ANTONIO PAULO CARRETTA

RECUPERAÇÃO DE INFORMAÇÃO EM JORNAIS ON-LINE
Percepção sobre Atributos de Pesquisa em Mecanismos de Busca

[versão corrigida]

MESTRADO

Programa de Pós-Graduação em Ciência da Informação

UNIVERSIDADE DE SÃO PAULO

ESCOLA DE COMUNICAÇÕES E ARTES

São Paulo

2015

ANTONIO PAULO CARRETTA

RECUPERAÇÃO DE INFORMAÇÃO EM JORNAIS ON-LINE
Percepção sobre Atributos de Pesquisa em Mecanismos de Busca

[versão corrigida]

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, como exigência para a obtenção do título de Mestre em Ciência da Informação, da Escola de Comunicações e Artes – USP. Área de Concentração: Cultura e Informação. Linha de Pesquisa: Organização da Informação e do Conhecimento.

Orientadora: Profª Drª Vânia Mara Alves Lima.

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÕES E ARTES

São Paulo

2015

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação

Serviço de Biblioteca e Documentação

Escola de Comunicações e Artes da Universidade de São Paulo

Dados fornecidos pelo(a) autor(a)

Carretta, Antonio Paulo

Recuperação de informação em jornais on-line: percepção sobre atributos de pesquisa em mecanismos de busca / Antonio Paulo Carretta. -- São Paulo: A. Carretta, 2015. 117 p.: il.

Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Informação - Escola de Comunicações e Artes / Universidade de São Paulo.

Orientadora: Vânia Maria Alves Lima

Bibliografia

1. Recuperação de Informação 2. Mecanismo de Busca 3. Jornalista 4. Jornal Online 5. Informação Jornalística I. Lima, Vânia Maria Alves II. Título.

CDD 21.ed. - 020

Folha de aprovação

CARRETTA, Antonio Paulo.

Recuperação de informação em jornais on-line: percepção sobre atributos de pesquisa em mecanismos de busca.

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, como exigência para a obtenção do título de Mestre em Ciência da Informação, da Escola de Comunicações e Artes – USP.

Área de Concentração: Cultura e Informação

Linha de Pesquisa: Organização da Informação e do Conhecimento

Orientadora: Profa. Dra. Vânia Mara Alves Lima

Aprovado em: ____/____/____

Banca Examinadora

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Para Conceição, uma grande amiga.

Agradecimentos

À minha família – presente em todos os momentos e grande apoiadora dessa caminhada.

À minha orientadora – à professora Vânia Mara Alves Lima pela confiança e sensível objetividade para nortear os caminhos dessa pesquisa acadêmica.

Aos profissionais, amigos e motivadores – em especial aos jornalistas e pesquisadores: Ana Estela de Sousa Pinto, da Folha de S.Paulo; Edmundo Leite e Carlos Eduardo Entini, de O Estado de S.Paulo; Fabio Ponso, do jornal O Globo.

Aos professores e colegas de pós-graduação da ECA-USP – por compartilharem seu conhecimento e pela visão crítica, em especial as professoras Cibele Araújo Camargo Marques dos Santos e Marilda Lopes Ginez de Lara, pelas contribuições durante a qualificação, e aos colegas jornalistas Mayanna Estevanim e Jade Augusto de Macedo Gola Fernandes.

À CAPES pelo apoio financeiro – bolsa de estudo concedida durante o período de janeiro a agosto de 2015.

Agradeço a todos que, interessados ou não pelo tema, simpatizaram e tiveram paciência quando ouviram minhas reflexões sobre a pesquisa e permitiram, a partir desses pensamentos, construir argumentos.

“Se o conhecimento pode criar problemas, não é através da ignorância que podemos solucioná-los.” (Isaac Asimov)

CARRETTA, Antonio Paulo. **Recuperação de informação em jornais on-line**: percepção sobre atributos de pesquisa em mecanismos de busca. São Paulo, 2015. 117p. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2015.

