querem ter certeza de que estão utilizando dados de boa qualidade”.

Com vistas às questões regulatórias, foi mencionada a influência destes aspectos para as atividades de governança de dados da organização. O comentário abaixo de um líder de governança de dados na empresa 1 evidencia o papel da regulação setorial:

“Nós somos parte do grupo de regulatório. Este era o nome antigo para “governança de dados”. Então existe uma forte visão da governança de dados como uma função de regulatório no banco”, cita o “data governance officer” da empresa 1.

Aspectos operacionais com maior nível de granularidade não foram evidenciados na pesquisa piloto. Para tais atividades, são contemplados agentes do departamento de TI para auxiliar em atividades estritamente operacionais com relação aos dados.

3.2.2 Empresa 2

A segunda empresa consultada para a pesquisa piloto corresponde a uma instituição financeira de um conglomerado norte-americano que produz bens de capital. A empresa fornece empréstimos comerciais e operações de *leasing*, além de suportar diversas outras operações financeiras do conglomerado. Em 2014, a empresa operava com 35.000 funcionários em mais de 40 países, com total de ativos avaliado em US\$499 bilhões.

Trata-se de uma organização voltada aos serviços financeiros que suportam o setor de aviação, energia e industrial. Apresenta portfólio de serviços complexo, podendo ser classificada no quadrante estratégico do *Grid* de MacFarlan. Neste caso, a TI assume perspectiva estratégica para a organização. Além disto, a organização contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

A organização contempla capacidades analíticas transformadas segundo o modelo de LaValle et al. (2011). Ela almeja expandir o uso analítico e propiciar a constituição de ferramentas da *advanced analytics*. Para tanto, a estrutura de governança de dados foi adotada com o intuito de melhor organizar os ativos de dados e propiciar a constituição de melhores indicadores e relatórios para a alta gestão.

A pesquisa piloto foi realizada no início de 2017, onde foram consultados o gerente de governança de dados e um analista de qualidade dos dados. A estrutura de governança de dados se localiza no escritório em Londres e atende as operações e os processos globais da organização.

Como pode ser visto pela Figura 21, a estrutura de governança de dados está sob responsabilidade de um *Chief Data Officer*, desvinculado das unidades de negócios e do departamento de TI. Tal CDO se reporta ao CFO da organização, o qual também contempla a TI sob sua responsabilidade. A equipe de governança de dados contempla analistas de qualidade de dados, sendo a atividade de controle e monitoramento da qualidade dos ativos de dados central no modelo de governança de dados em operação.

A estrutura de governança de dados focaliza no gerenciamento de dados para aprimorar a qualidade dos ativos de dados que são utilizados para as análises de riscos da instituição. São contempladas análise de risco de capital e de crédito. Além disto, a definição de meta-dados também correspondeu uma atividade central no modelo de governança de dados da organização.

Figura 21 – Estrutura organizacional para a empresa 2. Fonte: elaborado pela autora

Com vistas as atividades operacionais mais específicas, estas não foram reportadas dentre o modelo de governança da empresa 2. Foi mencionado apoio do departamento de TI para implementar os requerimentos de qualidade de dados definidos pela equipe de governança de dados. Além disto, os entrevistados reportaram a necessidade de transformação da cultura da organização no sentido *data-driven*. Eles retrataram a atuação pró-ativa para que mudança do comportamento das pessoas com relação aos dados.

Foi também observada com destaque a relevância das práticas estruturais do modelo de governança de dados dentre a pesquisa piloto 2. Estas incluem a definição da estrutura para a governança de dados, englobando CDO, gerente de governança de dados, analista de dados, comitê de dados e demais definições de responsabilidades com relação aos dados e as informações no âmbito organizacional.

Com vistas à pesquisa piloto realizada, as variáveis 35, 36, 37, 38 e 39 do Quadro 14 não apresentaram relevância. Em geral, são discutidos tópicos mais estratégicos

e gerenciais com os *stakeholders* da estrutura de governança de dados ou de informação. Demais agentes na organização, como aqueles do departamento de TI, contemplam atividades operacionais que suportam as atividades de gerenciamento de dados ou de informações.

Logo, o Quadro 16 lista as sub-práticas operacionais retiradas do roteiro de pesquisa dos estudos de casos.

Quadro 16 – Sub-práticas operacionais descartadas do roteiro de pesquisa.

Prática	Sub-práticas
Operacional	<ul style="list-style-type: none"> • Estabelecer procedimentos de cobrança retroativa para recuperar os custos operacionais e comunicar aos donos dos dados e usuários (TALLON, 2013; TALLON et al., 2014) • Planejar e disponibilizar capacidade de armazenamento (TALLON, 2013) • Decretar políticas de retenção de dados, de possível auto-destruição ou arquivamento dos dados (TALLON, 2013; TALLON et al., 2014) • Classificar a informação de acordo com seu valor e alocar em infraestruturas estratificadas de armazenamento (TALLON et al., 2014) • Migrar os dados e as informações entre as camadas de armazenamento com base em critérios definidos (TALLON, 2013; TALLON et al., 2014)

Fonte: elaborado pela autora.

3.3 Roteiro de pesquisa

O roteiro de pesquisa consiste no material base para orientar a coleta de informações a campo. Ele serve como instrumento para a obtenção das informações necessárias para responder às perguntas de pesquisa preliminarmente definidas. O roteiro de pesquisa foi definido com auxílio da pesquisa piloto realizada, onde foram eliminadas cinco sub-práticas estritamente operacionais sugeridas pela literatura.

A Figura 22 esquematiza o processo de coleta de informação a campo definido em três momentos. A primeira frente do roteiro de pesquisa foi embasada em perguntas generalistas para a compreensão do contexto organizacional, da estratégia e dos negócios, podendo-se lançar mão das informações disponibilizadas nas páginas eletrônicas da organização e nas redes sociais para respondê-las:

- Quais são os produtos ou serviços oferecidos pela organização? A organização possui estratégia de negócios focada em alguns produtos ou serviços ou em múltiplos, constituindo um portfólio complexo?

- Qual é o tamanho e abrangência da organização? Qual é o posicionamento da empresa em seu segmento?

Figura 22 – Frentes do roteiro de pesquisa. Fonte: elaborado pela autora.

A segunda frente do roteiro de pesquisa contemplou os aspectos relacionados aos elementos e conteúdos dos modelos de governança de dados ou de informação no ambiente *big data*, como constam as perguntas no Quadro 17.

Em particular, as questões que compõem o Quadro 17 foram embasadas nos trabalhos de Sambamurthy e Zmud (1999); Weill e Woodham (2002); Kirsch (2004); Rau (2004); Weill (2004); Weill e Ross (2005); Raghupathi (2007); Malakar (2007); Xue et al. (2008); Willson e Pollard (2009); Weber et al. (2009); Khatri e Brown (2010); Rosenbaum (2010); Kooper et al. (2011); Fu et al. (2011); Otto (2011); Otto et al. (2011); Xue et al. (2011); LaValle et al. (2011); Nfuka e Rusu (2011); McAfee e Brynjolfsson (2012); Ross et al. (2013); Tallon et al. (2013); Haider e Haider (2013); Tallon et al. (2014); Kurti et al. (2014); Thompson et al. (2015); Sun et al. (2015a); Koltay (2016); Demarquet (2016); Berkooz (2017); Lismont et al. (2017).

No terceiro momento da entrevista, foram aplicadas as questões com foco nas competências analíticas da organização e extração de valor dos dados no ambiente *big data*. O Quadro 18 lista tais perguntas com base nos trabalhos de LaValle et al. (2011), McAfee e Brynjolfsson (2012); Ross et al. (2013), Sun et al. (2015a) e Lismont et al. (2017).

Quadro 17 – Questões sobre os elementos e conteúdos da governança de dados ou de informação.

Segunda fase da entrevista
<ul style="list-style-type: none"> • Quanto tempo existe na organização a estrutura da governança de dados ou de informação? Quais <i>stakeholders</i> foram definidos na estrutura? • A estrutura é centralizada em algum líder ou possui múltiplas lideranças? • Qual é a abrangência da estrutura de governança de dados ou de informação na organização? • Quais são os objetivos da governança de dados ou de informação na organização? • A organização possui “cultura com base em evidências” ou <i>data-driven</i>? A organização trata os dados como “ativos de dados”? • Quais foram os principais fatores que motivaram a adoção da governança de dados ou de informação? • Os aspectos de regulação setorial foram decisivos na adoção da governança? O aumento no volume de dados motivou a adoção das governanças de dados ou de informação? • Qual é a relação com a TI? Qual é o posicionamento e a estratégia da TI com relação à governança de dados ou de informação? • Como é o ambiente organizacional quanto ao envolvimento e suporte da alta gestão, engajamento dos <i>stakeholders</i>, comunicação e colaboração quanto às práticas de governança de dados ou de informação? • Os usuários são envolvidos para a formulação de políticas e avaliação? A supervisão é compartilhada? • Os usuários são treinados sobre os riscos e custos relativos aos ativos de dados? Os usuários são solicitados para colocar um valor financeiro aos dados? • Como é o relacionamento entre os <i>stakeholders</i> da governança de dados ou de informação com demais <i>stakeholders</i> da organização? • São contempladas mudanças na organização? São necessárias habilidades de gestão de mudanças e gerenciamento de conflitos? • São conduzidas atividades de gerenciamento de meta-dados? Qual relevante elas são? • Os dados ou as informações são avaliadas ao longo de seus ciclos de vida? Como a organização define e acompanha o ciclo de vida da informação? • São feitas ações para eliminar práticas que governam em demasia? Quão burocráticos são os procedimentos da governança de dados ou de informação? • São permitidas exceções nas políticas globais para atender especificidades locais? • São estabelecidos e monitorados os direitos de acessos? Os custos e os riscos são monitorados com relação aos ativos de dados? • Quão relevante são os aspectos de qualidade dos dados para a governança de dados? O que é feito no campo do gerenciamento da qualidade?
Fonte: elaborado pela autora

Quadro 18 – Questões com foco em BI.

Terceira fase da entrevista
<ul style="list-style-type: none"> • Existem alguma estrutura específica na organização para os profissionais de BI? • Como é a comunicação interna entre profissionais de BI e o restante da organização? • Como o(a) Sr(a) caracteriza os níveis de capacidades analíticas da organização? • A organização possui experiência com projetos de BI? Há interesse em expandir o uso da BI na organização? • São obtidas vantagens competitivas através do uso da BI na organização? • Como é possível descrever as competências organizacionais da BI antes e depois da adoção da governança de dados ou de informação? • A adoção da governança de dados ou de informação contribui para as práticas de BI? • Pela visão do Sr(a), as competências organizacionais da BI viabilizaram a adoção da governança de dados ou de informação? • O(a) Sr(a) acredita que a experiência preliminar em BI contribuiu para melhor estabelecer os procedimentos das governanças de dados ou de informação?
Fonte: elaborado pela autora

3.4 Estudos de casos

A pesquisa em campo contemplou organizações dos setores farmacêutico, financeiro (bancos e seguradoras) e bens de consumo. As pesquisas foram realizadas dentre os anos 2017 e 2018 em diferentes localidades do mundo: Itália, Suíça e Brasil.

Dos casos estudados, três deles correspondem a organizações do setor de farmacêutico (casos A, B, C). Dentre as organizações consultadas, duas delas são indústria farmacêutica suíças situadas na Basiléia (local de condução das entrevistas) e estão no *top-3* das maiores indústrias do setor. A outra empresa estudada corresponde a uma organização norte-americana focada em biotecnologia, cujas entrevistas foram realizadas em Zug, na Suíça, em um dos mais relevantes escritórios da organização.

Outros quatro casos (D, E, F, G) são organizações do setor financeiro, compreendendo bancos múltiplos e uma seguradora. A organização D se trata de um dos maiores bancos múltiplos europeus, situado em Turim, na Itália (local das entrevistas); o caso E corresponde a uma das maiores seguradoras no mundo, com sede em Milão (local das entrevistas); o caso F corresponde a um banco múltiplo brasileiro, estando no *top-3* dos maiores bancos da América Latina; e o caso G corresponde a um banco europeu do setor de alimentos e agronegócio, cujas pesquisas foram realizadas na subsidiária brasileira do grupo financeiro.

Em última instância, no caso H, foi contemplada uma organização norte-americana do setor de bens de consumo que possui estrutura de governança de informação no âmbito global. Esta organização é uma *family-business* que produz produtos destinados à indústria química e de limpeza doméstica. As pesquisas foram realizadas no escritório de Arese, na Itália.

Os nomes das empresas e dos profissionais consultados foram omitidos nesta pesquisa por fins legais e de sigilo. O Quadro 19 resume os oito casos contemplados na pesquisa, os quais foram tratados segundo a simbologia A, B, C, D, E, F, G e H. As organizações consideradas para os estudos de casos atendem integralmente as premissas para seleção dos casos, mas apresentam particularidades, como é possível observar pelo Quadro 12.

Quadro 19 – Casos contemplados na pesquisa em campo.

Caso	Descrição das empresas consultadas
A	Uma das empresas pioneiras em biotecnologia fundada em 1978 nos EUA. Apresentou crescimento acelerado após os anos 2000 e lançou mais de 5 novos medicamentos após 2010, sendo estes os anos mais produtivos da história da organização. Em 2015, teve faturamento de US\$10 bilhões. O estudo de caso foi realizado no escritório suíço, com líderes da governança de dados.
B	Empresa farmacêutica global e pioneira, fundada na Suíça no final dos anos 1800. Compreende as funções de biotecnologia e de diagnóstico. Está dentre as três maiores empresas farmacêuticas e de biotecnologia do mundo, com faturamento na ordem dos CHF 53 bilhões em 2017. Foi investigada a estrutura de governança de dados da divisão de fármacos na sede da organização, na Suíça.
C	Grupo farmacêutico suíço com mais de 120 mil funcionários, presente em mais de 150 países e com vendas de mais de US\$48 bilhões para o ano de 2015. Está no <i>top-3</i> das maiores empresas farmacêuticas e de biotecnologia do mundo. O estudo de caso foi conduzido na sede da organização, com o grupo de governança de dados de uma unidade de negócios;
D	Grupo bancário líder na Itália, com operações em varejo, corporativo e gestão de riquezas. É um dos maiores grupos bancários da União Europeia. Possui ativos na faixa de US\$ 800 bilhões e faturamento acima de EUR 17 bilhões (para 2016). O estudo de caso foi realizado na sede da organização, na Itália.
E	Grupo italiano de seguros, sendo um dos principais na Europa. Possui mais de 150 anos de história e está presente globalmente, com maior parcela do faturamento proveniente de países europeus. Apresentou faturamento de EUR 70 bilhões para o ano de 2016. O estudo de caso foi conduzido com os líderes da governança de informação, na sede da organização, na Itália.
F	Banco múltiplo brasileiro figurando dentre os três principais da América Latina. Conta com aproximadamente 90.000 funcionários e apresentou faturamento de mais de BRL 1,9 trilhões em 2017 e lucro de mais de BRL 4,5 bilhões em 2017. É uma das principais empresas do mundo, segundo <i>ranking</i> da Forbes. O estudo de caso foi realizado na sede da organização, localizado em São Paulo, no Brasil, com líderes da estrutura de governança de dados e qualidade de dados.
G	Banco europeu no <i>tier-1</i> que atende o mercado global de alimentos e agronegócio. Está presente em mais de 40 países, incluindo o Brasil. Possui mais de US\$ 900 bilhões de ativos, faturamento de mais de EUR12 bilhões (para 2014). No mundo, possui mais de 50 mil funcionários e, no Brasil, 600 funcionários e 15 agências. Possui estrutura de governança de dados no escritório-sede e nas subsidiárias, cujo estudo de caso foi realizado na subsidiária brasileira, localizada em São Paulo.
H	Organização norte-americana de produção de produtos químicos e para a limpeza doméstica. Apresentou faturamento acima de US\$11 bilhões (para 2013) e conta com aproximadamente 13 mil funcionários no mundo. O estudo de caso foi realizado com a liderança da governança de informação no escritório da Itália.

Fonte: elaborado pela autora

O Quadro 20 lista os profissionais consultados por empresa, com a omissão dos dados pessoais. As entrevistas foram realizadas com profissionais pré-selecionados e engajados no tema de pesquisa. Isto corresponde a um importante mecanismo

para a compreensão dos fenômenos que se encontram em estágios iniciais (McCRACKEN, 1988; TALLON et al. 2014).

Quadro 20 – Perfil das entrevistas realizadas nos estudos de caso.

Setor	Empresa	Cargo do entrevistado	Local
Farmacêutico	A	<ul style="list-style-type: none"> • Diretor, Gerenciamento de Meta-Dados • Gerente de Governança de Dados 	Suíça
	B	<ul style="list-style-type: none"> • Consultor líder de TI para <i>Big Data</i> • Gerente de Governança de Dados 	Suíça
	C	<ul style="list-style-type: none"> • Gerente Global, Modelagem de Gestão da Informação • Gerente de Governança de Dados 	Suíça
Financeiro	D	<ul style="list-style-type: none"> • <i>Chief Data Officer</i> 	Itália
	E	<ul style="list-style-type: none"> • Líder de Governança de Informação • Líder do Departamento de TI & Estratégia das Operações 	Itália
	F	<ul style="list-style-type: none"> • <i>Chief Data Officer</i> • Superintendente de Governança de Dados 	Brasil
	G	<ul style="list-style-type: none"> • Gerente de qualidade dos dados • Consultor de gestão de dados 	Brasil
Bens de consumo	H	<ul style="list-style-type: none"> • Diretor, Gerenciamento de Meta-Dados • Gerente de Governança de Informação 	Itália

Fonte: elaborado pela autora

3.4.1 Caso A

O caso A corresponde a uma das empresas pioneiras em biotecnologia fundada em 1978 nos Estados Unidos. Apresentou faturamento de US\$20 bilhões (em 2017) e foi responsável por tratar cerca de 38% dos pacientes globais com esclerose múltipla (aproximadamente 900 mil pacientes). A organização apresentou crescimento acelerado após os anos 2000, onde seu faturamento cresceu 14% entre os anos 2013 e 2017. Após os anos 2010, foram lançados mais de 5 novos medicamentos, sendo estes anos os mais produtivos da história da organização.

Seu *core business* contempla as seguintes doenças: esclerose múltipla, neuroimunologia, Alzheimer, demência, Parkinson, desordem de movimento e doenças neuromusculares. Outros campos emergentes na organização contemplam tópicos de dor, oftalmologia e neurologia aguda. É objetivo da organização maximizar o *core business* da organização e criar novas fontes de valor na neurociência, ao desenvolver e expandir o portfólio de terapias e medicamentos e

re-priorizar o capital alocado da organização. Ela almeja ser líder global em neurociência.

Dentre os objetivos estratégicos da organização está o desenvolvimento de novos tratamentos e medicamentos no campo da neurociência e, para isto, é necessário contemplar o ambiente *big data* em seus centros de P&D. Esta organização contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985). A organização apresenta capacidades analíticas transformadas com vistas aos investimentos em tecnologias e experiência em projetos de BI para alcance de vantagens competitivas. O uso de tecnologia de ponta é fundamental para a descoberta de novas terapias no campo da biotecnologia.