RESUMO

Estudo analisa questões de organização e recuperação de informação em repositórios de jornais on-line. Destaca aspectos do suporte hipermídia, estrutura informativa do documento digital e gênero do conteúdo da informação jornalística on-line; aborda a noção de memória como atributo de ativação e conexão de informações no contexto da Web; descreve a estrutura básica de mecanismos de busca e traça o perfil de jornalistas no âmbito da convergência digital. Para investigar potenciais dificuldades de pesquisa e recuperação de informação, adota-se pesquisa exploratória para inspeção das interfaces similares de mecanismos de busca de jornais selecionados, nacionais e estrangeiros, e questionário on-line para identificar a percepção de usuários especialistas, jornalistas, sobre o uso de mecanismos de busca interna na rotina de trabalho. Como resultado, discute-se sensibilidades dos atributos de pesquisa, padrões técnicos de tratamento da informação, carências do processo de pesquisa e fatores de satisfação para recuperação de informação em ambiente digital.

PALAVRAS-CHAVE: Recuperação de Informação. Mecanismo de Busca. Jornalista. Jornal Online. Informação Jornalística.

CARRETTA, Antonio Paulo. **Information retrieval in online newspapers: perceptions of search attributes in search engines**. São Paulo, 2015. 117p. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2015.

ABSTRACT

Study examines issues of organization and information retrieval in online newspapers' repositories. Highlights aspects of hypermedia, informative structure of the digital document and some genres of the online journalistic information; It addresses the concept of memory as attribute of activation and connection of information in the Web context; It describes the basic structure of search engines and traces the journalists' profile within the aspect of digital convergence. To investigate potential difficulties of search and information retrieval, exploratory research is adopted to inspect similar interfaces of search engine, in national and foreign selected newspapers; in addition, an online questionnaire is used to identify the perception of expert users about the use internal search engines, based on the work routine of journalists. As a result of these investigations, study shows some sensitivities of search attributes, standards of information processing, research process and satisfaction factors for information retrieval in digital context.

KEYWORDS: Information Retrieval. Search Engine. Journalist. Online Newspaper. Journalistic information.

LISTA DE TABELAS

Tabela 1 – Filtros de pesquisa: avaliação geral.....	76
Tabela 2 – Filtros de pesquisa: avaliação por região.....	77
Tabela 3 – Resultados de pesquisa: avaliação geral.....	78
Tabela 4 – Perfil.....	80
Tabela 5 – Uso do mecanismo de busca em pesquisas diárias.....	84
Tabela 6 – Conteúdo e categorias de pesquisa.....	85
Tabela 7 – Recursos da ferramenta de pesquisa.....	86
Tabela 8 – Resultados de pesquisa e visualização da informação.....	87
Tabela 9 – Percepção como pesquisador.....	88
Tabela 10 – Questão 24.....	89
Tabela 11 – Questão 25.....	89

LISTA DE ILUSTRAÇÕES

Figura 1 - Arquitetura de um mecanismo de busca simples.....	55
Figura 2 – Representação do processo de recuperação de informação.....	57
Figura 3 – Formulário de busca do site Folha.com.....	59
Figura 4 – Pesquisa de conteúdo na Web.....	66
Figura 5 – Pesquisa de conteúdo interno de um site.....	66
Figura 6 – Exemplos de áreas de inspeção em mecanismos de busca.....	74
Quadro 1 – Ocorrência de metadados em jornais on-line.....	40
Quadro 2 – Inspeção de mecanismos de busca de sites jornalísticos.....	75
Quadro 3 – Esquema de análise do questionário.....	79

SUMÁRIO

1 Introdução	1
1.1 Objeto [Justificativa e Problema]	3
1.2 Objetivos	6
1.3 Quadro Referencial Teórico.....	7
1.4 Procedimentos Metodológicos	13
 2. Perfil do Jornalista em Ambiente Digital	 21
 3. Aspectos da Informação Jornalística On-Line	 29
3.1 Suportes em Convergência	32
3.2 Estrutura Informativa [Conteúdo, Gênero e Documento On-line] ...	33
3.5 Memória (Digital), Recuperação de Informação e Conectividade	46
 4. Mecanismo de Busca e Recuperação de Informação na Web	 52
4.1 Características Gerais de um Mecanismo de Busca	52
4.2 Aspectos da Recuperação da Informação na Web	56
4.3 Usabilidade e Avaliação da Qualidade de Mecanismos de Busca	64
 5. Análise dos Resultados	 72
5.1 Inspeção de Interfaces Similares: Mecanismos de Busca	72
5.2 Questionário: Recuperação de Informação em Jornais On-Line	79
 6. Considerações Finais	 93
 Referências	 98
 Apêndice	 111

1. INTRODUÇÃO

Desde as primeiras gerações de jornais digitais, uma significativa mudança tem ocorrido na produção, distribuição e consumo da informação jornalística disponível na Web. Segundo Pariser (2012), está ocorrendo uma queda constante no valor para produção e distribuição de qualquer tipo de mídia (palavras, imagens, vídeos e áudio). Esta redução de valor, cada vez mais próxima do custo zero, ampliará a oferta de informações na Web e, conseqüentemente, provocará uma constante necessidade de escolhas para filtrar o gigantesco conteúdo gerado. Como efeito inicial dessa mudança nas condições de produzir informações, observamos uma maior apropriação de ferramentas de busca pela dimensão de sua aplicação e uso: cerca de 40%¹ da população mundial é usuária da internet e mais de 90%² utiliza frequentemente um mecanismo de busca para obter informações/notícias. Além disso, o acesso e consumo de notícias no mundo é ampliado pelo uso de diferentes plataformas e dispositivos, por exemplo, em grandes centros urbanos do Brasil, estudo revela que 8% dos brasileiros acessa notícias via tablets e 18% por smartphones (NEWMAN; LEVY, 2014).

Nesse atual panorama, surgem no jornalismo on-line novas tendências de organização da informação (divulgação e classificação de conteúdo), projetos de digitalização de documentos, incorporação de arquivos retrospectivos, personalização da disseminação de notícias e, como resultado dos esforços para melhorar e agilizar a capacidade de tecnologia de pesquisa: inovações em mecanismos de busca e modificações na experiência de pesquisa para recuperação e divulgação de acervos digitais. Estes indicadores de evolução (técnica ou estratégica) são percebidos na grande maioria de jornais, nacionais ou estrangeiros, com produção de conteúdo digital.

No entanto, mesmo observando que as ferramentas de busca são frequentes em sites com permanente produção e arquivamento de informação, os serviços para recuperação de informação possuem reduzidos, irregulares e insatisfatórios atributos de pesquisa. Conseqüentemente, profissionais que atuam em redações on-line lidam

¹ Cf. INTERNET LIVE STATS. Disponível em: <www.internetstats.com> Acesso em: 20 out. 2014

² Cf. GRUPO DE MÍDIA SÃO PAULO. **Mídia Dados Brasil 2014**. São Paulo: Grupo de Mídia São Paulo, 2014. Disponível em: <www.gm.org.br> Acesso em: 20 out. 2014

diariamente com este problema que, carente de uma análise mais profunda dentro do contexto profissional, necessita de uma avaliação sobre seu impacto na rotina de trabalho. Portanto, observada pela relação entre mídia impressa e digital, que provoca uma mudança de cultura do saber profissional, o principal argumento deste projeto é ampliar estudos em tecnologia de pesquisa e acesso à informação. Para isso, a proposta do estudo está delimitada na avaliação dos atributos de pesquisa da informação jornalística on-line, comparação dos mecanismos de busca de sites jornalísticos e perfil dos profissionais-pesquisadores (relação ferramenta-usuário), inseridos no atual ambiente de web jornalismo e durante a prática de pesquisa de conteúdo e recuperação de notícias.

Esta pesquisa insere-se na área de concentração denominada Cultura e Informação, mantida pelo Programa de Pós-Graduação em Ciência da Informação (PPGCI) da ECA-USP, que promove reflexões epistemológicas, definições teóricas e propostas metodológicas voltadas ao estudo da informação no contexto social-cultural.

Dentro deste contexto, o estudo que aqui desenvolvemos encontra-se localizado na linha de pesquisa do PPGCI ECA-USP designada Organização da Informação e do Conhecimento. Essa linha mantém um foco nas reflexões sobre organização do conhecimento e da informação e de sua circulação para fins de acesso, recuperação e uso.