A empresa A possui 7.350 funcionários altamente qualificados e tecnologias de ponta para criação, comercialização e manufatura de terapias e medicamentos. A organização possui laboratórios de P&D no estado-da-arte da pesquisa em neurociência e fábricas próprias na Suíça e nos Estados Unidos. Ela pode ser classificada no quadrante estratégico no *Grid* de MacFarlan, já que a TI é fundamental e estratégica para o desenvolvimento dos produtos da organização. Ela contempla ambiente *big data* e portfolio de produtos focalizado no campo da neurociência. Apesar disto, seus negócios requerem alta especificidade e individualidade devido às particularidades dos tratamentos, dos projetos e da regulação dos países.

No âmbito desta pesquisa, foram entrevistados um diretor de gerenciamento de meta-dados e um gerente de governança de dados que executam suas atividades na escala global da organização no grupo funcional da cadeia de suprimentos (*supply chain*) e *procurement*. Tal estrutura se reporta ao CFO da organização, não pertencendo ao departamento de TI e nem às unidades de negócios que desenvolvem terapias e medicamentos por nicho de doenças.

As atividades do gerenciamento de meta-dados são orientadas por processos, independente do produto. A única particularidade corresponde aos produtos comerciais, que já estão sendo vendidos, e aos clínicos, que estão em desenvolvimento e ainda não são comercializados. A partir do momento que os produtos clínicos estão em fase de teste em pacientes, devem ser contemplados os esforços de gerenciamento de meta-dados para avaliar e monitorar a qualidade dos produtos.

A estrutura de governança de dados possuía 3 anos de existência com referência à data das pesquisas em campo (1º semestre de 2017). A equipe possui 8 indivíduos atuando na esfera global para a definição e o gerenciamento de meta-dados para produtos e materiais. No âmbito dos produtos, podem ser empregadas diferentes formas para descrever um produto farmacêutico, como o uso do nome comercial, nome genérico, tamanho da embalagem, dentre outras características. O grupo de governança de dados estabelece os padrões dentre diferentes cenários de negócios. Além disto, o controle dos materiais relaciona-se às matérias-primas e os intermediários, como químicos, substâncias específicas, embalagens. Para o líder da governança na organização:

“O objetivo da governança de dados é realmente decidir quem faz o que e como padronizar todos os dados, como gerenciar as mudanças nessas estruturas de dados. Temos que alinhar os papéis, as responsabilidades e os padrões de dados”.

Dentre as atividades da governança, são contemplados o atendimento aos aspectos legais, de *compliance* e operacionais, como o alinhamento de diferentes departamentos em como os dados devem ser definidos, como os dados devem ser gerados, como os dados são armazenados e como os dados podem ser utilizados em diferentes aplicações de TI e objetivos. Devem ser determinados quem na empresa, qual departamento, qual função é responsável por fornecer determinada informação, aprovar ou revisar. Considera-se também o nível de sensibilidade do dado, o qual pode influenciar diretamente na qualidade de um produto ou na aprovação de um novo medicamento pelo órgão regulador ou portar alguma informação confidencial.

O motivo para a adoção da governança de dados foi a necessidade de melhor organizar os processos internos em função do rápido crescimento que a organização apresentou em escala global:

“Precisamos de processos eficientes para estruturar, centralizar e estabelecer o fluxo de trabalho de diferentes pessoas através dos dados”, comenta o líder do gerenciamento de meta-dados na organização.

Diferentemente das outras organizações, a empresa A passou de uma *start-up* para uma organização global em um curto período de tempo (nos últimos 10 anos), perdendo eficiência operacional:

“Cinco anos atrás, a operação ainda era pequena. Éramos uma equipe de 5 pessoas para planejamento, regulatório e qualidade e elas se conheciam muito bem, de forma que os problemas eram resolvidos à medida que apareciam. Agora a organização cresceu, as equipes cresceram, as pessoas não se conhecem mais tão bem. E o fluxo de trabalho não está mais funcionando. Então tornou-se mandatório aprimorar o fluxo dos processos”, explica o gerente de governança de dados.

A equipe de governança de dados consultada localiza-se no escritório internacional da organização na Suíça, onde são empregados 400 funcionários representando os negócios em 33 países no mundo. Este escritório é responsável por fornecer suporte comercial e funcional aos negócios fora dos Estados Unidos. Os funcionários deste escritório colaboram com parceiros de distribuição na América Latina, Oriente Médio e Leste Europeu que operam na venda das terapias e dos remédios. Trata-se de um escritório de referência para as operações externas dos EUA, com impacto global.

3.4.2 Caso B

O caso B corresponde a um grupo farmacêutico global e pioneiro, fundado na Suíça no final dos anos 1800. O grupo contempla as funções de biotecnologia e de diagnósticos sob duas divisões: a B Farmácia e a B diagnóstico, sendo estes nomes fantasia para proteger os aspectos legais da marca. O grupo apresentou crescimento nas vendas em 5% para o ano de 2017 em ambas as divisões de negócios e faturamento na ordem dos CHF 53 bilhões em 2017. O caso B está dentre os três maiores grupos farmacêuticos e de biotecnologia do mundo.

A divisão B Farmácia contempla áreas de negócio distribuídas em oncologia, neurociência, doenças infecciosas, imunologia, cardiologia e metabolismo, oftalmologia, hematologia, doenças respiratórias. Dentre esses segmentos de negócios são contempladas atividades de P&D de novas terapias e manufatura dos produtos farmacêuticos. Esta divisão contempla intensamente o uso de *big data*

para as atividades de descoberta de novo produtos e possui capacidades analíticas transformadas.

A divisão B Diagnósticos do grupo contempla as atividades de diagnósticos que visam a construção e a configuração de instrumentos ou soluções para laboratórios e hospitais para executar diagnósticos baseados em química, mineralogia, análise de tecidos, análises moleculares, dentre outros parâmetros.

Ambos os segmentos de negócios, B Farmácia e B diagnóstico tendem a se unir cada vez mais no ambiente *big data* para desenvolver e testar novas terapias. Em particular, no campo da oncologia, o tratamento dos indivíduos requer terapias mais individualizadas e, para tanto, os testes na divisão de diagnósticos devem evoluir para fornecer parâmetros corretos e mais específicos para suportar as exigências na B Pharma em oncologia.

Tanto a B Farmácia quanto a B diagnóstico possuem departamentos de TI independentes que almejam suportar os negócios em diferentes atividades, como gerenciamento das vendas, controle das finanças, gerenciamento da cadeia de suprimentos, suporte ao desenvolvimento de novas terapias, novas soluções, dentre outros campos. No topo da organização, o CEO do grupo é suportado pelos CEO's das divisões Farmácia e Diagnósticos. Sob o CEO de cada divisão estão compreendidos os líderes das unidades de negócios e dos grupos funcionais como RH, Finanças.

O departamento de TI reporta ao CFO de cada divisão. E, de forma matricial, os departamentos de TI da B Farmácia e da B diagnóstico constituem um grupo de TI que possuem como base serviços de infraestrutura (oferecido por um grupo específico). Tais serviços fornecem a infraestrutura mais baixa e as plataformas empregadas pelo grupo. Em particular, no departamento de TI da B Farmácia existe um grupo dedicado a BI que trabalha em conjunto ao grupo de governança de dados estudado nesta pesquisa.

Com vistas ao posicionamento estratégico da TI na organização, o grupo pode ser classificado no quadrante estratégico no *Grid* de McFarlan (1984). Além disto, o grupo apresenta alto conteúdo de informação no produto e alta intensidade de informação na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985). Trata-se de um grupo com portfolio complexo de produtos e serviços, onde foi localizada uma estrutura de governança de dados dentre os negócios farmacêuticos (B Farmácia). O segmento de diagnósticos (B Diagnósticos) não

contempla estrutura de governança de dados ou de informação, embora seja um negócio orientado aos dados, pois são coletados inúmeros dados dos serviços de diagnósticos oferecidos pelo grupo a nível global.

A estrutura da governança de dados foi adotada em 2017 e possuía 6 meses de existência na organização (com relação à data das entrevistas). O líder da governança de dados responde ao CFO da organização. Seu escopo de atuação restringe à área de suprimentos e *procurement* (aquisições) e tem como objetivo central o gerenciamento de meta-dados de produtos e materiais a nível global. Podem existir outros grupos de governança de dados na organização para outros segmentos de negócios, embora na ocasião não tenham sido identificadas outras estruturas.

O motivo para a criação da estrutura de governança de dados correspondeu a necessidade de melhor organizar o fluxo das informações para assegurar a singularidade das informações, garantir o aumento da qualidade dos ativos de dados e melhor fundamentar os processos de tomada de decisões quando se trata da manufatura das terapias e dos remédios. Logo, trata-se de um grupo de trabalho orientado por processos que focaliza a manufatura e não ao P&D de novas drogas e terapias. Foram contempladas entrevistas com um consultor líder de TI para *big data* na divisão B Diagnósticos e um gerente de governança de dados da B Farmácia.

O grupo possui aproximadamente 94.000 funcionários no mundo. Devido ao seu tamanho e complexidade, existem diversos silos operacionais nas divisões. As vendas do grupo para o ano de 2016 foram de CHF 50 bilhões, onde B Farmácia contemplou cerca de CHF 39 bilhões e a B Diagnósticos 11 bilhões. Cerca de 27 milhões de pacientes foram tratados com um dos top 25 produtos da B Farmácia. A organização investiu significativamente em tecnologias de BI nas últimas décadas e investe em tecnologias de *big data* tanto na divisão Farmácia quanto na Diagnósticos. Trata-se de um caso com capacidades analíticas transformadas, segundo o modelo de LaValle et al (2011).

3.4.3 Caso C

O caso C corresponde a um grupo farmacêutico suíço com mais de 126 mil funcionários e está presente em 155 países. Trata-se de uma organização com vendas de mais de US\$49 bilhões em 2017 e está no *top-3* das maiores empresas farmacêuticas e de biotecnologia do mundo. Dentre os valores corporativos, está a inovação, a qualidade, colaboração, desempenho, coragem e integridade. A missão do grupo é descobrir novas formas de melhorar e prolongar a vida humana.

A empresa C se estrutura em três unidades de negócios que contemplam a oncologia, os genéricos e a oftalmologia. Além disto, fazem parte da empresa grupos funcionais como o P&D e as atividades operacionais de manufatura. Em particular no ano de 2017 foram conduzidas iniciativas estratégicas para melhor integrar as atividades de P&D com as manufaturas da organização.

A empresa C pode ser classificada no quadrante estratégico no *Grid* de MacFarlan, já que a TI é fundamental e estratégica para o desenvolvimento dos produtos na organização. Além disto, ela contempla o ambiente *big data* e contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

A organização contempla portfolio de produtos diversificado e complexo, além de possuir capacidades analíticas desenvolvidas e distribuídas nas unidades de negócios da organização. Estas competências de BI são empregadas para o desenvolvimento de novos produtos, análise do ambiente competitivo, na cadeia de suprimentos, dentre diversos campos da organização. O caso C contempla capacidades analíticas transformadas, segundo o modelo de LaValle et al (2011).

Com vistas às buscas realizadas na plataforma do LinkedIn, foram identificados *stakeholders* de governança de dados. Logo, as premissas para seleção do caso foram atendidas integralmente. Ao conduzir as entrevistas com a estrutura de governança de dados, foi observado que ela está vinculada a uma unidade de negócios, denominada por “Desenvolvimento Global de Drogas”. Trata-se de uma divisão vinculada ao P&D da organização que aborda holisticamente as unidades de negócios para o desenvolvimento de novas terapias e fármacos. Esta divisão avalia regularmente o potencial de novos produtos para compor o portfolio da organização e assegura que recursos sejam alocados nos projetos mais promissores.

A estrutura de governança de dados consultada foi estabelecida em Janeiro de 2017 na divisão de Desenvolvimento Global de Drogas, cujo líder da governança de dados se reporta matricialmente ao CFO da organização e ao *head* da divisão de desenvolvimento de drogas.

O modelo de governança de dados da organização “Desenvolvimento Global de Drogas” considera seus esforços voltados aos negócios (*business-driven*) e orientados aos processos (*process-driven*). O modelo estabelece padrões e procedimentos comuns na unidade de negócios com relação aos ativos de dados com o intuito de aumentar a eficiência operacional e maiores índices de eficácia.

A estrutura foi criada com o intuito de harmonizar os dados entre os diferentes *stakeholders* da divisão, definir padrões nos dados, certificar que tais padrões estão sendo aplicados corretamente às informações coletadas, gerenciar as mudanças com relação às informações no ecossistema de desenvolvimento de drogas da organização a nível global. Cabe a equipe de governança definir as regras e os procedimentos com base nos pontos críticos observados no fluxo de trabalho dos *stakeholders* da divisão, repassado tais mudanças aos líderes dos negócios e ao departamento de TI.

“Trata-se de uma função de negócios e não uma função de TI. Nós definimos os padrões e as regras de negócio mas não gerenciamos o trabalho operacional. Nós não selecionamos ferramentas. Nós suportamos mas não estamos fazendo o trabalho de TI. Nosso trabalho é orientado aos processos”, cita o gerente de governança de dados.

Esta estrutura interage com diferentes *data stewards* das unidades de negócios que podem estar dispersos no globo. Trata-se de um ponto focal e único na organização que possui a visão geral dos processos para o desenvolvimento de novas drogas. Esta estrutura se posiciona no escritório sede da organização pela exposição que possui perante os demais *stakeholders* distribuídos no mundo.

“O processo de gerenciamento de dados não se relaciona ao país. É orientado pelos negócios. Podem existir algumas particularidades, mas a forma como gerenciamos os dados na Itália ou no Brasil deve ser a mesma”, comenta o gerente de dados da empresa.

Além do grupo dedicado à governança de dados, a organização de Desenvolvimento de Drogas contempla grupos de TI, testes clínicos, *reporting* e

operações. O grupo de TI suporta a equipe de governança em todas as fases do trabalho, pois auxilia na extração de informações dos sistemas para estudo pela equipe de governança e operacionaliza as mudanças nos sistemas de informação apontadas pela estrutura de governança de dados.

Foram consultados o gerente global de gerenciamento de informação e modelagem e o gerente de governança de dados situados no escritório na Suíça. Tais *stakeholders* desenvolvem ações relativas ao desenvolvimento de drogas para todos os países em que há operação do grupo. Em particular, a estrutura de governança de dados reside na sede da organização e atende os processos independente da geografia.

3.4.4 Caso D

O caso D corresponde a um grupo bancário europeu líder na Itália, contemplando negócios no varejo, no segmento corporativo e para a gestão de riquezas. É um dos maiores grupos bancários da União Europeia, apresentando 4.700 *branches* e 12,3 milhões de clientes. O grupo bancário possui aproximadamente 97.000 funcionários, apresenta ativos na faixa dos US\$ 800 bilhões e faturamento acima de EUR 17 bilhões para 2016. A formação do grupo ocorreu em 2007 após a fusão de dois grandes bancos italianos. No entanto, os primórdios dos bancos que deram origem à série de fusões e aquisições que levaram a constituição deste grupo se remete aos anos 1823.

O grupo possui presença considerável na Europa Central, Oriente Médio e Norte da África e seus clientes corporativos são atendidos em mais de 25 países, incluindo o Brasil. Trata-se de uma organização voltada aos serviços financeiros, com portfólio complexo de serviços, podendo ser classificada no quadrante estratégico do *Grid* de MacFarlan. Neste caso, a TI é estratégica para a organização.

O caso D corresponde a um banco múltiplo, cujos objetivos estratégicos contemplam a extração de valor dos dados para orientar os negócios com base em dados. O banco apresentou aumento de 15% na eficiência operacional através da digitalização dos processos. A instituição está rigorosamente investindo em técnicas de *big data* para extração de valor dos dados. O banco possui um “*data lake*” com

robusto modelo de governança de dados com dedicado *Data Officer*, 150 *data owners* e 30 especialistas de qualidade dos dados. São gerados mais de 5.000 relatórios oriundos da descoberta de dados.

O banco também estabeleceu em 2016 um *Big Data Lab* para fomentar o avanço analítico da organização. Este laboratório conta com cientistas de dados dedicados para conduzir iniciativas piloto dentre as várias divisões de negócios do banco e fomentar o avanço analítico. A organização D contempla alto conteúdo de informação no produto e alta intensidade de informação na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

A estrutura de *Chief Data Officer* foi criada no *head office* da organização, sob responsabilidade do CFO da organização. Dentre as unidades de negócios compreendem-se a divisão de varejo, investimentos e corporativo, subsidiárias internacionais, *private banking*, gestão de ativos, seguros e capitais. A função de CDO apresentava dois anos de existência com relação à data de realização das entrevistas (1º semestre de 2017). Como apontado pelo CDO:

“Diferentemente das organizações focadas em dados como Facebook, Google e Amazon, os dados para os bancos (múltiplos) sempre foram tratados como sub-produtos de processos e produtos. O foco dos bancos (múltiplos) nunca foi nos dados. Nosso objetivo é mover o banco para a perspectiva data-driven e promover mudanças na mentalidade organizacional com relação aos dados”.

A pesquisa foi conduzida com o *Chief Data Officer* da instituição, localizado na Itália, que ressalta que a estrutura de governança de dados possuía 2 anos de existência (com relação ao momento das pesquisas em campo). A estrutura foi criada em 2015. Vale mencionar que ao CDO da empresa compreende-se os desafios de governança de dados, gerenciamento de dados, aumento da qualidade dos dados e atividades de ciência dos dados no âmbito do *Big Data Lab*.

“O Chief Data Officer assume o papel de empurar a inteligência dos dados para a inteligência artificial, aprendizado de máquina. Nosso papel contempla, também, o estudo e o desenvolvimento de novos casos para o avanço analítico da organização”, comenta o CDO sobre os objetivos do big data lab.

A decisão de alocar a governança de dados no âmbito financeiro se deve à perspectiva orientada aos negócios e não orientada à TI. O departamento de TI é fundamental para a organização e fornece a infraestrutura tecnológica básica que sustenta os bancos de dados, que viabiliza as operações e a oferta dos serviços financeiros. O departamento de TI é único para a organização e se reporta ao *C-level* de tecnologia (e não para o CFO para o caso do *Chief Data Officer*).

Como comentado pelo entrevistado, quando a estrutura de governança está sob um *Chief Financial Officer*, trata-se de um projeto de negócios e isto requer profissionais focados em negócios e na monetização dos projetos. Por outro lado, modelos de governança subordinados a um *Chief Operations Officer* possuem expectativas sobre o serviço oferecido para outros departamentos da organização.

A organização D apresentou investimentos significativos em BI e possui experiência significativas em projetos e tecnologias de BI. O interesse do grupo é expandir as capacidades analíticas na orientação do *advanced analytics*. Corresponde a um caso com capacidades analíticas transformadas, segundo o modelo de LaValle et al. (2011).

3.4.5 Caso E

O caso E corresponde a um grupo italiano de seguros constituído em 1831, tornando-se uma das principais seguradoras na Europa. Está presente em mais de 60 países, possui mais de 74.000 funcionários no mundo e mais de 55 milhões de clientes. Apresentou faturamento de EUR 70 bilhões para 2016, sendo a maior parte deste faturamento proveniente de três países europeus: Alemanha, França e Itália.

A organização contempla negócios em seguros que são utilizados para gerenciar incertezas e riscos futuros das mais diversas formas sobre as pessoas, organizações, bens. Os dados desempenham papel estratégico para a organização comercializar seus produtos e antever situações futuras. A organização contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

Para a organização, a TI desempenha papel estratégico, podendo classificá-la no quadrante estratégico do *Grid* de MacFarlan (1984). Além disto, a organização

contempla ambiente *big data*, com o interesse em expandir a coleta de dados das redes sociais e da IoT para o desenvolvimento de modelos analíticos avançados.