Por este foco, a abordagem exploratória deste estudo visa ampliar as reflexões sobre questões da Ciência da Informação voltadas à organização do conhecimento que, de acordo com San Segundo (2013), enfrentam modificações paradigmáticas impulsionadas por tecnologias de informação, dinâmicas da interação humano-computador e o entorno digital que, conseqüentemente, afetam o processo de recuperação de informação e caminham para uma nova concepção da representação do conhecimento.

1.1. OBJETO [Justificativa e Problema]

Diariamente, sites de notícias produzem milhares de informações que são publicadas e divulgadas em suas diversas áreas editoriais. Para produção desse conteúdo, um grupo grande de jornalistas e profissionais de apoio técnico (webdesigners e programadores) trabalha para publicar e atualizar este material, com a maior rapidez possível. Além disso, para o acompanhamento diário da notícia, a organização da homepage (página principal do site) é pensada como um guia de conteúdo que facilita o destaque das informações de grande repercussão ou ineditismo. Podemos também adicionar neste processo as recentes tendências de interação e divulgação do noticiário que agregaram recursos de organização, por exemplo, o uso de páginas especiais para cobertura de eventos, como eleições e competições esportivas, ou rankings do conteúdo, que exibem as notícias mais lidas do dia ou mais curiosas ou mais enviadas pelo internauta. Considerando esta situação, parece simples e eficiente, para jornalistas e redatores de jornais on-line, localizar as informações mais importantes divulgadas no mesmo dia.

Agora, consideremos que será necessário recuperar uma informação publicada na semana anterior, ou no mês anterior, ou no ano passado. Em princípio, um profissional de jornal vai utilizar o mecanismo de busca oferecido pelo site. O mesmo também vai ocorrer com os demais jornalistas que alimentam o site com informações.

Abrimos parênteses aqui para esclarecer que há uma impressão errada de que o profissional de redações on-line possui um recurso de busca diferente do leitor comum e, com vantagem sobre ele, utilizaria recursos avançados de busca para sua rotina de trabalho. Isto não ocorre nos jornais brasileiros. É fato que para um jornalista que lida com a busca de notícias retrospectivas, cuja origem de produção é apenas on-line, o ambiente de navegação e seus recursos são os mesmos do leitor de um portal de notícias que procura uma informação qualquer. Para este profissional recuperar algum dado que o ajude, por exemplo, a compor um texto de memória, será necessário iniciar uma busca no conteúdo on-line do jornal. É neste momento que surgem imprecisões e falhas de recuperação.

Em meio a milhões de páginas com versões e atualizações sistemáticas, links antigos (fora de padrão) que ficaram inacessíveis³ e informações antigas substituídas por novas, tanto o leitor comum, quanto o profissional acostumado com a estrutura do site e do fluxo de informações, encontrarão dificuldades em explorar, localizar e determinar a relevância das matérias apresentadas nos resultados de busca. Na melhor das hipóteses, o mecanismo de busca terá uma opção para pesquisa avançada, mas os atributos definidos pelo sistema e a identificação estrutural do documento permitirão um resultado adequado? Será necessário optar pelo uso de um buscador de caráter global (como Google ou Google News)? A pesquisa precisará ser refeita e reformulada muitas vezes? O usuário, no caso deste estudo, o profissional de jornal, desistirá da busca e utilizará outras estratégias, ou outros caminhos de navegação do site, para encontrar a informação? Todas estas questões orbitam o problema da pesquisa: a exploração da produção existente em jornais on-line.