Desde 2016 a organização tem implementado reestruturações em suas operações com foco em inovação, digitalização e transformação cultural *data-driven*. Neste contexto, foi criada a estrutura de governança de informação na organização para atuar sobre os ativos de dados e prepará-los para que técnicas analíticas mais avançadas possam ser empregadas na corporação.

No contexto do estudo de caso, foram entrevistados o líder de Governança de Informação e o líder do Departamento de TI & Estratégia das Operações na sede da organização, em Milão (Itália). A estrutura de governança de informação está alocada sob responsabilidade do CFO do grupo, ao passo que o departamento de TI está alocado sob o COO.

A estrutura de governança de informação trabalha em conjunto com o departamento de *analytics* para criar valor dos dados e das informações. Almeja-se transformar a organização para o contexto *data-driven* e tornar as tomadas de decisão mais embasadas em análises quantitativas avançadas. Para a empresa E, a governança de informação deve estabelecer um modelo operacional definindo papéis e responsabilidades com relação aos dados e assegurar que os dados e as informações possuam alta qualidade.

A empresa E ressalta a relevância da governança de informação. Para o líder da governança de informação:

“Governança de dados corresponde a uma suposição genérica com o intuito de fornecer dados com alta qualidade. Mas nós queremos ir além disto. Queremos transformar dados em informação para extrair valor para a organização e, portanto, almejamos a governança de informação”.

Para o contexto do modelo de governança de informação, tem-se a definição das responsabilidades com relação aos ativos de dados, ao definir *data owners* e *data stewards* dentre as unidades de negócios da organização, nas diversas partes do mundo. Outras atividades centrais do modelo de governança de informação em vigor na organização incluem: a definição e o gerenciamento de meta-dados, atuar sobre a qualidade dos dados, definindo princípios, ferramentas e formas de controle.

A preocupação regulatória também influencia nas atividades de governança de informação, onde deve existir cooperação entre as demais operações no mundo para que particularidades legais de cada país possam ser contempladas. Em particular, é mencionado o trato com informações sensíveis de seus clientes, onde as regulações do sistema financeiro são mais severas quanto ao uso e ao armazenamento destas informações.

A estrutura de governança de informação desenvolve ações para as diversas partes da organização. Os líderes ressaltaram o papel do corpo executivo para a facilitação da adoção das medidas de governança de informação pela organização. Quanto mais descentralizada a organização, como é o caso da empresa E, maiores níveis de negociação são necessários entre as diversas partes. Uma série de reuniões são necessárias entre as partes para alinhar as medidas e convencê-los sobre os benefícios dos métodos de governança de informação.

A organização contempla casos de sucesso com ferramentas de BI e possui competências desenvolvidas em termos de pessoas, ferramentas e processos. Segundo o modelo de LaValle et al. (2011) referente às capacidades analíticas, o caso E apresenta características na faixa experiente. A organização almeja a expansão do uso da BI e das técnicas analíticas avançadas, onde os esforços de governança são necessários para melhor constituir e organizar os ativos de dados.

3.4.6 Caso F

O caso F corresponde a um banco múltiplo brasileiro que figura dentre os três principais da América Latina. Conta com aproximadamente 90.000 funcionários e apresentou faturamento de mais de BRL 1,9 trilhões em 2017 e lucro de mais de BRL 4,5 bilhões em 2017. É uma das principais empresas do mundo, segundo *raking* da Forbes.

Trata-se de uma organização voltada aos serviços financeiros, com portfólio complexo de serviços, podendo ser classificada no quadrante estratégico do *Grid* de MacFarlan. A TI é estratégica para a organização e recebe intensivos investimentos anuais. Além disto, o caso F contempla o ambiente *big data* e possui alto conteúdo

de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

A estrutura de *Chief Data Officer* foi criada em 2017 no escritório sede da organização, sob responsabilidade do CRO da organização. Ela possuía 4 meses de existência com relação à data de realização das pesquisas em campo.

O departamento com o CDO é denominado por “Departamento de Gestão de Dados”. Este departamento compreende duas superintendências: uma para o desenvolvimento de modelos e algoritmos e a outra para a governança e qualidade de dados. No caso, foram consultados o *Chief Data Officer* e o Superintendente de Governança e Qualidade de Dados.

Dentro da superintendência de gestão de dados, existem os domínios de negócios: *Analytics*, BI, Governança de dados e Qualidade. Para a *analytics* compreende-se o preparo de dados para servir a superintendência de algoritmos, também do departamento de gestão de dados. Para a BI compreende-se as funções de visualização de dados para o alto escalão executivo, utilizando-se ferramentas como TableauTM, Qlik ViewTM. Para a governança de dados compreende-se as funções de acessibilidade dos dados na corporação, transformação da cultura da organização para *data-driven*, definição e gerenciamento de meta-dados. O domínio qualidade de dados é contemplando a parte da governança de dados devido à sua relevância para o negócio e o volume de trabalho para desenvolvimento e adoção das ferramentas de qualidade dos dados. Apesar de ser tratado por um domínio a parte da governança de dados, o líder da área entende que a qualidade é um campo da governança de dados.

O grupo de governança e qualidade de dados atua sobre toda a organização, é orientada aos negócios e aos processos. Trata-se de uma estrutura desvinculada de outras unidades de negócios da organização ou grupos funcionais, sendo suportada pelo alto escalão executivo. Dentre seus objetivos estratégicos compreende a transformação cultural no sentido *data-driven*.

Dentre os objetivos da governança de dados, são contempladas atividades de definição e gerenciamento de meta-dados, atuar sobre dados sensíveis para proteção regulatória e atendimento das exigências de sigilo, definir papéis e responsabilidades com relação aos dados. No entanto, foram reportadas dificuldades operacionais com vistas ao volume de dados e tamanho organizacional:

“Às vezes se torna impossível estabelecer meta-dados devido ao alto volume de dados e bases de dados utilizadas”, comenta o líder da governança de dados do caso F.

A atuação sobre a qualidade dos dados é tida como fundamental para o aumento da eficiência operacional no desenvolvimento de novos modelos analíticos, sob responsabilidade da outra superintendência do departamento de gestão de dados. Logo, o estabelecimento de medidas de governança de dados na organização almeja contribuir para o avanço analítico da corporação e aumentar a eficiência operacional para desenvolvimento de novos modelos.

Apesar disto, a organização pode ser classificada no nível transformada com relação ao modelo de classificação das capacidades analíticas proposto por LaValle et al. (2011). Ela contempla casos de sucesso com BI e possui interesse em melhor utilizar a *analytics* em seus negócios.

3.4.7 Caso G

O caso G trata de um dos bancos líderes no mercado global de alimentos e do agronegócio. A organização é provedora internacional de serviços financeiros para mais de 8 milhões de clientes no mundo (em particular fazendeiros, produtores de alimentos, organizações do agronegócio ou de alimentos) e com mais de US\$ 900 bilhões de ativos. Trata-se de um banco europeu constituído em 1972 a partir de cooperativas de créditos para produtores rurais. No entanto, suas origens se remetem aos anos 1800 quando fazendeiros fundaram a primeira cooperativa de crédito para a agricultura.

A organização G possui mais de 380 escritórios no mundo, está presente em 40 países e possui aproximadamente 50 mil funcionários no mundo. No Brasil, conta com 600 funcionários e 15 agências. Trata-se de um banco classificado no *Tier-1* do mercado financeiro, apresentando faturamento de mais de 12 bilhões de euros (para 2014).

No ano de 2016 o banco começou a implementar seu modelo estratégico 2016-2020, o qual descreve as ações para aperfeiçoar o serviço aos clientes e a base financeira da organização para assegurar um futuro de sucesso. Em particular, o

ano de 2017 foi relevante para o estabelecimento de ações estratégicas para a transformação digital da organizacional, onde foram implementadas novas estruturas na organização com vistas aos dados. A organização contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

Além disto, a TI desempenha papel estratégico na organização e suporta todas as operações e produtos financeiros da organização. Logo, a organização G pode ser classificada no quadrante estratégico do *Grid* de MacFarlan (1984). Quanto as capacidades analíticas da organização, estas foram observadas transformadas segundo o modelo de LaValle et al. (2011).

Quanto as buscas pelos profissionais na ferramenta LinkedIn, foram identificados “*business data stewards*” e “*data quality manager*” em diversas localidades da empresa G. Logo, a organização foi considerada para estudo de caso para a investigação do modelo de governança de dados adotado. Foram conduzidas entrevistas no escritório de São Paulo (Brasil) da empresa G com a gerente de qualidade e o consultor de gestão de dados. Logo, foi objeto de investigação a estrutura de governança de dados na subsidiária brasileira.

A governança de dados corresponde a um tema recente na organização e encontra-se em fase de estruturação. Ela está se tornando relevante para garantir a qualidade da informação que é usada para fins comerciais, regulatórios e para elaborar os relatórios internos e externos da organização. Ela almeja promover a transformação cultural da organização em *data-driven*.

A função “qualidade dos dados” surgiu em 2017 na organização a nível global, inclusive para as operações no Brasil. Esta área se reporta ao CFO de cada escritório regional ou da matriz. No caso brasileiro, a gerente de qualidade de dados se reporta ao CFO das operações no Brasil. Complementarmente, as questões de gestão de dados, definição de meta-dados, acessibilidade dos dados, integração das bases são conduzidas pela equipe de gestão de dados alocada no departamento de TI da organização, que se reportam ao COO da organização. Os cargos relacionados aos “*data stewards*” também foram estabelecidos em 2017 no departamento de TI.

Diferentemente do observado na subsidiária brasileira, o escritório sede da organização contempla “*data stewards*” alocados no “*Data Reporting Center*”, uma estrutura em formação na organização que responde ao Centro de Finanças &

Risco, sob responsabilidade do CFO. Na sede da organização, o *Data Reporting Center* contempla também profissionais para a gestão da qualidade dos dados e gerentes para reportar problemas de qualidade dos dados. Logo, existem diferenças estruturais quanto ao escritório sede e a subsidiária brasileira no que tange ao posicionamento dos “*data stewards*”.

A estrutura “qualidade dos dados” está sob responsabilidade de um executivo de finanças pois ele é o responsável contábil na organização e deve assegurar que os dados reportados para a sede da organização estejam corretos. Esta área deve assegurar que os dados empregados pelos negócios e aqueles utilizados nos relatórios possuam qualidade. A qualidade é fundamental para os serviços financeiros, embora seja um tema bastante complexo da organização G, já que os fluxos de dados permeiam elevada quantidade de sistemas desintegrados. Com vistas a este cenário, a área de “qualidade dos dados” atua com forte colaboração com a área de gestão de dados, uma área da TI. As atividades contempladas pelo COO são orientadas aos projetos, enquanto as atividades do CFO incluem a contabilidade e finanças.

São objetivos da governança de dados no país: estabelecer responsáveis no departamento de TI para atuar como *data stewards* e para o gerenciamento de dados; estabelecer equipe para a gestão da qualidade dos dados para implementar ferramentas de qualidade dos dados, gerenciar e controlar a qualidade dos dados nas operações na subsidiária brasileira; atuar sobre a definição de métricas e parâmetros para a promoção da integridade, corretude e veracidade dos dados; definir e controlar meta-dados; suportar as unidades de negócio da organização para o tema qualidade e gestão de dados.

Além disto, o grupo do *Data Reporting Center* na sede da organização também atua ativamente para a adoção de tecnologias para suportar a gestão de dados e o controle da qualidade dos dados a nível global. Um exemplo de tecnologia adotada pelo escritório sede foi a plataforma CollibraTM. Apesar de ser uma plataforma que permite o mapeamento do fluxo dos dados no ambiente *big data*, os entrevistados na subsidiária brasileira reportaram dificuldades operacionais para alimentar a plataforma com os fluxos das operações visto a enorme quantidade de sistemas e dados empregados pelas unidades de negócios.

3.4.8 Caso H

O caso H corresponde a uma empresa de bens de consumo de uma *holding* norte-americana fundada em 1886, cujos produtos são distribuídos para mais de 160 países. A *holding* possui de 120.000 funcionários no mundo e apresentou faturamento acima de US\$70 bilhões no ano de 2016. O caso H contempla a manufatura de produtos para a indústria química e para a limpeza doméstica, apresentando faturamento acima de US\$11 bilhões para o ano de 2013 e mais de 13 mil funcionários no mundo. Tanto o caso H quanto a *holding* correspondem a uma *family-business*, isto é, uma organização privada controlada por uma família.

O caso H contempla unidades de negócios para a produção e distribuição dos produtos em escala global, cujas operações e produtos devem cumprir políticas de sustentabilidade definidas pela organização. A sustentabilidade corresponde a um tópico estratégico para a organização, onde reformulações de seus produtos ou processos produtivos ocorrem com frequência para diminuir impactos ambientais. Além disto, a organização retratou a necessidade de atendimento das regulações do mercado químico global, sendo estas bastante exigentes e intensas.

O departamento de TI da organização conta com a liderança do CIO, sendo a maior parte do departamento alocado na sede da organização nos Estados Unidos. No entanto, alguns líderes da governança de informação encontram-se dispersos no mundo pois suas funções foram resultado da reestruturação de algumas funções de BI. No estudo de caso conduzido, foram consultados o diretor de gerenciamento de meta-dados e a gerente geral de governança de informação, que se localizam no escritório da Itália do caso H. Diferentemente dos outros casos, a governança de informação do caso H está integralmente alocada no departamento de TI.

A empresa H pode ser classificada no quadrante em transição no *Grid* de MacFarlan, já que a TI migra de uma situação mais discreta para uma de maior destaque na estratégia da empresa devido às novas aplicações de TI. Além disto, ela contempla o ambiente *big data* e contempla alto conteúdo de informação no produto e alta intensidade de informações na cadeia de valor na Matriz de Intensidade de Informação de Porter e Millar (1985).

A estrutura de governança de informação consultada atende globalmente os escritórios e as operações e podem atuar pró-ativamente na avaliação de processos

de negócios internos para sugerir mudanças com relação aos padrões, siglas e meta-dados para um processo de negócio específico. Neste caso, podem ser contemplados aspectos da manufatura, da cadeia de suprimentos, como a distribuição de produtos, compra de matérias-primas, embalagens, dentre outros tópicos. Além disto, alguma unidade de negócios ou grupo funcional (como RH, jurídico, *marketing*) pode solicitar o engajamento da equipe de governança de informação para desenvolver projeto sob demanda para promover a melhoria de algum aspecto sobre os ativos de dados por eles identificado.

Dentre os objetivos da estrutura de governança de informação do caso H cabe definir responsabilidades com relação aos dados, assegurar que as informações estejam corretamente disponibilizadas para os *stakeholders* na organização e assegurar a qualidade dos ativos de dados. Além disto, foi identificada a atuação da estrutura de governança de informação na transformação cultural da organização para o contexto *data-driven* e engajamento no atendimento e no monitoramento dos aspectos regulatórios do setor químico.

Tais pontos podem ser observados através do comentário da gerente geral de governança de informação:

“O objetivo principal da estrutura é assegurar que cada função possui a informação correta para fazer seu trabalho e que as decisões sobre os dados sejam feitas nos mesmos lugares onde são tomadas as decisões de negócios”, como menciona a gerente geral de governança de informação.

No âmbito das competências analíticas, o caso H apresentou características aspiracionais do modelo de LaValle et al. (2011). A organização busca eficiência operacional, automação dos processos existentes e redução de custos, estando mais distantes de obter resultados da BI. Um dos objetivos para a adoção da estrutura de governança de informação correspondeu a necessidade de avançar a capacidade analítica da organização. Para isto, os esforços organizacionais para melhor organizar e gerenciar os ativos de dados são fundamentais para que técnicas analíticas mais avançadas possam ser empregadas sobre ativos de dados.

4. Resultados e discussões

Para responder à questão geral de pesquisa: “como as organizações podem adotar as governanças de dados ou de informação no ambiente *big data*”, foram definidas duas perguntas de pesquisa e delineadas três proposições de pesquisa conforme revisão de literatura. Para tanto, este capítulo se estrutura em cinco momentos para apresentar os resultados e expor as análises desta pesquisa:

1. Objetivos gerais das estruturas de governança de dados ou de informação (sessão 4.1);
2. Elementos e conteúdos das governanças de dados ou de informação (sessão 4.2);
3. Capacidades analíticas e a governança de dados ou de informação (sessão 4.3);
4. Proposições de pesquisa (sessão 4.4);
5. Fatores motivadores, viabilizadores e determinantes para a adoção das governanças de dados ou de informação no ambiente *big data* (sessão 4.5).

4.1 Objetivos gerais das estruturas de governança de dados ou de informação

Os casos foram selecionados segundo os critérios listados pelo Quadro 12. Além disso, foram notadas características comuns nas organizações, como é possível observar pelo Quadro 21.

Quadro 21 – Características comuns das organizações da pesquisa em campo.

1	Todos os casos organizacionais identificados são empresas grandes, complexas, globais, tradicionais e líderes em seus segmentos de negócios;
2	São empresas centenárias ou que apresentaram rápido crescimento no passado próximo;
3	Possuem portfólio de produtos ou de serviços complexos, difusos e dispersos dentre subsidiárias;
4	Possuem estruturas organizacionais complexas e equipes globalmente dispersas, com baixa interação;
5	Possuem faturamento acima de US\$10 bilhões.

Fonte: elaborado pela autora

As organizações contempladas nos estudos de casos são globais, grandes, complexas e apresentaram faturamento acima de US\$ 10 bilhões. São empresas

centenárias que contemplam a transformação digital como elemento estratégico para seus negócios. Os portfólios de serviços ou produtos são complexos e heterogêneos. As estruturas de governança de dados ou de informação podem estar localizadas nos escritórios-sede ou estarem dispersos em subsidiárias.

Dentre os objetivos gerais identificados pelas estruturas de governança de dados ou de informação avaliadas, o Quadro 22 lista os principais itens observados. Na sequência, o Quadro 23 resume as estruturas organizacionais dos casos contemplados, assim como evidencia a abrangência da governança de dados ou de informação.

Quadro 22 – Objetivos gerais identificados pelas estruturas de governança de dados ou de informação.

Caso	Objetivos gerais da governança de dados ou de informação
A	<ul style="list-style-type: none"> • Estabelecer processos eficientes para melhor definir e gerenciar dados; • Definir e gerenciar meta-dados para a cadeia de suprimentos global da organização;
B	<ul style="list-style-type: none"> • Contribuir para o gerenciamento dos meta-dados na área de suprimentos de um dos segmentos de negócios; • Estabelecer denominações comuns e assegurar a singularidade dos dados de clientes, fornecedores, pagamentos, dentre outros temas da área de suprimentos;
C	<ul style="list-style-type: none"> • Harmonizar e padronizar os dados entre os diferentes <i>stakeholders</i> dos negócios globais de um segmento de negócio (estabelecer meta-dados); • Definir padrões nos dados e certificar se tais padrões foram adotados;
D	<ul style="list-style-type: none"> • Transformar a cultura organizacional para a perspectiva <i>data-driven</i>; • Explorar o gerenciamento dos dados (incluindo meta-dados), aumentar e assegurar a qualidade dos dados; contribuir para a monetização dos dados através de seu uso em análises avançadas;
E	<ul style="list-style-type: none"> • Assegurar que os dados e as informações sejam precisos e que possuam alta qualidade, permitindo a extração de valor e uso nos negócios; • Definir papéis, responsabilidades e processos com relação aos dados e as informações na organização; • Transformar a cultura da organização com relação ao uso dos dados (<i>data-driven</i>);
F	<ul style="list-style-type: none"> • Contemplar políticas, métricas e ferramentas de governança de dados e qualidade de dados para as áreas do banco; • definir, gerenciar e controlar meta-dados; • definir papéis e responsabilidades com relação aos ativos de dados; atuar sobre dados sensíveis na organização; • promover a transformação cultural da organização no sentido <i>data-driven</i>;
G	<ul style="list-style-type: none"> • Estabelecer responsáveis para a governança de dados na subsidiária brasileira; • Implementar ferramentas para controlar, monitorar e gerenciar a qualidade dos dados; • Atuar sobre a definição de métricas e parâmetros para a promoção da integridade, correte e veracidade dos dados; • Definir e controlar meta-dados; • Suportar as unidades de negócio para a qualidade e gestão de dados;
H	<ul style="list-style-type: none"> • Definir os princípios na organização para a governança de informação e para o gerenciamento de dados (meta-dados) na perspectiva global dos negócios da organização; • Definir responsabilidades com relação aos ativos de dados; • Atender aos requisitos regulatórios e de <i>compliance</i>; • Promover a transformação cultural da organização com relação aos dados (<i>data-driven</i>).