Considerando este quadro, levantamos duas indagações (e respectivas hipóteses) que acreditamos precisariam ser respondidas e justificam o estudo proposto:

³ Independente do tipo de falha técnica, expurgo automático, erro de arquivamento em migração de páginas ou ausência de metadados, existem lacunas em repositórios de sites jornalísticos que a tecnologia não resolveu. Um bom exemplo desse fato é descrito pelo jornalista Edmundo Leite que, a partir de tentativas para localizar suas próprias matérias, produzidas exclusivamente em ambiente online, percebeu o desaparecimento de textos desenvolvidos durante sua trajetória profissional no portal de notícias do Estadão. Recorrendo ao seu arquivo pessoal, parte do material foi recuperada, e o jornalista iniciou registro das ocorrências de textos perdidos em seu blog pessoal (<https://edmundoleite.wordpress.com/about/>) que, posteriormente, ganhou espaço no blog profissional “Memória, Gente e Lugares” (<http://brasil.estadao.com.br/blogs/edmundo-leite/>). Um dos relatos que ilustra bem essa situação é a cobertura do atentado terrorista ocorrido em 11 de setembro de 2001, nos EUA. Segundo observações de Leite (2011), o conteúdo da cobertura on-line se perdeu quase por completo e poucos textos ainda podem ser encontrados como originalmente publicados na Web. Leite (2011) também reconhece o sentido paradoxal da situação que, apesar de todos os avanços tecnológicos, torna “mais fácil encontrar uma informação publicada por um jornal em papel há um século do que achar o que foi publicado exclusivamente no formato digital” na data do evento mencionado. Além disso, conclui que, “apesar de ser possível encontrar parte dos textos da cobertura digital” do Estadão, perdeu-se também “todo o esforço multimídia originalmente realizado naqueles dias.”

- 1) Questão: Considerando a crescente formação de repositórios de informação jornalística (entendidos aqui apenas como um conjunto de documentos digitais, gerados em ambientes on-line), podemos acreditar que os mecanismos de busca utilizados para recuperação destas informações possuem estruturas e atributos de pesquisa adequados?

Problema 1: Considerando que os atuais mecanismos de busca de grandes jornais não possuem atributos suficientes para recuperar as informações armazenadas em áreas-repositórios, podemos admitir que o uso dos sistemas de busca local de jornais on-line está prejudicado e exige que diferentes estratégias de pesquisa sejam adotadas pelo usuário (jornalista) para contornar o problema: recuperar informação em um ambiente que contém milhões de páginas, produzidas anualmente pela equipe de jornalismo.

- 2) Questão: A ferramenta para recuperação de informação proporciona uma resposta pertinente e eficaz para satisfazer o usuário (jornalista) em sua rotina de trabalho e sua dinâmica de pesquisa?

Problema 2: Considerando que os profissionais da equipe de jornalismo on-line possuem uma percepção clara das dificuldades que encontram para atender suas necessidades de informação, observadas e declaradas pelo uso dos buscadores locais como ferramenta de trabalho, podemos admitir que ao identificar um padrão em seu comportamento de pesquisa é possível indicar falhas nos atributos do mecanismo de busca e, conseqüentemente, se a ferramenta de pesquisa satisfaz (ou não) suas necessidades de informação e precisa ser melhorada.

Atentos ao problema da exploração de conteúdo dos sites jornalísticos, finalizamos a justificativa acrescentando que as duas questões norteiam os experimentos desenvolvidos neste trabalho e pontuam o embasamento teórico, assim como métodos de pesquisa, observações, análises e tópicos selecionados.

1.2. OBJETIVOS

A condução de explorações e experimentações propostas neste estudo têm por objetivo geral analisar o comportamento de pesquisa de usuários especialistas para recuperação de informação em jornais on-line. Para isso, os objetivos específicos do estudo são: analisar atributos de pesquisa da ferramenta de busca; identificar necessidades de informação e tarefas de pesquisa do usuário; reconhecer dificuldades no processo de pesquisa; identificar características de usabilidade da ferramenta e satisfação do usuário. Todos estes detalhes têm como foco a percepção e experiência de jornalistas sobre o uso desta ferramenta em sua prática de pesquisa.

Neste estudo o mecanismo de busca será entendido como uma ferramenta (interface) para comunicação entre objeto (informação) e indivíduo (usuários que fazem parte de um grupo social formado por profissionais atuando em redações de jornais).

Como resultado final da pesquisa, apresentamos uma avaliação dos atuais atributos de informação disponíveis para pesquisa e recuperação de conteúdo jornalístico e uma síntese do comportamento de pesquisa dos usuários especialistas. Estas propostas visam à identificação de parâmetros de avaliação da ferramenta e indicadores (da satisfação, limitação ou modificação) das práticas de uso, que permitam reduzir problemas nas estratégias de recuperação de informação e propor recomendações para melhorar a capacidade informativa e usabilidade da ferramenta de trabalho.