Fonte: elaborado pela autora

Quadro 23 – Estrutura organizacional e abrangência das estruturas de governança de dados ou de informação.

Caso	Estrutura	Abrangência
A	<ul style="list-style-type: none"> Equipe com líder para o gerenciamento de meta-dados para a cadeia de suprimentos da organização; Estrutura alocada em suprimentos (negócios), reportando ao CFO; Existem outras equipes envolvidas em governança de dados ou de informação para outros segmentos de negócios; 	<ul style="list-style-type: none"> Global para o tema cadeia de suprimentos da organização; Existem outras equipes de governança de dados ou de informação na organização
B	<ul style="list-style-type: none"> Equipe com líder da governança de dados para a cadeia de suprimentos de um segmento de produtos; Alocada em suprimentos (negócios), que reporta ao CFO; Existem outras equipes envolvidas em governança de dados ou de informação para outro segmento de negócio; 	<ul style="list-style-type: none"> Global para uma unidade de negócio e para o tema cadeia de suprimentos; Existem outras equipes de governança de dados ou de informação na organização;
C	<ul style="list-style-type: none"> Equipe com líder da governança de dados, gerente de meta-dados para a cadeia de suprimentos de um segmento de produtos; Alocada em um segmento de negócios específico na organização; Foco em harmonização e padronização das informações, aspectos de qualidade e aprimoramento de processos; Existem outras equipes envolvidas em governança de dados ou de informação para outro segmento de negócio; Comitê de dados 	<ul style="list-style-type: none"> Global para uma unidade de negócio Existem outras equipes de governança de dados ou de informação na organização;
D	<ul style="list-style-type: none"> Estrutura com <i>Chief Data Officer</i> que responde ao CFO da organização; Contempla tópicos de governança de dados e de <i>analytics</i>; comitê de dados 	<ul style="list-style-type: none"> Equipe interna e centralizada que trabalha para todas as partes da organização;
E	<ul style="list-style-type: none"> Estrutura descentralizada com Líder em Dados e <i>Analytics</i>, reportando-se ao CFO da organização; <i>Data owners</i> e <i>data stewards</i> pertencem às unidades de negócios nos diferentes países, não sendo parte formal da estrutura de governança de dados; Comitê de dados 	<ul style="list-style-type: none"> Global para todas as partes da organização; Contempla assunto de <i>analytics</i> e trabalha em conjunto com departamento de BI;
F	<ul style="list-style-type: none"> <i>Chief Data Officer</i> com duas superintendências: governança e qualidade de dados; desenvolvimento de modelos e algoritmos; Comitê de dados 	<ul style="list-style-type: none"> Global para todas as partes da organização;
G	<ul style="list-style-type: none"> Qualidade de dados subordinada ao CFO Gestão de dados subordinada ao COO, sob responsabilidade da TI; Comitê de dados 	<ul style="list-style-type: none"> Para todas os negócios na subsidiária brasileira
H	<ul style="list-style-type: none"> Governança de Informação alocada no departamento de TI; reporta-se ao CIO; Líderes da governança de informação alocados na Europa e contemplam atividades a nível global; Comitê de dados. 	<ul style="list-style-type: none"> Atuação global para todos os segmentos de negócios e produtos; Atende projetos sob demanda de algum negócio em específico.

Fonte: elaborado pela autora

Com base nas características dos casos apresentadas pelo Quadro 22, foi possível identificar que a governança de dados ou de informação se estabelece em organizações grandes, complexas, globais, com portfólio de produtos ou serviços complexos, difusos, com inúmeras estruturas, departamentos ou unidades de negócios. As organizações com baixa complexidade organizacional e aquelas com pequeno a médio porte não foram notadas adeptas às estruturas formais de governança de dados ou de informação. Isto está em linha ao comentado por líderes de estruturas de governança de dados ou de informação.

“A empresa E é muito grande. Acredito que os programas de governança de dados ou de informação são válidos para grandes organizações e não para pequenas ou médias organizações” (líder da governança da informação da empresa E).

“A adoção da governança de dados faz sentido somente para organizações a partir de um determinado tamanho e nível de complexidade. Então, tamanho e complexidade organizacional são variáveis chave para determinar se faz sentido ou não adotar a estrutura de governança de dados” (líder da governança de meta-dados da empresa A).

Assim, pode-se sugerir como um fator que determina a adoção dos modelos formais de governança de dados ou de informação:

→ Organizações grandes, globais, difusas, com estruturas descentralizadas de negócios e portfolio complexo de produtos ou serviços.

Com vistas aos casos observados, o departamento de TI corresponde a um agente que efetivamente atua na gestão de dados, contemplando atividades de extração de dados, manipulação, integração de bases, adequação de dados nas bases, dentre outros pontos operacionais. Além disto, foi observado que organizações cuja estrutura de governança de dados ou de informação é desvinculada do departamento de TI, as atividades deste departamento restringem-se às atividades operacionais acima citadas com relação aos dados. No entanto, foi observado dentre os casos G e H que parte da estrutura de governança de dados ou de informação, ou toda a estrutura, está alocada no departamento de TI. Neste caso, o

departamento de TI é engajado em temas estratégicos com relação aos dados e as informações, cuja atuação contempla também a definição dos princípios dos dados. Para os casos A, B, C, D, E, F, o departamento de TI contempla competências técnicas específicas que suportam as atividades nas bases de dados, como extração de dados, modificação, integração, dentre outras atividades. Neste caso, o departamento de TI é visto como um grupo de apoio às atividades da estrutura de governança de dados ou de informação. Para estes casos, cabe ao departamento de TI implementar as mudanças definidas pelo grupo de governança de dados ou de informação, assim como auxiliá-los nas questões de extração dos dados das bases de dados. Como mencionado pelo CDO da empresa D, *“TI é crucial para a governança de dados, mas atua como um viabilizador”*. Logo, pode-se sugerir que a atuação do departamento de TI é imprescindível para o suporte das atividades operacionais com relação aos dados e as informações estabelecidas pela governança de dados ou de informação.

→ Atuação do departamento de TI nas atividades de gerenciamento de dados ou de informação, viabilizando e executando atividades operacionais com relação aos dados e as informações dentre as bases de dados e sistemas de informação.

Para o caso G, parte da estrutura de governança de dados na subsidiária brasileira está alocada no departamento de TI, com os *“data stewards”*. Outra parte da estrutura de governança, relacionada à qualidade de dados, está alocada no departamento financeiro. Já para a empresa H, a estrutura de governança de informação está subordinada ao departamento de TI. Para o caso H, o departamento pode ser dividido em dois grupos: um relacionado às tecnologias propriamente ditas e o outro focalizado nos negócios, que é onde o grupo de governança de informação se encaixa.

Esta análise traz à luz que agentes do departamento de TI podem fazer parte da estrutura de governança de dados ou de informação na organização. Não necessariamente todos os profissionais do departamento de TI devem estar submetidos à governança de TI, ou unicamente à governança de TI. Quando Kooper et al. (2011) menciona que a governança de dados ou de informação almeja se encaixar nas limitações da governança de TI (pois a governança de TI justamente não se preocupa com os dados e as informações), não se pode inferir

que os agentes do departamento de TI não possam fazer parte da governança de dados ou de informação.

A literatura se preocupa em discutir a correlação da governança de TI com a governança de dados ou de informação. No entanto, com vistas às evidências práticas, é pertinente a investigação preliminar das estruturas corporativas que envolvem o departamento de TI e a estrutura de governança de dados ou de informação. Com base nesta investigação, pode-se melhor compreender a forma com que tais estruturas se interagem e colaboram no âmbito dos modelos de governança de TI, de dados ou de informação.

Do ponto de vista das unidades de negócios, a estrutura de governança de dados ou de informação pode ser estabelecida em uma unidade de negócios específica ou ser independente de uma área de negócio e atender múltiplas unidades de negócio. Observa-se que o desenho organizacional com relação às estruturas de governança de dados ou de informação apresentam diferenças e encontram-se em fase de consolidação.

As estruturas de governança de dados ou de informação são orientadas ao negócio (*business-driven*). São focadas em dados mas voltadas aos negócios. Aos casos cujo departamento de TI englobam também estruturas ou *stakeholders* de governança de dados ou de informação, a parcela do departamento de TI responsável por tais atividades é também *business-driven*, embora o departamento de TI seja voltado para a tecnologia (*technology-driven*).

Além disto, os agentes da estrutura de governança de dados ou de informação no ambiente *big data* atuam em prol de mudanças na organização a fim de torná-las “*data-driven*”, validando a relevância da sub-prática listada pelo Quadro 9. A transformação cultural da organização para o *data-driven* foi notada central dentre os objetivos dos casos A, B, C, D, E, F, G, H. Isto pode ser observado pelo comentário do *Chief Data Officer* da empresa D:

“Antes de investir em arquitetura, definição de papéis e processos, o que realmente importa é mudar a mentalidade na organização [...] o maior desafio é mudar a mentalidade das pessoas e da organização para transformar a organização em data-driven”, aponta o Chief Data Officer da empresa D.

Os *stakeholders* da governança de dados ou de informação atuam ativamente na transformação cultural da organização pois dependem disto para que suas políticas,

práticas e procedimentos com relação aos dados e as informações sejam aceitos e adotados com mais facilidade pelas diversas equipes no ambiente organizacional. Assim, a atuação ativa na mudança cultural corresponde a um fator que determina na adoção da governança de dados ou de informação no ambiente *big data*. Este elemento está em linha ao apresentado por autores do campo do *big data*, como McAfee e Brynjolfsson (2012), Ross et al. (2013) e por Tallon et al. (2014) do campo da governança de TI, de dados e de informação. A cultura baseada em evidência ou *data-driven* correspondeu a um ponto de convergência entre as literaturas de extração de valor de *big data* e governança de dados ou de informação. Tal constatação sustenta a tese de revisar ambas as literaturas para responder à questão de pesquisa.

→ A atuação ativa na transformação cultural da organização para *data-driven* corresponde a um fator que determina a aceitação e adoção das políticas, práticas e procedimentos definidos para governar os dados e as informações no contexto organizacional.

Retomando o modelo de Khatri e Brown (2010), notou-se nos casos que as estruturas de governança de dados ou de informação são governadas por princípios, como previsto pela teoria. Como propõem Khatri e Brown (2010), cabe ao domínio “princípios dos dados” definir os objetivos estratégicos da estrutura na organização e, em particular, com relação à divisão ou unidade de negócio ao qual é vinculada. Neste domínio são definidas a estrutura organizacional para governar os dados ou as informações, as responsabilidades com relação aos ativos de dados e as medidas para atendimento às regulações setoriais. Tais aspectos foram notados em todos os casos organizacionais contemplados.

A regulação setorial corresponde a um parâmetro externo que influencia no processo de adoção das governanças de dados ou de informação. Ao mesmo tempo que corresponde a uma pressão do ambiente competitivo, que pode ameaçar o desempenho da corporação, ela positivamente influencia na adoção do modelo de governança de dados ou de informação pela organização. Os *stekeholders* possuem maior aceitação e maior nível de adequação às práticas ou exigências estipuladas pela estrutura de governança de dados ou de informação caso a justificativa ou o motivo sejam oriundos da adequação das questões regulatórias.

“Cada país possui uma perspectiva diferente com relação à regulação no mercado de seguros. Podemos estabelecer um framework de governança de informação, mas cada país terá que adequar para sua realidade”, comenta o líder da governança de informação da empresa E.

“Reguladores do sistema financeiro podem requisitar mapas com os fluxos dos dados entre aplicações do banco, então trabalhamos nisso”, comenta o líder da governança de dados da empresa F.

“Meu melhor parceiro é o departamento jurídico pois eles são sensíveis com relação aos aspectos regulatórios e ao compliance”, cita a líder da governança de informação da empresa H.

A regulação setorial está cada vez mais severa e corresponde a uma preocupação constante para as organizações, particularmente as do setor financeiro, farmacêutico e de biotecnologia. No caso europeu, a regulação GDPR define princípios de confidencialidade dos dados, segurança e ética, cujos setores financeiro e de saúde estão diretamente relacionados. Além disto, uma série de particularidades são eminentes nas regulações destes setores.

A equipe de governança de dados ou de informação monitora continuamente as exigências regulatórias com o apoio do departamento jurídico da organização. Trata-se de um aporte sinérgico com relação às equipes de suporte ao negócio, como é o caso do jurídico, RH, com os *stakeholders* de negócios. Isto está em linha ao modelo apresentado por Cabral Netto (2010), o qual denota o monitoramento do ambiente externo como uma área de decisão. Em específico, o caso 1 (da pesquisa piloto) indicou que o nome antigo da estrutura de governança de dados na organização era “grupo regulatório”. Logo, a governança de dados ou de informação é orientada para o atendimento dos requisitos regulatórios setoriais.

A regulação setorial foi citada por Tallon et al. (2014) como um fator antecedente e viabilizador para a adoção da governança de dados ou de informação (vide Quadro 2). Trata-se de uma força externa que motiva a adoção formal da estrutura.

→ A regulação setorial foi notada como uma força do ambiente competitivo que motiva e viabiliza a adoção da governança de dados ou de informação.

Demais pontos evidenciados nos casos incluem a necessidade de comunicação e colaboração entre diferentes *stakeholders* da organização para compreender pontos críticos com relação ao fluxo dos dados, divulgar as recomendações a serem seguidas com relação à estratégia dos dados ou informações e a necessidade,

gerenciar potenciais conflitos entre equipes. No âmbito da colaboração e da comunicação almeja-se promover a sinergia entre os diferentes *stakeholders* da cadeia produtiva e servir de *stakeholders* focal para os assuntos relativos aos dados e as informações. A necessidade dos esforços de colaboração e comunicação para garantir o bom funcionamento do modelo de governança foi previsto pelo referencial de governança de TI nos trabalhos de Weill e Woodham (2002), Rau (2004), Weill e Ross (2005), Tallon et al. (2014), Nfuka e Rusu (2011) e Kurti et al. (2014). Como dimensão deste campo pode ser atribuída a capacidade de gerenciamento de conflitos, como sugerido por Koltay (2016).

“Como estrutura de governança de dados, nós temos em mente a estratégia com relação aos dados e as informações em nossa unidade de negócios. Nós temos a visão geral das diferentes necessidades que existem em nossa unidade de negócios. E nós só temos essa visão geral pois conversamos com todo mundo”, comenta o gerente de governança de dados da empresa C.

“Um CDO é inútil sem o apoio de outros gerentes (da alta gestão)”, cita o Chief Data Officer da empresa D.

Além disto, foram ressaltados os aspectos do gerenciamento dos dados (e das informações) e o gerenciamento da qualidade dos dados (e das informações) nos casos contemplados. Como propõem Khatri e Brown (2010), a qualidade pode ser tratada como um domínio do modelo de governança e assume um dos objetivos centrais da governança de dados ou de informação. Por meio do gerenciamento da qualidade, os dados desqualificados, repetidos, incompletos, desnecessários, ultrapassados ou sem valor podem ser removidos ou impedidos de serem armazenados nos sistemas de informação.

No campo do gerenciamento dos dados ou das informações, podem ser contempladas as atividades com relação aos meta-dados, acessibilidade, ciclo de vida da informação em um domínio único do modelo, como propõem a Figura 23.

Dentre as contribuições ao modelo de Khatri e Brown (2010) incluem:

- Agrupamento das atividades com relação aos meta-dados, ciclo de vida dos dados ou das informações e definição da acessibilidade dos dados em um domínio único intitulado por “gerenciamento dos dados”;

- Adição de um domínio exclusivo com relação à mudança cultura da organização em direção ao *data-driven*;
- Adição de um domínio único e exclusivo com relação à mudança cultural da organização.

Figura 23 – Contribuições ao modelo de Khatri e Brown (2010) para a governança de dados ou de informação. Fonte: elaborado pela autora

No âmbito dos modelos teóricos de governança de dados, a organização G apresentou o domínio qualidade de dados desmembrado do domínio gerenciamento de dados na subsidiária brasileira. Isto é, as atividades de qualidade de dados estão sob um CFO e as atividades de gestão de dados sob um COO, alocadas no departamento de TI. No entanto, o escritório sede da organização estabeleceu uma estrutura única para a Governança de Dados, contemplando profissionais da qualidade dos dados e para o gerenciamento de dados, com os *data stewards*. Apesar das diferenças estruturais na subsidiária, a colaboração entre as equipes de gestão de dados e qualidade dos dados é intensa e diária.

No caso F, a qualidade de dados também adquiriu um domínio a parte do modelo de governança de dados para ter pessoal dedicado no estabelecimento de métricas de qualidade e desenvolvimento de ferramentas para mensurar a qualidade das bases de dados. No entanto, governança de dados e qualidade de dados corresponderam a dois domínios sob responsabilidade de uma única liderança.

A questão sobre a qualidade dos dados apresentou alta relevância dentre os casos organizacionais e está em linha à discussão dos autores Weber et al. (2009), Khatri

e Brown (2010), Otto et al. (2011), Otto (2011), Fu et al. (2011) e Tallon et al. (2014). Aprimorar a qualidade dos dados permite constituir ativos de dados mais íntegros, completos e precisos, tornando as análises mais assertivas do conjunto de dados e de informações.

Quando se tem estrutura de governança de dados, assegurar a qualidade dos dados está no topo das prioridades deste grupo. A determinação de um *stakeholder* específico para atuar sobre a qualidade dos dados é relevante para o trabalho contínuo no tema. Como apontado pelo caso G, o modelo de governança de dados na subsidiária brasileira focalizou em primeira instância no monitoramento e no controle da qualidade dos dados. Em segunda instância foram determinados agentes para o gerenciamento de dados, como os *data stewards*, no departamento de TI.

A perspectiva que o tema qualidade dos dados adquire no modelo de governança de dados ou de informação pode ser observada pelos comentários dos líderes dos casos C e D:

“O foco sobre a qualidade dos dados é extremamente alta na indústria farmacêutica. Lidamos com a saúde das pessoas. Temos que assegurar a conformidade em nossos produtos”, comenta o gerente da governança de dados da empresa C.

“A qualidade dos dados é um desafio para a organização, pois os dados são inúteis se eles não forem confiáveis ou bons”, como comentado pelo *Chief Data Officer* da empresa D.

Além disto, somente a empresa H mencionou o benefício da redução dos custos de armazenamento dos dados ao aprimorar a qualidade dos dados, já que dados incompletos podem ser descartados das bases de dados. As demais organizações focalizaram nos benefícios para os relatórios de BI e a possibilidade de desenvolver técnicas avançadas de *analytics*. Isto está em linha ao apresentado por Berkooz (2017), *“em analytics, nada importa mais que qualidade”*.