1.3. QUADRO REFERENCIAL TEÓRICO

Iniciamos esta primeira reflexão teórica da pesquisa com um breve posicionamento de amparo ao objeto deste estudo e a intenção de promover associações com o pensamento da chamada “era da informação” e a Ciência da Informação. Este estudo perpassa necessariamente disciplinas diversas que, ao lado de abordagens afeitas diretamente à Ciência da Informação, colaboram na análise de dados empíricos e no exame de temas e teses que permitem observações e relações da investigação com aspectos da informação jornalística on-line; recuperação da informação (por meio de um mecanismo de busca); noções de usabilidade de busca e perfil do jornalista em ambientes digitais, considerando a prática profissional e comportamento de pesquisa focado no uso de mecanismo interno de busca em jornais on-line.

Feito este esclarecimento, como primeira abordagem, tratamos de aspectos que envolvem tecnologia, informação e os paradigmas de referência relacionados. Numa década repleta de informação, conteúdos diversos são estocados em banco de dados, que, a partir de seu uso, poderiam facilitar a geração de novo conhecimento. Entretanto, a

[...] massa de dados digitais disponíveis se infla o tempo todo. E quanto mais ela cresce, mais é preciso estruturá-la, cartografá-la, criar uma matriz com estradas expressas e avenidas lógicas; mais as interfaces para a caça eficaz e o garimpo furioso devem ser aperfeiçoadas (LÉVY, 2002, p. 108).

Neste sentido, entre as ferramentas que surgem para facilitar este garimpo de conteúdo estão os mecanismos de busca de informação, geradores de novos hábitos de pesquisa.

O contexto (usuário/informação/recuperação) que abordamos neste estudo invade o tradicional tratamento de informação e expande para o contexto digital (de rede, hiperlinks e navegação) que conecta informações e pessoas. De forma mais ampla, por entender como resultado final uma apropriação transformadora de informação pelo usuário, a fundamentação para desenvolvimento deste estudo leva em consideração a tecnologia como um dos caminhos de transformação social. De forma complementar,

mas pela relação de informação e tecnologia, Castells (1999, p. 78-79) apresenta neste sentido o “paradigma da tecnologia da informação” cujas características são: “tecnologias para agir sobre a informação, não apenas informação para agir sobre a tecnologia”; a influência da tecnologia na atividade do homem por meio da informação, indicador da “penetrabilidade dos efeitos das novas tecnologias”; a complexa interação em um sistema de informação resultado da “lógica das redes”, presente nestas novas tecnologias; a “flexibilidade” do sistema de redes para melhorar a comunicação e a “convergência de tecnologias” como característica da revolução tecnológica, neste caso é observada pela relação entre mídia impressa e digital, que provoca uma mudança de cultura do saber profissional.

A literatura registra que o aparecimento da internet introduziu novos paradigmas que esclarecem os efeitos da tecnologia da informação sobre o homem e são apresentados, discutidos e desdobrados em diversas áreas do conhecimento. Para posicionamento histórico do estudo, acompanhando esta evolução de opiniões e modelos de análise, após a década de 70, a Ciência da Informação, por sua vez, introduziu conceitos sobre informação e a recuperação de informações, que refletem as novas perspectivas tecnológicas e sua influência sobre a dinâmica que estuda a relação entre humano-computador-informação. Dentro desta relação, para Ciência da Informação o termo “recuperação de informação” (RI), de acordo com Capurro e Hjørland (2007), possivelmente é uma das noções mais importantes da área e provoca uma série de discussões que entrelaçam questões sobre armazenamento, gestão e recuperação de conteúdo (que observado e classificado por diferentes pontos de vista pode ser identificado como dado, informação ou documento). Além disso, diversas visões (por exemplo, semiótica ou cognitiva) e teorias da informação (de ordem social, cultural ou científica) presentes na área, acrescentam novos conflitos na definição de informação e, conseqüentemente, no sentido dado ao processo de recuperação.

Frente à diversidade de vertentes existentes, utilizamos como ponto de partida as considerações de Capurro e Hjørland (2007) que, em seu estudo sobre o conceito de informação, expõem sua percepção sobre a distinção entre informação como objeto (neste caso dotado de qualquer forma, por exemplo, um conjunto de bits) e informação como conceito subjetivo (apoiado na interpretação de um agente

cognitivo). Para os autores, surgiu um deslocamento dos atributos apresentados ao estado de coisa/objeto (da informação) para um estado de relevância, muito mais difícil de compreender porque exige um processo de interpretação determinado por contextos sociais e culturais e não apenas pela estrutura. Desta forma,

A medida que os sistemas de informação tornam-se mais globais e interconectados, a informação implícita [ou seja, aquela que precisa ser interpretada e possibilita definir a relevância de um documento e quais perguntas do usuário ele poderá responder] é, muitas vezes, perdida. Esta situação desafia a CI [Ciência da Informação] a ser mais receptiva aos impactos sociais e culturais dos processos interpretativos e, também, às diferenças qualitativas entre diferentes contextos e mídias (CAPURRO, HJORLAND, 2007).

Esta percepção do deslocamento de atributos da informação, desenvolvido e analisado pelo viés epistemológico de Capurro (2003) e explicado por Almeida e outros (2007), propõe que a Ciência da Informação pode ser caracterizada pela existência de três paradigmas: *físico, cognitivo e social*. O paradigma físico (com origem presumida em 1945, quando Vannevar Bush imaginou, em seu ensaio “As We May Think”, o processamento e recuperação de informação por um sistema mecânico) possui um sentido mais técnico da informação e está centrado em sistemas informatizados. Influenciado pelo desenvolvimento tecnológico, este modelo possui limitações que concentram seu objetivo na aplicação de métodos voltados para gerir informação (entendida como coisa ou objeto) de uma forma mais eficiente.

Ultrapassando estas limitações, na década de 70 (marcada pela Conferência de Copenhagen, onde teorias de ordem racionalistas foram confrontadas com perspectivas psicossociais focadas no usuário) uma nova abordagem é apresentada como: paradigma cognitivo. O comportamento e percepções do usuário no processo de recuperação de informação dão o tom desta análise. Embora esta nova visão acrescente um importante aspecto comportamental que determina as necessidades do usuário, não é considerado o contexto social em que os usuários estão inseridos e,

a partir deste limite, surge um terceiro paradigma nas investigações da Ciência da Informação contemporânea: o paradigma social (ALMEIDA *et al.*, 2007).

Capurro (2003) esclarece que esta evolução de investigações não se dá de forma linear, mas surge entre os limites entre um e outro paradigma, com funções explicativas e complementares. Dito de outra forma, o paradigma físico estabelece uma relação enorme com o objeto. A interação com este objeto (informação) é quase mecânica e atenua a ação da consciência sobre ele. O posicionamento cognitivo maximiza a operação mental realizada pelo indivíduo na busca desse objeto (informação) e fortalece a análise dos processos interiores que relacionam objeto e indivíduo (informação e usuário). O paradigma social se apresenta como ações exteriores que envolvem este indivíduo no processo de atender suas necessidades de informação.

Segundo Almeida *et al.* (2007), esta abordagem sócio-cognitiva, introduzida pelo teórico Birger Hjørland, foca a interação entre usuário (indivíduo) e ambiente social (organização). O paradigma social é denominado como “análise de domínio” e a informação melhor compreendida quando avaliada por “domínios do conhecimento” relacionados a “comunidades discursivas”, entendidas aqui como grupos sociais formados pela sincronia de seus pensamentos, linguagem e conhecimento. Desta forma, a informação se configura como um fenômeno social coletivo, estruturado pelo conhecimento e memória de comunidades distintas. Neste sentido,

Informação não é algo que comunicam duas cápsulas cognitivas com base em um sistema tecnológico, visto que todo sistema de informação está destinado a sustentar a produção, coleta, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso de conhecimentos e deveria ser concebido no marco de um grupo social concreto e para áreas determinadas (CAPURRO, 2003).