Dentre os casos contemplados, não foram enfatizadas as atividades com relação ao ciclo de vida das informações, como preveem Khatri e Brown (2010) e como consta no Quadro 9. A partir dos casos, foi notado que a análise do ciclo de vida da informação não necessariamente é uma atividade da governança de dados ou de informação. Segundo a empresa A, trata-se de uma atividade conduzida por

stakeholders próximos ao departamento jurídico da organização e contemplam aspectos regulatórios e o trato com dados sensíveis. Segundo a empresa A, este grupo classifica os dados, define como melhor gerenciá-los ao longo do ciclo de vida da informação, observa se estão em conformidade com os diferentes reguladores. Já a governança de dados da empresa C contempla o tópico de ciclo de vida do dados ou da informação de forma sistêmica, contemplando os processos para serem criados, alterados e desativados os dados. Assim, pode-se sugerir a análise do ciclo de vida da informação como um sub-grupo do gerenciamento de dados, embora não tenha apresentado relevância dentre os casos.

Além das questões do gerenciamento da qualidade, o gerenciamento de metadados apresentou também alta relevância nos casos organizacionais. Demais atividades operacionais mencionadas nas entrevistas incluem:

- Definir, gerenciar e controlar meta-dados, incluindo parâmetros, métricas e siglas de referência para uso pelas diversas equipes de um determinado negócio (empresas A, B, C, E, F, G, H)
- Analisar criticamente os processos organizacionais e os fluxos dos dados entre diferentes *stakeholders* para identificar possíveis melhorias para o gerenciamento dos dados (empresa E, F, G, H)
- Identificar gargalos operacionais sob a perspectiva dos dados envolvendo diferentes *stakeholders* de uma cadeia produtiva (pode-se focar na cadeia de suprimentos de um segmento de produtos) (empresa A, B, C, H)
- Elaborar propostas de melhoria das bases de dados e definir planos de trabalho para a correção dos problemas identificados (empresa E, H);
- Foco no gerenciamento de dados para suportar as áreas de risco da organização (empresa F, G);
- Reduzir a redundância das informações, remover dados desqualificados, reduzindo gastos operacionais com bases de dados desnecessárias (empresa F, G, H);
- Atualizar as equipes de negócios sobre as adequações necessárias no *modus operandi* para atender exigências regulatórias (empresa A, B, C, D, E, F, G, H).

Com vistas ao observado pelo caso G, os modelos de governança de dados ou de informação estão em desenvolvimento, podendo haver diferenças entre os escritórios-sede e as subsidiárias. A estrutura de governança de dados ou de informação pode até mesmo estar restrita aos escritórios sede, como observado pelas empresas B, C, E.

4.2 Elementos e conteúdos da governança de dados ou de informação

A Figura 24 esquematiza a estrutura organizacional dos casos estudados com relação à governança de dados ou de informação.

Figura 24 – Estrutura organizacional dos casos com relação à governança de dados ou de informação. Fonte: elaborado pela autora

Os casos A e B contemplam estruturas de governança de dados sobre os aspectos do gerenciamento da cadeia de suprimentos ou da área de *procurement* de uma

dada unidade de negócios ou segmento corporativo. A diferença destas estruturas para o caso C é que nesta terceira organização, a estrutura de governança de dados está vinculada a uma unidade de negócios e não se restringe ao campo da gestão da cadeia de suprimentos ou aquisições.

Nas organizações A, B e C podem existir múltiplas estruturas de governança de dados, já que tal estrutura se estabelece por divisão negócio ou um domínio específico. Tal constatação está em linha ao apresentado por Donaldson e Walker (2004), que indicam que as governanças de dados ou de informação focalizam porções discretas de informação na corporação, não abrangendo todo o conteúdo organizacional. Trata-se de um desafio estabelecer estruturas de governança de dados ou de informação com alta abrangência organizacional, pois existem diversas particularidades do negócio que a estrutura deve contemplar.

No âmbito das organizações do sistema financeiro, foi notada a participação do CFO, CRO e COO na liderança das atividades de governança de dados ou de informação. Tanto o caso D quanto o F possuem dois níveis de *C-level*: um CDO se reportando ao CFO no caso D e ao CRO no caso F. A diferença destas organizações para o caso E é a inexistência da posição de *Chief Data Officer*, embora exista o gerente geral de governança de informação. Complementarmente, o caso G foi o único cujo modelo de governança de dados foi contemplado em duas estruturas organizacionais distintas: uma sob o CFO e outra sob o COO, através de *stakeholders* do departamento de TI.

A importância de um CDO para a organização pode ser retratada pelo comentário de um líder de governança de dados. Esta posição exerce influência positiva para empoderar as atividades de governança de dados ou de informação:

“Quando se tem um CDO, provavelmente também se tem uma política forte de gerenciamento de dados em toda a organização. A organização também sabe que tem que trabalhar em novo modo para fornecer dados com boa qualidade para suportar o gerenciamento de dados”, menciona o líder da governança de dados da empresa F.

No caso E, a liderança da governança de informação se reporta diretamente ao CFO da organização e pode atender todas as unidades de negócios, sem limitação de tema ou localidade. Trata-se de uma posição global, cujo foco se estabelece nas regiões cuja organização possui maior rentabilidade de seus negócios. No caso,

focaliza seus esforços em três países europeus. Como comentado pelo entrevistado:

“O cenário é bastante descentralizado na organização [...] Para organizações centralizadas, o grupo de governança possui bastante poder do nível corporativo e pode ser mais incisivo nas medidas: de agora em diante, as regras A, B, C devem ser seguidas. No caso das organizações descentralizadas, é necessária muita negociação entre as partes. Tem que se reunir com as partes, sentar, conversar e discutir com as diferentes equipes”, como menciona o líder da governança de informação da empresa E.

O caso H contempla uma estrutura de governança de informação alocada em um departamento de TI, reportando-se ao CIO. O caso retratou intensa necessidade das habilidades relacionais, contemplando gerenciamento de conflitos constantes, comunicação e necessidade de colaboração dos diferentes *stakeholders* da organização. Apesar disto, a líder mencionou que possui boa parceria com os departamentos de finanças e jurídico, recorrendo a eles quando necessita de ajuda para o convencimento interno com relação às medidas de governança de informação.

Foram observados que os casos A, B, C, D e G contemplam estruturas de governança de dados ou de informação sob liderança do CFO da organização. A relevância do CFO na organização foi prevista por Demarquet (2016), o qual cita a migração do papel com foco em contabilidade para um agente digital. Especialmente no ambiente *big data*, o CFO se torna mais estratégico para apoiar e executar a transformação da organização.

Apesar da possibilidade de existir diferentes equipes de governança de dados ou de informação em uma organização, foi notada baixa colaboração e comunicação entre elas. A atuação de uma determinada equipe acaba sendo delimitada ao negócio ao qual é vinculada, possuindo fraca troca de experiências ou lições aprendidas com relação à definição de políticas e práticas no âmbito organizacional.

Os casos D, G, E, F H reportaram maiores conteúdos relacionais que os demais casos. Isto pode ser oriundo do fato destas estruturas de governança de dados ou de informação não serem vinculados a um segmento de negócio específico. Pelo fato de poderem contribuir com diferentes unidades de negócios e grupos

funcionais, foi reportada a necessidade de maiores competências de gestão de conflitos, comunicação e interação.

As práticas relacionais do modelo de governança de dados ou de informação são tão mais relevantes quanto mais desvinculada é a estrutura de governança de dados ou de informação do negócio e quanto menor for o poder da alta gestão que suporta a governança de dados ou de informação. Como cita o entrevistado da empresa F:

“Especialmente se o grupo por trás da governança de dados não for visível para as unidades de negócios, ou se não tiver um Chief Data Officer, por exemplo, assegurando poder para a organização, então a estrutura de governança sofre para ter as orientações adotadas na organização”, gerente de governança de dados da empresa F.

Com vistas aos demais *stakeholders* da governança de dados ou de informação, foi observado que os *data owners* e os *data stewards* não necessariamente estão subordinados à estrutura de governança de dados ou de informação. Como visto pelas empresas C e E, tais *stakeholders* pertencem às unidades de negócios distribuídas entre os diferentes países, não sendo subordinados à liderança de governança de dados ou de informação. No entanto, a colaboração destes *stakeholders* com a equipe global de governança de dados ou de informação é fundamental para que as estratégias com relação aos dados e as informações possam ser adaptadas para as realidades de cada ambiente do negócio e da regulação ao qual está inserida.

“Nós como governança de informação estamos operando no alto nível [...] Não queremos tocar nos ativos de cada país. Os dados estão sob responsabilidade dos data owners de cada país. Nós focamos em estabelecer o modelo operacional, que deve ser adaptado para cada mercado”, cita o líder da governança de informação da empresa E.

“Os data stewards são experts nas áreas de negócios que compreendem as especificidades de sua área. Nós como governança de dados precisamos do suporte deles para assegurar que os processos definidos estão em conformidade com as especificidades de sua área”, menciona o gerente de governança de dados da empresa C.

Aos “*data stewards*” do caso G cabem as seguintes responsabilidades:

- Analisar, compreender e aprovar as mudanças necessárias quanto aos requerimentos de novos dados, requerimentos de monitoramento da qualidade dos dados, novos sistemas;
- Participar de projetos e assegurar que os dados serão disponibilizados;
- Gerenciar a completude e integridade dos dados;
- Executar correção nos dados nas bases de dados.

O caso G foi o único que contemplou a prática da governança de dados sob a perspectiva de uma subsidiária. A diferença do modelo estrutural da governança de dados entre o escritório sede e a subsidiária brasileira pode ser eminente da diferença na quantidade de funcionários e de volume das operações. Em geral, os funcionários na subsidiária brasileira contemplam uma gama maior de atividades que no escritório sede. Com vistas a este fato, profissionais de TI acumularam as funções de *data stewards* para suportar a gestão de dados.

Com vistas à primeira sub-prática operacional listada no Quadro 9, que diz respeito à criação dos direitos e das responsabilidades de "*data ownership*" e "*data stewardship*" (MALAKAR, 2007; ROSENBAUM, 2010; TALLON, 2013; TALLON et al., 2014; THOMPSON et al., 2015; KOLTAY, 2016), foi possível sugerir as seguintes adaptações com base nas evidências práticas coletadas:

- Deve-se estabelecer uma estrutura independente e singular na unidade de negócio ou na organização com relação à definição das políticas de governança de dados ou de informação;
- Esta estrutura deve ter poder e visibilidade na organização, o que torna imprescindível a figura de um *C-level* suportando as iniciativas da governança de dados ou de informação;
- A criação de *steward* e *owners* não garante o bom funcionamento do modelo de governança de dados ou de informação; além da liderança estratégica, devem ser contemplados gerentes para conduzir as atividades de gestão dos dados e das informações, promoção da transformação cultural da organização, identificação de pontos críticos e gargalos operacionais com relação ao fluxo dos dados.

Com base no acima exposto, definir um *C-level* e criar direitos e responsabilidades com relação aos dados, incluindo gerentes da governança de dados ou de informação, *data owners* e *data stewards* são necessários para que o modelo de governança de dados ou de informação possa ser adotado na organização. Esta consideração está em consonância aos trabalhos de Malakar (2007), Rosembaum (2010), Tallon (2013), Tallon et al. (2014), Thompson et al. (2015), Demarquet (2016) e aos casos contemplados nesta pesquisa. Assim, sugere-se como fator determinante para a adoção das governanças de dados ou de informação:

➔ **Apontar um *C-level* responsável pela governança de dados ou de informação, definir gerentes na estrutura para conduzir as atividades com relação aos dados e as informações, determinar *data owners* e *data stewards*.**

Além disto, os comitês de dados apresentaram relevância dentre os casos organizacionais. Malakar (2007) tratou o comitê de dados como o corpo central das decisões da governança de dados ou de informação e isto foi observado nos casos. Os comitês são utilizados para “colocar todos os líderes dos negócios na mesma página”.

“Nós criamos um conselho e colocamos pessoas chave ao redor da mesa. Apresentamos a estratégia e o plano de negócio”, comenta o líder da governança da informação da empresa E.

Em particular, a empresa F menciona que o *board* da governança de dados contempla a alta cúpula da organização pois eles reconhecem o valor da boa governança de dados para a adequação aos requisitos regulatórios e de *compliance*. A empresa G reportou o uso dos comitês de qualidade para promover o engajamento dos líderes das unidades de negócios do banco na adoção das ferramentas de qualidade dos dados.

Os comitês de dados servem como um ambiente para a comunicação e a colaboração entre os diferentes *stakeholders* da organização, permite engajar as unidades de negócios na definição dos princípios de governança, encorajar as mudanças culturais na organização, servir como espaço para o gerenciamento de conflitos ou identificação de oportunidades de melhorias nos fluxos de dados. E tais conteúdos estão em consonância aos trabalhos de Rau (2004), Weill e Ross (2005),

Malakar (2007), Weber et al. (2009), Otto (2011), Tallon (2013), Tallon et al (2014). Apesar de Rau (2004), Weill e Ross (2005) se referirem à governança de TI, o conceito é válido para o modelo de governança de dados ou de informação, como discutem Weber et al. (2009) e Otto (2011) no campo da governança de dados.

Com vistas às evidências coletadas, sugere-se também como um fator que determina a adoção da governança de dados ou de informação:

→ Estabelecer um comitê de dados ou outros meios para reunir a alta cúpula e os principais líderes da organização para definir as estratégias e as ações com relação aos dados e as informações na organização.

Partindo-se para as práticas relacionais, estas apresentaram significativa parcela dos resultados coletados em campo. Conforme mencionado pelo líder da governança de dados da empresa F, as organizações Fortune 500 possuem um ambiente interno de comunicação e intensas trocas das melhores práticas. E, portanto, contemplam ambientes mais propícios para o estabelecimento da governança de dados. Observa-se, portanto, a referência de práticas relacionais como base para o estabelecimento da governança de dados.

Logo, as práticas relacionais devem ser contempladas de forma explícita nos modelos de governança de dados ou de informação, como prevê a Figura 23 ao introduzir os domínios relativos à comunicação, colaboração e mudança cultural. É através da comunicação formal e da colaboração que podem ser compreendidos os cenários corporativos, os pontos críticos, as oportunidades de melhoria. Sem tal domínio relacional, não há elo entre os princípios dos dados definidos e a entrega dos resultados para a organização. Isto justifica a classificação deste campo como um domínio do modelo de governança de dados ou de informação. Além disto, sugere-se como dimensão deste domínio as habilidades para gerenciamento de conflitos.

Na literatura de governança de TI, Weill e Woodham (2002), Weill e Ross (2005) consideram a comunicação formal como um mecanismo para a governança de TI. Tallon et al. (2014) contemplam o relacionamento, a colaboração, a troca de ideias, a comunicação e a gestão de conflitos como práticas relacionais do modelo da governança de TI. Além disto, Nfuka e Rusu (2011) e Kurti et al. (2014) sugerem cinco fatores críticos de sucesso (FCS's) baseados em práticas relacionais,

incluindo: o envolvimento e o suporte da alta gestão; o encorajamento e o apoio à comunicação e à parceria entre TI e negócio; o engajamento dos *stakeholders* principais; a conscientização dos indivíduos da organização sobre a governança de TI e o treinamento para o uso ótimo da TI; a consolidação, a comunicação e o fazer cumprir a política e as práticas que propiciam a aquisição e o uso da TI na organização com custo efetivo.

Com vistas às práticas de comunicação e colaboração listadas pelo Quadro 9 e as evidências coletadas em campo, foi possível considerar como fatores relevantes para o modelo de governança de dados ou de informação: promover a colaboração interna para desenho da governança de informação (TALLON et al., 2013); promover a supervisão compartilhada para a formulação de políticas e avaliação (TALLON et al., 2014); desenvolver a comunicação com relação à eficácia das políticas e a necessidades de usuários (TALLON, 2013; TALLON et al., 2014). Assim, sugere-se como fatores determinantes para a adoção das governanças de dados ou de informação:

→ Promover a comunicação e a colaboração interna para definir o modelo de governança e assim identificar pontos críticos no trato com os dados esolucionar os entraves; Desenvolver a comunicação com relação à eficácia das políticas e a necessidade de adequação dos *stakeholders*.

Além disto, a cultura baseada em evidências ou *data-driven* (listada no Quadro 9) foi citada como um fator determinante para o estabelecimento do modelo de governança de dados ou de informação, como apresentado na sessão 1.1. Tal elemento requer, portanto, capacidades para gerenciar mudanças, como aponta Koltay (2012). As questões relativas à educação e aos treinamentos apontadas por Tallon (2013), Tallon et al. (2013) e Tallon et al. (2014) não apresentaram significância dentre os casos organizacionais considerados.

No âmbito das práticas operacionais, as questões burocráticas foram notadas como parte dos princípios dos modelos de governança de dados ou de informação para determinar os níveis, as barreiras necessárias para o gerenciamento dos dados, das responsabilidades, da acessibilidade aos dados ou as informações. Assim, não foi notada a preocupação em eliminar práticas que governam em demasia, como proposto por Tallon (2013). No entanto, foi observada a preocupação em permitir

exceções nas políticas globais para atender as necessidades das políticas locais, como sugerem Tallon et al. (2013), particularmente quando envolve regulação setorial.

“Certo grau de burocracia é necessário devido ao tamanho e a complexidade da organização [...] burocracia não é algo ruim. Somente é ruim se é demasiada e desgovernada. É uma linha tênue entre assegurar o que tem que ser feito e prejudicar o andamento dos trabalhos internos na organização”, menciona o líder do gerenciamento de meta-dados da empresa A.

“Burocracia é gerada, mas porque precisamos colocar níveis de controle. O ponto de haver governança é assegurar que as mudanças são corretamente gerenciadas, que as pessoas possuem a mesma compreensão dos atributos, e isto requer um pouco de burocracia. Isto para evitar que alguém possa ter alguma atitude deliberada que possa prejudicar o fluxo de trabalho de outros departamentos, outros grupos de trabalho”, comenta o líder da governança da empresa C.

“Também tentamos impactar positivamente os negócios através de mini-projetos [...] Assim, os negócios não nos veem como burocracia”, cita o data governance officer do caso 1 (pesquisa piloto)

Dentre as práticas operacionais listadas pelo Quadro 9, o gerenciamento de meta-dados e o gerenciamento da qualidade foram os mais relevantes nos casos estudados. Estabelecer e gerenciar meta-dados foi observado como objetivo estratégico em todas as empresas consultadas (A, B, C, D, E, F, G, H). Apesar de ser uma atividade operacional, os esforços com relação aos meta-dados é um desafio estratégico para a corporação, pois permite ganho de eficiência e transparência no fluxo dos dados pela organização. Além disto, permitem-se adequações com relação às exigências regulatórias, padronização no fluxo dos dados, dentre outros benefícios internos para a organização. Tal atividade está prevista pelos trabalhos de Khatri e Brown (2010), Fu et al. (2011), Otto et al. (2011) e Koltay (2016).

Devido a relevância das atividades com relação aos meta-dados no modelo de governança de dados ou de informação, foi possível considerar como fator determinante:

- ➔ **Definir, gerenciar e controlar meta-dados, incluindo parâmetros, métricas e siglas de referência para uso pelas diversas equipes de um determinado negócio ou pela organização como um todo.**

A qualidade dos dados foi observada como um tópico central no modelo de governança de dados ou de informação, como preveem Weber et al. (2009), Khatri e Brown (2010), Fu et al. (2011), Kooper et al. (2011), Otto (2011), Otto et al. (2011) e Tallon et al. (2014) (vide Quadro 9).

Atuar sobre a qualidade dos ativos dos dados na organização viabiliza laços entre as unidades de negócios e os *stakeholders* financeiro, agregando valor às iniciativas da estrutura perante os negócios. Berkooz (2017) menciona que assegurar a qualidade dos dados determina o valor dos ativos dos dados e Fu et al. (2011) relacionam a baixa qualidade dos dados de posse das organizações como um influenciador negativo no lucro da corporação.

Logo, as ações relativas à qualidade dos dados podem motivar, viabilizar e determinar a adoção das governanças de dados ou de informação, pois correspondem a uma iniciativa de alto valor agregado para a organização e um defensor estratégico da estrutura de governança de dados ou de informação. Ao contemplar os aspectos de qualidade dos dados, a estrutura de governança de dados ou de informação pode explorar as limitações da governança de TI e colocar no topo da prioridade a constituição de melhores ativos de dados e de informações para a organização.

Logo, pode-se sugerir como fator determinante para o modelo de governança de dados ou de informação:

- ➔ **Definir os padrões, as exigências e o controle sobre a qualidade dos dados, assegurando a corretude, integridade, completude e transparência dos dados e das informações; contemplar as atividades de gerenciamento da qualidade dos dados.**

A acessibilidade foi apontada como uma prática operacional comum dentre as estruturas e está em linha ao previsto por Tallon (2013) e Tallon et al. (2014). Ela também pode ser interpretada como os direitos de acesso. Outras sub-práticas que compõem o Quadro 9 e que não exprimiram relevância na esfera estratégica, são: procedimentos de cobrança retroativa, políticas de retenção, auto-destruição ou

arquivamento, alocar a informação em camadas estratificadas, migrar os dados entre camadas de armazenamento.

Com vistas aos pontos apresentados e discutidos nas sessões 4.1 e 4.2, foi possível refutar parcialmente a proposição 2 e validar a proposição 3. Quanto a proposição 2: “as ações organizacionais para governar os dados ou as informações no contexto do *big data* ainda são pouco consolidadas e conhecidas no ambiente organizacional”. Foi notado que as ações são pouco consolidadas, embora elas sejam conhecidas pelas organizações. A dificuldade está na implementação dos modelos e das ações pré-estabelecidas, o que pode culminar com a baixa consolidação das ações notadas dentre os casos organizacionais. Tal desafio também pode ser notado pela relevância que as práticas relacionais assumem no modelo de governança, incluindo: a mudança de mentalidade organizacional, o papel da comunicação e da colaboração entre equipes.

A proposição de pesquisa 3: “apesar das particularidades organizacionais, existem práticas, sub-práticas e fatores comuns que frequentemente são adotados para viabilizar a adoção das governanças de dados ou de informação”, foi validada nesta pesquisa. Por isto, foi possível identificar sete fatores determinantes para a adoção das governanças de dados ou de informação no ambiente big data.

4.3 Capacidades analíticas e a governança de dados ou de informação

A governança de dados ou de informação contempla o fluxo de trabalho com relação aos dados em uma organização ou uma unidade de negócios. O processo de extração de valor é realizado pelas equipes de negócios com foco em BI para extrair conteúdos dos dados, analisá-los e conferir subsídios aos tomadores de decisão. Como comentou um *Chief Data Officer*:

“Os dados são apenas o caminho, apenas o começo. Os dados são importantes somente se a organização conseguir utilizá-lo, explorá-lo e realizar o potencial de valor deles”, comenta o CDO da empresa D.

Em geral, as equipes de BI estão distribuídas na organização dentre diferentes unidades de negócios ou departamentos. Elas almejam desenvolver suas

capacidades analíticas e melhor fundamentar os processos de tomadas de decisão. Como mencionam Bekooz (2017) e Lismont et al (2017), o gerenciamento dos dados e os problemas de qualidade ainda desafiam a *analytics*. Estas atividades são contempladas pelos modelos de governança de dados ou de informação e almejam contribuir que os ativos de dados ou de informação adquiram maiores níveis de qualidade e confiabilidade, viabilizando análises mais assertivas e sustentando melhores tomadas de decisão. Logo, a governança de dados ou de informação atua como pano de fundo do ciclo da inteligência competitiva da Figura 5, podendo contribuir para o desenvolvimento analítico da organização.

“Temos que assegurar a integridade e a consistência dos dados para que os relatórios sejam constituídos com base em informações corretas”, comenta o líder da governança de dados da empresa C.

“A qualidade dos dados determina a qualidade dos relatórios (para tomadas de decisão pela alta gestão)”, cita a líder da governança de informação da empresa H.

Dentre os casos contemplados, somente ao CDO da empresa D cabem as funções de governança de dados e da *advanced analytics*. As funções de governança de dados são necessárias para definir uma base de procedimentos que permitam constituir ativos de dados íntegros, completos, organizados e relevantes. Devido a tal preparo, as funções de *advanced analytics* podem ser desenvolvidas. Vale mencionar que as funções tradicionais de BI estão dispersas na organização, somente cabem ao CDO da empresa D as questões do *advanced analytics*, como forma de testar novas soluções e então alocá-las nas equipes de BI dispersas nas unidades de negócios.

Revisitando a definição adotada para as governanças de dados ou de informação, pode-se definir dois níveis de responsabilidade: o primeiro trata da estrutura organizacional e de suas responsabilidades e, o segundo, das atividades de apoio e de suporte para a garantia da qualidade, confiabilidade, transparência e singularidade dos dados ou das informações na organização. Tais partes foram retratadas através da Figura 25.

Figura 25 – Definição adotada para as governanças de dados ou de informação. Fonte: elaborado pela autora

Na segunda parcela da definição adotada consta: “também envolvem o estabelecimento de um conjunto de competências ou práticas para a avaliação, criação, coleta, análises, distribuição, armazenamento, uso, controle, acesso, arquivamento e exclusão dos dados ou das informações ao longo de seus ciclos de vida” (KOOOPER et al., 2011; KHATRI; BROWN, 2010; TALLON et al., 2014)”. Algumas análises podem ser feitas desta definição ao considerar as competências analíticas da organização:

- As competências ou práticas empregadas na definição acima relacionam-se com a qualidade, segurança, singularidade, transparência, acessibilidade, por exemplo, dos dados ou das informações. Estas não englobam as competências ou práticas analíticas sob reponsabilidade das equipes de BI;
- As competências relativas à criação circundam na esfera dos meta-dados que, em outras palavras, podem ser explicados como dados que representam outros dados. Tratam-se de símbolos ou siglas criados pelo grupo de governança de dados ou de informação que, em geral, são

diretores, gerentes ou analistas de *master data*. O uso de meta-dados almeja a construção de uma biblioteca de dados singulares na organização, permitindo o melhor armazenamento e localização de um conjunto de dados quando necessário;

- Avaliação, análise e uso são termos que aparecem tanto na conceituação da governança de dados ou de informação quanto na de BI. No entanto, no campo da governança, as “análises” se dão em relação aos ativos de dados e, no campo da BI, as análises são eminentes do uso dos ativos de dados para gerar relatórios gerenciais para fundamentar processos futuros de decisão.

Logo, para melhor definir as governanças de dados ou de informação, sugere-se melhor especificar a segunda parte da definição adotada para não confundir com possíveis atributos da inteligência de negócios:

- [...] Também contemplam atividades de apoio e de suporte na garantia da qualidade, confiabilidade, transparência e singularidade dos dados dentre os diferentes grupos da organização, permitindo a constituição de indicadores confiáveis para os processos de tomadas de decisão. Dentre tais atividades de apoio ou suporte, contemplam: a análise e a avaliação dos ativos de dados com relação à qualidade, segurança, singularidade, transparência, acessibilidade; criação de meta-dados com fins de promoção de uma representação única das informações na organização.

Foi notado também que os profissionais abordam indistintamente “governança de dados” ou “governança de informação”. Dentre os casos estudados, somente um profissional trouxe à luz uma possível diferenciação dentre os conceitos, como foi o caso de um líder da prática global de governança de informação em serviços de seguradora. Para ele, o objetivo da governança de informação é assegurar a entrega de informações precisas, seguras e com qualidade para os tomadores de decisão na organização. Já a governança de dados não se compromete com esta entrega, pois contempla a parte de governança de dados, e não de informações válidas para as tomadas de decisões.

Com vistas às empresas consultadas, o Quadro 24 classifica a capacidade analítica com base em LaValle et al. (2011). Dos casos contemplados, cinco deles correspondem a organizações transformadas, uma experiente e uma aspiracional.

Quadro 24 – Classificação das capacidades analíticas das organizações estudadas.

Capacidade analítica	Organizações	Justificativa
Transformada	A	• Organizações experientes em projetos de BI;
	B	• Possuem competências desenvolvidas em termos de pessoas, processos e ferramentas;
	C	• Contemplam aporte significativo para o desenvolvimento de novas soluções;
	D	• Interesse na expansão do uso da BI;
	F	
	G	
Experiente	E	• Organização do setor financeiro;
		• Possui alguns casos de sucesso na fase aspiracional e investimentos em BI;
		• Possui algum tipo de experiência em BI e algumas competências desenvolvidas em termos de pessoas, processos e ferramentas;
		• Almeja a expansão do uso da BI;
Aspiracional	H	• Corresponde a uma organização de bens de consumo.
		• Foco em eficiência; automação; redução de custos; poucas competências desenvolvidas; poucos projetos e exemplos no campo da BI.

Fonte: elaborado pela autora

As organizações farmacêuticas e de biotecnologia (A, B, C) e os bancos (D, F G) apresentaram capacidades analíticas transformadas. Para estas organizações, a BI é fundamental para os negócios e as competências analíticas estão dispersas na organização. Além disto, preocupam-se em extrair valor dos dados no ambiente *big data* e investem em novas competências tecnológicas e gerenciais para o alcance de vantagens competitivas.

Em particular, o caso D apresentou esforços específicos no campo do *big data* ao introduzir um *Big Data Lab* para a inovação analítica na organização e alcance de novas técnicas avançadas para suportar tomada de decisões dentre os negócios do banco. No contexto do banco, as competências tradicionais de BI estão dispersas nas unidades de negócios e suportam o gerenciamento de riscos (de crédito, de mercado), dos ativos e passivos, o desempenho da organização, dentre outros campos, estando em consonância ao discutido por Ubiparipovic e Durkovic (2011). As organizações farmacêuticas contemplam dentre a estratégia corporativa o uso

intensivo de dados para o desenvolvimento de novos produtos e novas terapias para atender as demandas de saúde da população.

O caso E apresentou capacidades analíticas na faixa experiente do modelo de LaValle et al. (2011). Trata-se de uma seguradora, cujos negócios almejam compreender os indivíduos e possivelmente oferecer serviços de seguro personalizados, no presente e possivelmente no breve futuro. Foram observadas limitações analíticas nestas organizações, apresentando alguns casos de sucesso de projetos de BI. Inclusive, a adoção formal da estrutura de governança de dados ou de informação foi motivada para aperfeiçoar os ativos de dados para permitir o avanço analítico da organização.

O caso H foi classificado com capacidade analítica aspiracional. Apesar de sua limitação no campo da BI, ela se mostrou altamente engajada no desenvolvimento da prática de governança de dados ou de informação na organização. E a motivação para tal foi justamente o desenvolvimento de competências de *analytics* na organização.

A capacidade analítica aspiracional do caso H exige significativas reformulações no ambiente organizacional para o avanço analítico e isto foi central para a adoção da estrutura de governança de informação. Como mencionado pela entrevistada:

“Não é possível avançar na maturidade analítica sem que ações básicas com relação aos dados sejam adotadas na organização”, comenta a líder da governança de informação da empresa H.

Independentemente da capacidade analítica da organização, as estruturas de governança de dados ou de informação apareceram relevantes e apresentaram objetivos comuns dentre os domínios do modelo de governança de dados ou de informação sugerido por este trabalho, incluindo: melhores práticas para o gerenciamento de dados, melhores índices da qualidade dos dados ou das informações, incentivo à colaboração e à comunicação de melhores práticas, transformação da organização com relação à mentalidade.

Com vistas à primeira proposição de pesquisa, refuta-se a ideia de que capacidades analíticas transformadas sejam um fator determinante para a adoção das governanças de dados ou de informação. Com vistas aos estudos de casos, foi possível observar que organizações com diferentes níveis de capacidades analítica tem adotado a governança de dados ou de informação para

atuar sobre seus ativos de dados. Inclusive, notou-se a governança de dados ou de informação corresponde a uma estrutura relevante para o avanço analítico independentemente do nível da capacidade analítica.

4.4 Proposições de pesquisa

Esta sessão resume as considerações com relação as proposições de pesquisa, as quais foram discutidas nas sessões anteriores. O Quadro 25 identifica as análises finais das proposições de pesquisa.

Quadro 25 – Análise das proposições de pesquisa.

	Proposições de pesquisa	Análise
1	Capacidades analíticas transformadas correspondem a um fator determinante para a adoção das governanças de dados e de informação.	Refutada
2	As ações organizacionais para governar os dados ou as informações no contexto do <i>big data</i> ainda são pouco consolidadas e conhecidas no ambiente organizacional.	Refutada parcialmente
3	Apesar das particularidades organizacionais, existem práticas, sub-práticas e fatores comuns que frequentemente são adotados para viabilizar a adoção das governanças de dados ou de informação.	Validada

Fonte: elaborado pela autora

Quanto à primeira proposição de pesquisa, as organizações podem adotar as estruturas de governança de dados ou de informação independente de suas capacidades analíticas. Inclusive quanto mais restritas forem suas capacidades analíticas, maiores são os trabalhos e os esforços organizacionais para adequar o fluxo dos dados e das informações na organização, viabilizando os avanços analíticos.

A segunda proposição foi refutada parcialmente, pois as ações são pouco consolidadas, embora sejam conhecidas pelas organizações. A dificuldade está na fase de implementação dos modelos e das ações pré-estabelecidas, o que pode

culminar com a baixa consolidação das ações notadas dentre os casos organizacionais.

A terceira proposição de pesquisa foi validada, o que permite constituir uma lista com fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*, independente da capacidade analítica da organização, do setor a que pertence e se são negócios voltados à manufatura ou oferta de serviços.

4.5 Fatores motivadores, viabilizadores e determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*

Para responder à questão geral de pesquisa: “como as organizações podem adotar as governanças de dados ou de informação no ambiente *big data*”, vale-se compreender que fatores motivadores, viabilizadores ou determinantes para a adoção das governanças de dados ou de informação não necessariamente são os mesmos ou estão alinhados na prática organizacional.

Por fatores motivadores compreendem-se aqueles que expressam os motivos pelos quais a organização deve aderir uma nova estrutura com foco na governança de dados ou de informação. Revisitando os objetivos destas governanças, tais motivações podem ser: atender as regulações, aumentar a qualidade dos dados, gerenciar meta-dados, dentre outros objetivos. Neste momento, são contempladas aspirações da organização, antes da etapa de adoção.

Os fatores viabilizadores são aqueles que facilitam a adoção das políticas, das práticas e das sub-práticas da governança de dados ou de informação. Aqui já se fala na etapa de implementação do modelo, cujos elementos facilitadores podem compreender aspectos estruturais, tecnológicos ou setoriais. Não necessariamente os fatores viabilizadores são realidade nas empresas, pois há elementos desafiadores na prática organizacional que prejudicam a adoção das governanças de dados ou de informação. Dentre os fatores listados pelo Quadro 2, foram constatados fatores viabilizadores para a adoção das governanças de dados ou de informação com base nos estudos de caso:

- Estrutura organizacional centralizada;

- Alta padronização da TI e alta integração dos processos;
- Cultura da TI que promove o uso estratégico da informação (cultura *data-driven*);
- Regulação setorial;
- Aumento na quantidade de dados disponível para gerenciamento e análise.

Em outra instância, os fatores determinantes são os elementos imprescindíveis para o estabelecimento do modelo de governança e que determinam o alcance dos objetivos pré-determinados das estruturas de governança de dados ou de informação. Os fatores determinantes podem ser contrários aos fatores viabilizadores ou estarem alinhados.

No âmbito da estrutura organizacional, as estruturas centralizadas viabilizam a adoção de políticas e práticas comuns para governar dados ou informações. No entanto, é justamente a descentralização organizacional, a departamentalização e os vários silos na organização que motivam ou justificam a necessidade de adoção dos modelos de governança de dados ou de informação (apesar de constituírem um fator inibidor). Como identificado na empresa A, a governança de dados somente se tornou relevante quando a organização cresceu, quando o número de pessoas aumentou e quando as pessoas já não se conheciam mais. Enquanto a organização A possuía um ambiente mais simples e com menos silos organizacionais, não era necessária a instituição formal da governança de dados ou de informação, já que as questões eram alinhadas dentro de um grupo seleto de *stakeholders* à medida que elas surgiam.

Além disto, o mesmo racional pode ser aplicado ao portfolio de produtos ou serviços oferecidos pelas organizações: quanto mais complexos, maior é a necessidade de adoção das governanças de dados ou de informação. A alta complexidade organizacional foi sugerida por Tallon et al. (2014) como um fator inibidor para a adoção das governanças de dados ou de informação, pois ela dificulta o consenso e o estabelecimento de práticas comuns. A partir do Quadro 26, é possível notar que a alta complexidade e o grande tamanho organizacional constituem um dos fatores determinantes para a adoção da governança de dados ou de informação no ambiente *big data* (fator determinante 1).

A alta padronização dos sistemas e a alta integração dos processos favorecem a adoção das práticas de governança, como apontado por Tallon et al. (2014) e, portanto, viabilizam a adoção de práticas comuns para a governança de dados ou de informação. No entanto, é a falta de padronização dos sistemas e a baixa integração dos processos que motivam a adoção das estruturas de governança de dados ou de informação. Assim, apesar da padronização dos sistemas e a alta integração dos processos viabilizarem a adoção das governanças de dados ou de informação, eles não motivam ou demarcam a necessidade da adoção da governança de dados ou de informação; é justamente a necessidade de harmonizar, padronizar e integrar o fluxo de dados entre sistemas e processos que motivam ou justificam a adoção das governanças de dados ou de informação.

A cultura *data-driven* foi prevista pela literatura como um fator viabilizador para a adoção da governança de dados ou de informação. Além disto, McAfee e Brynjolfsson (2012), Ross et al. (2013) a consideram essencial para a extração de valor dos dados e das informações no ambiente *big data*. Pelos casos contemplados nesta pesquisa, foi verificada a atuação viabilizadora e determinante da cultura *data-driven* para a adoção efetiva das políticas e práticas da governança de dados ou de informação (fator determinante 5). Para as organizações que não contemplam o ambiente *data-driven*, a atuação efetiva e pró-ativa dos *stakeholders* da governança na transformação cultural auxilia o processo de engajamento dos *stakeholders* da organização e no cumprimento das normas estabelecidas para governar e gerenciar os dados e as informações no ambiente organizacional. Para organizações que já contemplam o ambiente *data-driven*, as ações delineadas pela estrutura de governança de dados ou de informação podem ser adotadas com maior rapidez e eficácia. Portanto, além de constituir num fator viabilizador, a transformação cultural no sentido *data-driven* determina a adoção das medidas de governança de dados ou de informação.

A regulação setorial também exerce papel motivador e viabilizador para a adoção da governança de dados ou de informação: motivador pois a adequação dos fluxos dos dados com relação às exigências regulatórias é um dos objetivos do modelo de governança; e viabilizador pois os *stakeholders* são mais susceptíveis ao consenso e à adesão das políticas de governança de dados ou de informação a justificativa é o atendimento às exigências regulatórias. O aumento na quantidade dos dados no ambiente organizacional para gerenciar e analisar também motiva ou viabiliza a

adoção das governanças de dados ou de informação, como apontado por Tallon et al. (2014).

Quadro 26 – Fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*.

	Fatores determinantes	Casos práticos evidenciados	Autores
1	Organizações grandes, globais, difusas, com estruturas descentralizadas de negócios e portfólio complexo de produtos ou serviços;	A, B, C, D, E, F, G, H (todos)	Tallon et al. (2014)
2	Apontar um <i>C-level</i> responsável pela governança de dados ou de informação, definir gerentes na estrutura para conduzir as atividades com relação aos dados e as informações, determinar <i>data owners</i> e <i>data stewards</i> ;	A, B, C, D, E, F, G, H (todos)	Malakar (2007); Rosembaum (2007); Tallon (2013), Tallon et al. (2014), Thompson et al. (2015), Demarquet (2016); Koltay (2016)
3	Estabelecer um comitê de dados ou outros meios para reunir a alta cúpula e os principais líderes da organização para definir as estratégias e as ações com relação aos dados e as informações na organização;	C, D, E, F, G, H	Rau (2004); Weill, Ross (2005); Malakar (2007); Weber et al. (2009); Khatri, Brown (2010); Otto (2011); Tallon (2013); Tallon et al. (2014)
4	Atuação do departamento de TI nas atividades de gerenciamento de dados ou de informação, viabilizando e executando atividades operacionais com relação aos dados e as informações dentre as bases de dados e sistemas de informação;	A, B, C, D, E, F, G, H	Kooper et al. (2011), Tallon et al. (2014)
5	Atuar ativamente na transformação cultural da organização para <i>data-driven</i> ;	A, B, C, D, E, F, G, H (todos)	McAfee, Brynjolfsson (2012); Ross et al. (2013); Tallon et al. (2014)
6	Promover a comunicação e a colaboração interna para definir o modelo de governança e assim identificar pontos críticos no trato com os dados e solucionar os entraves; Desenvolver a comunicação com relação à eficácia das políticas e a necessidade de adequação dos <i>stakeholders</i>		Weill, Woodham (2002); Rau (2004); Weill, Ross (2005); Nfuka, Rusu (2011); Tallon (2013); Tallon et al. (2014); Kurti et al. (2014); Koltay (2016)
7	Definir, gerenciar e controlar meta-dados, incluindo parâmetros, métricas e siglas de referência para uso pelas diversas equipes de um determinado negócio ou pela organização como um todo;	A, B, C, D, E, F, G, H (todos)	Khatri, Brown (2010); Fu et al. (2011); Otto et al. (2011); Koltay (2016)
8	Definir os padrões, as exigências e o controle sobre a qualidade dos dados, assegurando a corretude, integridade, completude e transparência dos dados e das informações; contemplar as atividades de gerenciamento da qualidade dos dados.	A, B, C, D, E, F, G, H (todos)	Weber et al. (2009); Khatri, Brown (2010); Otto et al. (2011); Otto (2011); Kooper et al. (2011); Fu et al. (2011); Tallon et al. (2014)

Fonte: elaborado pela autora

O Quadro 26 identifica os oito fatores determinantes identificados nesta pesquisa para a adoção das governanças de dados ou de informação no ambiente *big data*. O primeiro fator determinantes trata dos aspectos da complexidade e do tamanho organizacional. Os fatores determinantes 2 e 3 relacionam-se às práticas estruturais e focalizam na estrutura de governança de dados ou de informação necessária para que o modelo de governança possa se estabelecer e operar no contexto organizacional.

O quarto fator determinante observado dentre os casos organizações correspondeu a atuação do departamento de TI nas atividades de gerenciamento de dados ou de informação, viabilizando e executando atividades operacionais com relação aos dados e as informações dentre as bases de dados e sistemas de informação.

O fator determinante 5 corresponde à transformação cultural no sentido *data-driven*, como discutido acima, e o fator determinante 6 contempla os aspectos de colaboração e comunicação entre *stakeholders*, fundamentais para que o modelo possa ser estabelecido. Sem práticas relacionais, torna-se impossível identificar os pontos críticos, estabelecer o consenso, divulgar as medidas necessárias para serem tomadas.

Os fatores determinantes 7 e 8 contemplam práticas operacionais fundamentais do modelo de governança de dados ou de informação. Através do sexto fator determinante foram contempladas as atividades de definição e gerenciamento de meta-dados. Tais atividades foram observadas necessárias e imprescindíveis para assegurar a singularidade dos dados e das informações na organização, constituindo a base do modelo de governança. Já o sétimo fator determinante contempla as atividades de gerenciamento da qualidade dos dados, notada fundamental para que possa ser extraído valor dos ativos de dados.

Com vistas às capacidades analíticas transformadas de *analytics*, esta por si só não corresponde a um fator determinante para a adoção das governanças de dados ou de informação. O alcance de maiores níveis de capacidades analíticas pode corresponder a um fator motivador para a adoção das governanças de dados ou de informação.

5. Conclusões

Esta pesquisa fornece uma consolidação teórica relevante para o campo da governança de dados ou da informação, contemplando vasta lista de variáveis da literatura de inteligência competitiva, em particular para a extração de valor dos dados no ambiente *big data*, e da literatura de governança de TI, governança de dados e governança de informação. Além disto, foram identificados oito fatores determinantes para a adoção das governanças de dados ou de informação no ambiente *big data*. Estes fatores compreendem práticas estruturais, relacionais e operacionais.

A estruturação das iniciativas de governança de dados ou de informação é emergente no contexto organizacional e motivada pela necessidade de melhor organizar e aprimorar o uso dos ativos de dados nos processos de tomada de decisões para a obtenção de vantagens competitivas. Foi notado que não há competição entre os campos inteligência de negócios e governança de dados ou de informação, pelo contrário: a governança de dados ou de informação pode servir como base fundamental para melhor contribuir para a formação e gestão dos ativos de dados na organização, contribuindo para o cálculo de indicadores mais assertivos.

Foi também possível observar que as ações organizacionais com relação à governança de dados ou de informação são pouco consolidadas, porém conhecidas pelas organizações. Em geral, as organizações sabem o que querem obter, mas as barreiras culturais e gerenciais dificultam o processo de implementação das iniciativas de governança de dados ou de informação.

As estruturas de governança de dados ou de informação são emergentes nas organizações e se apresentam sob algumas formas: podem estar alocadas integralmente no departamento de TI; podem estar parcialmente alocadas no departamento de TI; podem pertencer integralmente a uma unidade de negócios; podem contemplar múltiplas unidades de negócios e grupos funcionais da organização. Independente da forma que a estrutura de governança de dados ou de informação se posiciona na organização, o departamento de TI auxilia nas atividades de gestão de dados e conduz uma série de atividades operacionais nas bases de dados para promover a integridade, correteude, integração dos dados. Com

vistas a este ponto, a atuação do departamento da TI sobre questões operacionais dos dados e das informações foi tida como um fator determinante para a adoção dos modelos de governança de dados ou de informação no ambiente *big data*.

A adoção das estruturas formais de governança de dados ou de informação foi observada em organizações grandes, globais, difusas, com estruturas descentralizadas de negócios e portfolio complexo de produtos ou serviços. Tais elementos constituem o primeiro fator determinante proposto pela pesquisa. Dentre os casos contemplados nesta pesquisa, observou-se que os modelos de governança de dados ou de informação apresentaram grande importância nos ambientes complexos, difusos, descentralizados e heterogêneos. Estes ambientes necessitam de esforços organizacionais para definir as responsabilidades com relação aos dados, promover a qualidade dos dados, gerenciar os dados e promover a transformação cultural da organização. Apesar disto, os elementos relativos à alta complexidade e ao tamanho organizacional constituem fatores inibidores ao processo de adoção das governanças, embora determinantes, pois dificultam no estabelecimento de políticas e de práticas comuns no ambiente organizacional.

No âmbito estrutural, foram sugeridos dois fatores que determinam na adoção das governanças de dados ou de informação, sendo estes o segundo e o terceiro fator proposto pela pesquisa: 2º - Apontar um *C-level* responsável pela governança de dados ou de informação, definir gerentes na estrutura para conduzir as atividades com relação aos dados e as informações, determinar os “*data owners*” e “*data stewards*”; 3º - Estabelecer um comitê de dados ou outros meios para reunir a alta cúpula e os principais líderes da organização para definir as estratégias e as ações com relação aos dados e as informações na organização.

Os fatores determinantes 2 e 3 englobam práticas estruturais relevantes para que o modelo de governança de dados ou de informação possa se estabelecer e operar no contexto organizacional. A definição de um *C-level* responsável pela governança confere visibilidade e maior credibilidade do modelo de governança de dados ou de informação na organização. Em particular, foi notada a presença de dois *C-level* suportando as iniciativas de governança de dados ou de informação: *Chief Data Officer* (CDO) se reportando ao *Chief Financial Officer* (CFO) da organização. No contexto de *stakeholders* da estrutura de governança de dados ou de informação, foi notado relevante definir os gerentes da governança de dados para que estes

sejam responsáveis pela operação do modelo de governança de dados ou de informação. Além disto, apontar os “*data owners*” e os “*data stewards*” nas unidades de negócios assegura que as políticas e as práticas da governança de dados ou de informação sejam de fato adotadas na organização.

Os comitês de dados foram notados elementos também relevantes para que o modelo de governança de dados ou de informação tenha aderência na organização e engajamento dos líderes de negócios. Através desta estrutura é possível envolver os usuários na definição das políticas de governança, estabelecer um canal de comunicação e colaboração entre os *stakeholders* da organização, identificar oportunidades de melhoria no trato com os dados e gerenciar potencial conflitos.

No âmbito relacional, foram sugeridos outros dois fatores determinantes para a adoção das governanças de dados ou de informação: 5º - Atuar ativamente na transformação cultural da organização para *data-driven*; 6º - Promover a comunicação e a colaboração interna para definir o modelo de governança e assim identificar pontos críticos no trato com os dados e solucionar os entraves; desenvolver a comunicação com relação à eficácia das políticas e a necessidade de adequação dos *stakeholders*.

Em organizações que ainda não contemplam a cultura *data-driven*, a atuação sobre a transformação cultural determina a adoção das políticas e das práticas estabelecidas pelo modelo de governança de dados ou de informação. Além disto, a pré-existência da cultura *data-driven* pode também servir como um fator viabilizador para a adoção das governanças de dados ou de informação no ambiente *big data*. Tal fator foi intensamente discutido pelas literaturas de *big data* e corresponde a uma questão organizacional relevante para que os dados gerados no ambiente *big data* possam ter seu valor extraído e utilizado pela organização para fins estratégicos.

O quinto fator determinante estabelece o papel da comunicação e da colaboração para que o modelo de governança de dados ou de informação no ambiente *big data* possa ser estabelecido. Sem tais competências, os objetivos e os princípios do modelo de governança são pouco comunicados e dissipados pelos grupos de trabalho, inviabilizando a coleta dos benefícios com relação ao fluxo dos dados e das informações no contexto organizacional. As práticas relacionais do modelo de governança de dados ou de informação são tão mais relevantes quanto mais desvinculada for a estrutura de governança de dados ou de informação do negócio

e quanto menor for o poder da alta gestão que suporta a governança de dados ou de informação.

No âmbito operacional foram notados outros dois fatores que determinam a adoção dos modelos de governança de dados ou de informação: 7º - Definir, gerenciar e controlar meta-dados, incluindo parâmetros, métricas e siglas de referência para uso pelas diversas equipes de um determinado negócio ou pela organização como um todo; 8º - Definir os padrões, as exigências e o controle sobre a qualidade dos dados, assegurando a corretude, integridade, completude e transparência dos dados e das informações; contemplar as atividades de gerenciamento da qualidade dos dados.

Os esforços com relação aos meta-dados é um desafio estratégico para a corporação, pois permite ganho de eficiência e transparência no fluxo dos dados pela organização. Além disto, aprimorar a qualidade dos dados favorece constituir ativos de dados mais íntegros, completos e precisos, tornando as análises mais assertivas dos conjuntos de dados e das informações. Atuar sobre a qualidade dos ativos dos dados na organização viabiliza laços entre as unidades de negócios e os *stakeholders* financeiro, agregando valor às iniciativas da estrutura perante os negócios.

Os fatores determinantes aqui discutidos foram resultado de pesquisas em organizações voltadas aos produtos (manufaturas) e aquelas orientadas aos serviços. Eles independem do nível da capacidade analítica da organização e do setor a que pertence. A pesquisa também traz à tona diferenças entre fatores viabilizadores, inibidores, motivadores e determinantes para a adoção das governanças de dados ou de informação. Esta observação é fundamental para a literatura se desenvolver no sentido de identificação de fatores críticos de sucesso para o processo adotivo das governanças de dados ou de informação.

Estes fatores (inibidores, viabilizadores, motivadores ou determinantes) podem apresentar características convergentes ou divergentes. Podem estar alinhados quanto aos conceitos ou serem opostos. Os fatores motivadores contemplam os objetivos para a adoção das governanças de dados ou de informação, já os viabilizadores são aqueles que facilitam na adoção dos modelos de governança e os inibidores são aqueles que prejudicam ou inibem com que sejam estabelecidas práticas ou políticas comuns. Foram notados como fatores motivadores ou objetivos

para a adoção das governanças de dados ou de informação no contexto desta pesquisa:

- Definir as responsabilidades e os papéis com relação aos dados e as informações no segmento de negócio ao qual a estrutura de governança de dados ou de informação é vinculada ou no âmbito organizacional;
- Definir os padrões, as exigências e o controle sobre a qualidade dos dados e as informações, assegurando a correteude, integridade, completude e transparência dos dados e das informações;
- Estabelecer as atividades de gerenciamento dos dados e das informações, contemplando principalmente a definição e o gerenciamento de meta-dados;
- Estabelecer processos mais eficientes com relação ao fluxo dos dados entre *stakeholders* de uma cadeia produtiva;
- Adequar o trato com os dados e as informações com relação às exigências regulatórias;
- Fomentar a colaboração e a comunicação entre diferentes *stakeholders* de um segmento de negócios para identificar oportunidades de melhoria e contribuir para processos mais eficientes com relação aos dados e as informações;
- Atuar na transformação cultural da organização no sentido *data-driven*, permitindo tomadas de decisão mais assertivas e constituindo um ambiente para o avanço da capacidade analítica da organização.

O ambiente *big data* impõem capacidades de processamento e de análise de dados em uma dinâmica nunca antes vista ou necessária, fazendo com que organizações pouco preparadas neste contexto se preocupem em adotar iniciativas para acelerar o uso dos dados nos processos de tomada de decisões. Neste contexto, organizações aspiracionais no campo da *analytics* contemplam ambientes pouco consolidados com relação ao uso dos dados, cujos benefícios podem ser significativos ao adotar as governanças de dados ou de informação. Embora o processo de adoção das governanças de dados ou de informação nestas organizações seja complexo, os esforços são necessários para definir os papéis e as responsabilidades com relação aos dados, adequar os aspectos de qualidade

dos dados, estabelecer e gerenciar meta-dados contribuir para a transformação cultural da organização no sentido *data-driven*.

As organizações com capacidades analíticas transformadas também são beneficiadas dos modelos de governança de dados ou de informação. As ações de promoção de maiores níveis de qualidade dos dados, melhores formas para o gerenciamento dos dados, por exemplo, permitem obter maiores níveis de eficiência e eficácia com relação ao uso dos ativos de dados. Constituir melhores ativos de dados é a base para desdobramentos analíticos mais avançados.

Foi também possível concluir que a literatura de Inteligência de negócios ou *big data analytics* pode contribuir para a formação da governança de dados ou de informação, e vice-versa. Tais campos se complementam, onde os fundamentos que estabelecem os modelos de governança de dados ou de informação devem estar alinhados aos objetivos estratégicos da organização em explorar o valor dos ativos de dados. Sugere-se expandir a avaliação das competências analíticas organizacionais com outros modelos, já que a pesquisa realizada se baseou no modelo de LaValle et al. (2011).

Com vistas aos fatores identificados, foi possível expandir o modelo de governança de dados ou de informação contemplando os domínios relativos à colaboração e comunicação, mudança cultural. As atividades operacionais do modelo de governança defendido pela literatura puderam ser agrupadas em um domínio único denominado por “gerenciamento de dados”.

A pesquisa propõe expansão na conceituação geral dos termos governança de dados ou governança de informação. Em particular, a pesquisa sugere maior rigor e detalhamento das práticas operacionais da governança de dados ou de informação para diferenciar seus conteúdos daqueles relativos ao campo da inteligência de negócios. Logo, sugere-se como conceituação da governança de dados ou de informação:

- “As governanças de dados ou de informação envolvem o estabelecimento de regras e de procedimentos sobre quem detém os poderes de decisão e responsabilidades nas tomadas de decisão na organização com relação aos ativos de dados. Também contemplam atividades de apoio e de suporte na garantia da qualidade, confiabilidade, transparência e singularidade dos dados dentre os diferentes grupos da organização, permitindo a constituição de indicadores confiáveis para os processos de tomadas de decisão. Dentre

tais atividades de apoio ou suporte, contemplam: a análise e a avaliação dos ativos de dados com relação à qualidade, segurança, singularidade, transparência, acessibilidade; criação de meta-dados com fins de promoção de uma representação única das informações na organização”.

A consolidação teórica proposta pela tese pode ser avaliada e testada em outros ambientes organizacionais através de técnicas qualitativas ou quantitativas através de pesquisas *survey*. Sugere-se também que mais pesquisas possam contemplar o referencial teórico da inteligência competitiva para agregar na formação da literatura de governança de dados ou de informação e vice-versa. Existem sinergias entre as teorias já que contemplam o campo dos dados e das informações no ambiente *big data*.

Para fins práticos, os fatores determinantes identificados nesta pesquisa podem servir como orientação básica para agilizar e facilitar a adoção da estrutura de governança de dados ou de informação no ambiente organizacional. Para as organizações que almejam o alcance de vantagens competitivas no ambiente *big data*, sugere-se a incorporação dos modelos de governança de dados ou de informação como forma de aperfeiçoar os ativos de dados e viabilizar o uso da *advanced analytics*.

Para futuros trabalhos, sugere-se expandir e complementar as análises dos fatores determinantes sugeridos por esta pesquisa com auxílio de outras técnicas de pesquisa, como a pesquisa *survey*. Sugere-se que mais trabalhos possam explorar o campo e que assim possam ser identificados e propostos Fatores Críticos de Sucesso para a adoção das governanças de dados ou de informação no ambiente *big data*.

Além disto, sugere-se expandir os estudos dos fatores determinantes em organizações de outros nichos empresariais (como o varejo, telecomunicações), *start-ups*, organizações totalmente digitais, aquelas voltadas às redes sociais e aquelas de menor porte.

Referências

ADMA (*Association for Data-driven Marketin & Advertising*). **BEST PRACTICE GUIDELINE: BIG DATA** - A guide to maximising customer engagement opportunities through the development of responsible Big Data strategies. Sydney, 2013.

ALHASSAN, I.; SAMMON, D., DALY, M. Data governance activities: an analysis of the literature. **Journal of Decision Systems**. 2016.

ARNABOLDI, M. **Case study methodology**. Notas de aula. Politecnico di Milano, 2017.

ASSIS, C. B. **Governança e Gestão da tecnologia da informação: diferenças na aplicação em empresas brasileiras**. Dissertação de mestrado - Escola Politécnica da Universidade de São Paulo, 2011.

BARATA, A. M. Governança de dados em organizações brasileiras: uma avaliação comparativa entre os benefícios previstos na literatura e os obtidos pelas organizações. Universidade de São Paulo: Dissertação de mestrado. 2015.

BARATA, A. M.; PRADO, E. P. V.; Data Governance Framework: Case study in one Brazilian Organization. 2th INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS & TECHNOLOGY MANAGEMENT - **CONTECSI**. 2015.

BARTON, D.; COURT, D. Making advanced analytics work for you. **Harvard Business Review**. 2012.

BATRA, S. Big data analytics and its reflections on DIKW Hierarchy. **Review of Management**, v. 4, n. 1, 2014.

BERKOOZ, G. How Chief Data Officers Can Get Their Companies to Collect Clean Data. **Harvard Business Review**, 2017.

BERNHARDT, D. C. 'I want it fast, factual, actionable'-Tailoring Competitive Intelligence to Executives' Needs. **Long Range Planning**, v. 27, n.1, p. 12 - 14, 1994.

BLASIAK, K. **Big Data; A management revolution - The emerging role if big data in businesses**. Bachelor of Business Administration Thesis. Helsinki Metropolia University of Applied Sciences. 2014.

BOSE, R. Advanced analytics: opportunities and challenges. **Industrial Management & Data Systems**, v. 102, n. 1, p. 155 - 172, 2009.

BOSE, R. Competitive intelligence process and tools for intelligence analysis. **Industrial Management & Data Systems**, v. 108, p. 510 - 528, 2008.

BOYD, D.; CRAWFORD, K. Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. **Information Communication & Society**, v. 15, n. 5, p. 662 - 679, 2012.

BRAGANZA, A. Rethinking the data–information–knowledge hierarchy: towards a case-based model. **International Journal of Information Management**, v. 24, p. 347 - 356, 2004.

BRODHAG, C. Information, gouvernance et développement durable. **International Political Science Review**, v. 21, n. 3, p. 311- 327, 2000.

CABRAL NETTO, O. V. **Inteligência Competitiva: sistema aplicado ao mercado de crédito**. Trabalho de formatura - Universidade de São Paulo. 2007.

CABRAL NETTO, O. V. **Uma visão holística da inteligência competitiva para a construção de uma teoria**. Dissertação de mestrado - Universidade de São Paulo. 2011.

CARVALHO, M. M.; LAURINDO, F. J. B. **Estratégia competitiva: dos conceitos à implementação**. Editora Atlas: São Paulo, 2012.

CHEN, H.; CHAU, M.; ZENG, D. CI spider: a tool for competitive intelligence on the web. **Decision Support Systems**, v. 34, 2002.

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business Intelligence and Analytics: From big data to big impact. **MIS Quarterly**, v. 36, n. 4, p. 1165 - 1188, 2012.

CHEN, M.; MAO, S.; LIU, Y. Big Data: A Survey. **Mobile Networks & Applications**, v. 19, n. 2, p. 171 - 209, 2014.

DAVENPORT, T. H; BARTH, P.; BEAN, R. How "big data" is different. **Harvard Business Review**, 2012.

DAVENPORT, T. H. Analytics 3.0. **Harvard Business Review**, 2013.

DAVENPORT, T. H. Competing on Analytics. **Harvard Business Review**. 2006.

DAVENPORT, T. H.; BARTH, P.; BEAN, R. How big data is different. **Harvard Business Review**, 2012.

DEMARQUET, G. Five Key Reasons Enterprise Data Governance Matters to Finance ... and Seven Best Practices to Get You There. **The Journal of Corporate Accounting & Finance**, 2016.

DONALDSON, A.; WALKER, P. Information governance - a view from the NHS. **International Journal of Medical Informatics**, v. 73, p. 281 - 284, 2004.

DÖRNER, K.; EDELMAN, D. What 'digital' really means. **Mckinsey Quarterly**, 2015.

DU TOIT, A. S. A. Competitive intelligence in the knowledge economy: what is in it for South African manufacturing enterprises? **International Journal of Information Management**, v. 23, p. 111 - 120, 2003.

EELLS, R. S. F. **The Meaning of Modern Business: An Introduction to the Philosophy of Large Corporate Enterprise**. New York: Columbia University Press, 1960.

EISENHARDT, K. M. Building theories from case study research. **Academy of Management Review**, v. 14, n. 4, p. 532 - 550, 1989.

ESTEVEES, J. M. S. **Definition and Analysis of Critical Success Factors for ERP Implementation Projects**. Tese de doutoramento - Universidade Politécnica da Catalunya. 2004.

ESTEVEES, J. M. S.; PASTOR, J. Using a Multimethod Approach to Research Enterprise Systems Implementations. **Electronic Journal of Business Research Methods**, v. 2, n. 2, p. 69 - 84, 2004.

FU, X.; WOJAK, A.; RIDLEY, M.; TRAVIS, K. Data governance in predictive toxicology: A review. **Journal of Cheminformatics**, p. 3 - 24, 2011.

FURLAN, P. K.; LAURINDO, F. J. B. Agrupamentos epistemológicos de artigos publicados sobre *big data analytics*. **TransInformação**, Campinas, v. 29, ed. 1, p. 91 - 100, 2017a.

FURLAN, P. K.; LAURINDO, F. J. B. Estudo bibliométrico sobre as governanças de dados e de informação. **14th International Conference on Information Systems & Technology Management - CONTECSI**, São Paulo, 2017b.

FURLAN, P. K.; LAURINDO, F. J. B. Inteligência de Negócios, big data e analytics: estudo da tendência de publicação conjunta. **XXII Simpósio de Engenharia de Produção**, Bauru, 2015.

FURLAN, P. K.; LAURINDO, F. J. B. Um estudo bibliométrico sobre big data na literatura de gestão de projetos **13th International Conference on Information Systems & Technology Management - CONTECSI**, São Paulo, 2016.

GANDOMI, A.; HAIDER, M. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, v. 35, p. 135 - 144, 2015.

GOODHUE, D. L., KIRSCH, L. J., QUILLARD, J. A.; WYBO, M. D. Strategic data planning: lessons from the field. **MIS Quarterly**, v. 16, p. 11-34, 1992.

HAIDER, W.; HAIDER, A. Governance Structures for Engineering and Infrastructure Asset Management. **Proceedings of PICMET '13: Technology Management for Emerging Technologies**, 2013.

HAUPT, R.; SCHOLTZ, B.; CALITZ, A. Using Business Intelligence to Support Strategic Sustainability Information Management. **Annual Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT)**, 2015.

IBM. **Information Governance Principles and Practices for a Big Data Landscape**, 2014.

ITGI (*IT Governance Institute*). **Board Briefing on IT Governance**. IT Governance Institute, 2a edição, 2003.

JENSTER, P. V. Using Critical Success Factors in Planning. **Long Range Planning**, v. 20, n. 4, p. 102 - 109, 1987.

KAHANER, L. **Competitive intelligence: how to gather, analyse, and use information to move your business to the top**. New York, Touchstone Book, 1996.

KAMBATLA, K.; KOLLIAS, G.; KUMAR, V.; GRAMA, A. Trends in big data analytics. **Journal of Parallel and Distributed Computing**, v. 74, n. 7, p. 2561 - 2573, 2014.

KEMP, R. Legal aspects of managing Big Data. **Computer Law & Security Review**, v. 30, p. 482 - 491, 2014.

KHATRI, V.; BROWN, C. V. Designing Data Governance. **Communications of the ACM**, v. 53, n. 1, 2010.

KIRSCH, L. Deploying common systems globally: The dynamics of control. **Information Systems Research**, 15, 4, 374–395, 2004.

KOLTAY, T. Data governance, data literacy and the management of data quality. **International Federation of Library Associations and Institutions**, v. 42, n. 4, p. 303-312, 2016.

KOOIMAN, J. **Governing as Governance**. SAGE Publications: India, 2003.

KOOPER, M. N; MAES, R.; LINDGREEN, E. E. O. Roos. On the governance of information: Introducing a new concept of governance to support the management of information. **International Journal of Information Management**, v. 31, p. 195 - 200, 2011.

KORONIOS, A.; GAO, J.; SELLE, S. Big data project success - a meta analysis. **Pacific Asia Conference on Information Systems (PACIS) 2014 Proceedings**, 2014.

KOURTESIS, D.; ALAVAREZ-RODRIGUEZ, J. M.; PARASKAKIS, I. Semantic-based QoS management in cloud systems: Current status and future challenges. **Future Generation Computer Systems - The International Journal of Grid Computing and E-science**, v. 3, p. 307 - 323, 2014.

KURTI, I.; BAROLLI, E.; SEVRANI, K. Effective IT Governance in the Albanian Public Sector - a critical success factors approach. **The Electronic Journal of Information Systems in Developing Countries**, v. 63, n. 6, p. 1 -22, 2014.

KWON, O.; SIM, J. M. Effects of data set features on the performances of classification algorithms. **Expert systems with applications**. p. 1847 - 1857. 2013.

LAURINDO, F. J. B. **Tecnologia da Informação - Planejamento e Gestão de Estratégias**. Editora Atlas: São Paulo, 2008.

LAVALLE, S.; LESSER, E.; SHOCKLEY, R.; HOPKINS, M. S.; KRUSCHWITZ, N. Big Data, Analytics and the Path from Insights to Value. **MIT Sloan Management Review**, 2011.

LEIDECKER, J. K., BRUNO, A. V. Identifying and Using Critical Success Factors. **Long Range Planning**, v. 17, n. 1, p. 23-32, 1984.

LEMING, R. Why is information the elephant asset? An answer to this question and a strategy for information asset management. **Business Information Review**, v. 32, p. 212 - 219, 2015.

LIM, EE-P.; CHEN, H.; CHEN, G. Business Intelligence and Analytics: Research Directions. **ACM Transactions on Management Information Systems**, v. 3, n. 4, 2013.

LISMONT, J.; VANTHIENEN, J.; BAESSENS, B.; LEMAHIEU, W. Defining analytics maturity indicators: A survey approach. **International Journal of Information Management**, v. 37, p. 114-124, 2017.

MALAKAR, D. **Chief Data Steward or Chief Data Officer: Another C-Level Acronym?**. Disponível em: <<http://tdan.com/chief-data-steward-or-chief-data-officer-another-c-level-acronym/4581>>. Acessado em Abril 2017. Publicado em: Janeiro 2007.

MANYIKA, J.; CHUI, M.; BROWN, B. Big data: the next frontier for innovation, competition, and productivity. **McKinsey Global Institute**; 2011.

MARCAN, M. G. **The Chief Data Officer**. 2008

McAFEE, A.; BRYNJOLFSSON, E. Big Data: The Management Revolution. **Harvard Business Review**, v. 90, n. 10, p. 60, 2012.

McCRACKEN, G. **The Long Interview**. Newbury Park, CA: Sage, 1988.

MCFARLAN, W. E.: Information Technology Changes The Way You Compete. **Harvard Business Review**, v.62, n.3, p. 98-103, 1984.

META GROUP. **3D Data Management: Controlling Data Volume, Velocity, and Variety**, 2001.

MIGUEL, P. A. C. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. **Produção**, v. 17, n. 1, p. 216-229, 2007.

MIGUEL, P. A. C. **Metodologia de pesquisa em engenharia de produção e gestão de operações**. Rio de Janeiro: Elsevier, 2010.

MIMS, C. **Data is the new middle manager**. The Wall Street Journal, 2015.

NEVES, P. C.; SCHMERL, B.; CÁMARA, J.; BERNARDINO, J. Big Data in Cloud Computing: Features and Issues. **Proceedings of the International Conference on Internet of Things and Big Data (IoTBD 2016)**, p. 307-314, 2016.

NFUKA, E. N. **IT Governance in Tanzanian public sector organizations**. Tese de doutoramento - Stockholm University, 2012.

NFUKA, E. N.; RUSU, L. The effect of critical success factors on IT governance performance. **Industrial Management & Data Systems**, v. 111, n. 9, p. 1418 - 1448, 2011.

OLSZAK, C. M.; ZIEMBA, E. Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland. **Interdisciplinary Journal of Information, Knowledge, and Management**, 2012.

OTTO, B. Data Governance. **Business & Information Systems Engineering**, 2011.

OTTO, B.; HÜNER, K. M.; ÖSTERLE, H. Toward a functional reference model for master data quality management. **Information Systems and e-Business Management**, v. 10, p. 395 – 425, 2012.

PETERSON, R. Crafting information technology governance. **Information Systems Management**, v. 21, n. 4, 2004.

PORTER, M. E.; HEPPELMANN, J. E. How smart, connected products are transforming competition. **Harvard Business Review**, 2014.

PORTER, M. E.; MILLAR, V. E. How information gives you competitive advantage. **Harvard Business Review**, 1985.

QIU, T. Scanning for competitive intelligence: a managerial perspective. **European Journal of Marketing**, v. 42, p. 814 - 835, 2008.

RAGHUPATHI, W. Corporate governance of IT: A framework for development. **Communications of the ACM**, 50, 8, 94 – 99, 2007.

RANSBOTHAM, S.; NARASIMHAN, S.; LYNN, W. **Managing data in the era of the internet of things**. Webinar MIT Sloan Management Review, 2015.

RAU, K. G. Effective Governance of IT: Design objectives, roles and relationships. **Information Systems Management**, 2004.

RIBARSKY, W.; WANG, D. X.; DOU, W. Social media analytics for competitive advantage. **Computers & Graphics** - Special section on visual analytics. p. 328 - 331, 2014.

ROCKART, J. F. Chief Executives Define their Own Data Needs. *Harvard Business Review*, 1979.

ROCKART, J. F.; FORSTER, N. S. Critical success factors - an annotated bibliography. Sloan School of Management, 1989.

ROSENBAUM, S. Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. **Health Services Research**, v. 45, n. 5, 2010.

ROSS, J. W.; BEATH, C. M.; QUAADGRAS, A. You may not need big data after all. **Harvard Business Review**. 2013.

RUSSON, P. Big Data Analytics - Executive Summary. **TDWI Best Practices Reports**. 4th Quarter. TDWI Research, 2011.

SAMBAMURTHY, V.; ZMUD, R. Arrangements for information technology governance: A theory of multiple contingencies. **MIS Quarterly**, v. 23, n. 2, p. 261–290, 1999.

SCHWARZ, A.; HIRSCHHEIM, R. An extended platform logic perspective of IT governance: managing perceptions and activities of IT. **Journal of Strategic Information Systems**, v. 12, p. 129 - 166, 2003.

SIMONSSON, M.; JOHNSON, P.; EKSTEDT, M. The Effect of IT Governance Maturity on IT Governance Performance. **Information Systems Management**, v., 27, p. 10 - 24, 2010.

SUN, Y.; YAN, H.; LU, C.; BIE, R.; ZHOU, Z. Constructing the Web of Events from raw data in the Web of Things. **Mobile Information Systems**, v. 10, n. 1, p. 105 - 125, 2014.

SUN, Z.; PAMBEL, F.; WANG, F. Incorporating Big Data Analytics into Enterprise Information Systems. **International Federation for Information Processing**. 2015b.

SUN, Z.; ZOU, H.; STRANG, K. Big Data Analytics as a Service for Business Intelligence. **International Federation for Information Processing**, 2015a.

TALLON, P. P.; RAMIREZ, R. V.; SHORT, J. E. The Information Artifact in IT Governance: Toward a Theory of Information Governance. **Journal of Management Information Systems**, v. 30, n. 3, p. 141 - 177, 2014.

TALLON, P. P. Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. **IEEE Computer**, v. 46, n. 6, p. 32–38, 2013.

TALLON, P. P.; SHORT, J. E.; HARKINS, M. W. The Evolution of Information Governance at Intel. **MIS Quarterly Executive**, 2013.

TEO, T. S. H.; CHOO, W. Y. Assessing the impact of using the Internet for competitive intelligence. **Information & Management**, v. 39, p. 67 - 83, 2001.

THOMPSON, N.; RAVINDRAN, R.; NICOSIA, S. Government data does not mean data governance: Lessons learned from a public sector application audit. **Government Information Quarterly**, v. 32, p. 316 - 322, 2015.

TORRES, N. A. **Planejamento de Informática na empresa**. São Paulo: Atlas, 1989.

UBIPARIPOVIC, B.; DURKOVIC, E. Application of Business Intelligence in the Banking Industry. **Management Information Systems**, v. 6, n. 4, p. 023-030, 2011.

VARIAN, H. R. Big Data: New Tricks for Econometrics. **Journal of Economic Perspectives**, v. 28, n. 2, p. 3 - 27, 2014.

VIAENE, S.; VAN DEN BUNDE, A. The secrets to managing business analytics projects. **MIT SLOAN Management Review**, 2011.

WALLER, M. A.; FAWCETT, S. E. Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. **Journal of Business Logistics**, v. 34, n. 2, p. 77 - 84, 2013.

WEBER, K.; OTTO, B.; ÖSTERLE, H. One Size Does Not Fit All—A Contingency Approach to Data Governance. **ACM Journal of Data and Information Quality**, v. 1, n. 1, 2009.

WEILL, P. Don't just lead, Govern: How top-performing firms govern IT. **MIT Sloan School of Management**, Center for Information Systems Research, Working Paper n. 341, 2004.

WEILL, P.; ROSS, J. A matrixed approach to designing IT governance. **MIT Sloan Management Review**, 2005.

WEILL, P.; WOODHAM, R. Don't Just Lead, Govern: Implementing Effective IT Governance. **MIT Sloan School of Management**, 2002.

WILLSON, P., POLLARD, C. Exploring IT governance in theory and practice in a large multinational organisation in Australia. **Information Systems Management**, 26, 2, 98–109, 2009.

WU, X.; ZHU, X.; WU, G.-Q.; Ding, W. Data Mining with Big Data. **IEEE Transactions on Knowledge and Data Engineering**, v. 26, n. 1, p. 97 - 107, 2014.

XU, L. D.; HE, W.; LI, S. Internet of Things in Industries: A Survey. **IEEE Transactions on Industrial Informatics**, v. 10, n. 4, p. 2233 - 2243, 2014.

XUE, L.; RAY, G.; GU, B. Environmental uncertainty and IT infrastructure governance: A curvilinear relationship. **Information Systems Research**, 22, 2, 389–399, 2011.

XUE, Y.; LIANG, H.; BOULTON, W. Information technology governance in information technology investment decision processes: The impact of investment characteristics, external environment, and internal context. **MIS Quarterly**, 32, 1, 67–96, 2008.

YIN, R. K. **Case Study Research: Design and Methods**. 4th Edition. California: SAGE, 2009.

ZHANG, J.; WONG, J.-S.; LI, T.; PAN, Y. A comparison of parallel large-scale knowledge acquisition using rough set theory on different MapReduce runtime systems. **International Journal of Approximate Reasoning**, v. 55, n. 3, p. 896 - 907, 2014.

ZHANG, X.; YANG, L. T.; LIU, C.; CHEN, J. A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud. **IEEE Transactions on Parallel and Distributed Systems**, v. 25, n. 2, p. 363 - 373, 2014.

ZINS, C. Conceptual Approaches for Defining Data, Information, and Knowledge. **Journal of The American Society for Information Science and Technology**, v. 58, n. 4, p. 479 - 493, 2007.