Feito este posicionamento, esclarecemos que este estudo usa como referência teórica as proposições mais fortemente estabelecidas por um paradigma social. Além disso, reconhecemos que o paradigma tecnológico da atualidade, apontado por San Segundo (2013), apresenta desafios gerados por tecnologias da informação e pela interação humano-computador e, por isso, demanda novas formas de representação e

organização do conhecimento para mediação do acesso à informação. Definidas estas posições, concentramos as abordagens do estudo em um grupo de profissionais (jornalistas inseridos em um domínio do conhecimento) participantes de uma comunidade discursiva e analisados a partir de sua relação com uma ferramenta (mecanismo de busca) para recuperação de um objeto (informação) produzido pelo grupo.

Com base nestes aspectos, para definir os rumos empíricos da pesquisa, nos inspiramos no quadro de integração do processo de busca e recuperação da informação que, em modelo apresentado por Ingwersen e Järvelin (2005), cinco pontos complexos podem ser observados:

- Objeto informacional, composto por conhecimento representado, estruturas de dados e texto completo;
- Interface, que possui uma função de interação e opera em nível linguístico;
- Tecnologia de Informação, mecanismo que opera por estruturas lógicas e algorítmicas;
- Usuário (ator cognitivo), que busca por informação em um espaço de interesses pessoais, necessidades informacionais diversas, situações problema, tarefas de trabalho ou pesquisa;
- Contexto, organizacional, social e cultural, que estabelece domínios, metas, situações, tarefas de trabalho, estratégias, interesses e preferências.

Por último, salientamos que o estudo explora, principalmente, pontos de contato entre a Ciência da Informação e as Ciências da Comunicação. Dentro do contexto da “era da informação” e dos ambientes digitais, visualizamos esta aproximação entre as áreas a partir de três elementos indicados por Miège⁴ (2000, citado por SILVA, 2006, p.108) e propostos para necessidade de “delinear um programa de trabalho para as ciências da informação e comunicação”:

⁴ MIÈGE, Bernard. **O pensamento comunicacional**. Petrópolis: Vozes, 2000.

- “a articulação entre dispositivos tecnológicos da comunicação e a produção de mensagens e do sentido”, que associamos neste estudo pela relação dispositivo-informação (mecanismo de busca-informação jornalística);
- “a ‘inserção social’ das tecnologias e, principalmente, a atividade dos usuários-consumidores no aperfeiçoamento dos dispositivos”, no caso desta pesquisa, observada pela relação usuário-dispositivo (jornalista-mecanismo de busca);
- “a atenção aos ‘procedimentos’ de escrita das mensagens (icônicas, sonoras, gráficas...) e das condições que presidem sua concepção e realização”, que associamos ao ambiente digital e as condições de hipermídia e hipertexto para produção de informações jornalísticas.

Estes elementos constituem três dimensões de relação e análise da pesquisa: dispositivo (mecanismo de busca e sua interface de pesquisa), usuário (jornalista) e informação (jornalística).

1.4. PROCEDIMENTOS METODOLÓGICOS

Este é um estudo essencialmente exploratório e descritivo. O processo de construção teórico-empírico da pesquisa está dividido em quatro grandes etapas metodológicas, que são resumidamente apresentadas para visão geral da proposta e, na sequência, detalhadas da seguinte forma:

BASE TEÓRICA

Etapa 1: Análise de conceitos para fundamentação teórica da relação entre processo de pesquisa, comportamento e experiência no uso de informação por meio de uma ferramenta de busca, assim como aspectos da informação jornalística e análise do perfil do usuário especialista (jornalista) em sua prática de pesquisa.

BASE EMPÍRICA

Etapa 2: Análise de interfaces Similares - Elaboração de quadro com características dos buscadores dos jornais que são referência para pesquisa dos usuários estudados. Modelos selecionados destacando recursos de recuperação de informação.

Etapa 3: Elaboração de questionário (on-line) para identificação do comportamento de pesquisa, de jornalistas, assim como as impressões sobre o uso de buscadores como ferramentas de trabalho.

Etapa final: Análise dos questionários pontuando perfis de usuários, tarefas, tendências de pesquisa, recuperação de informação e usabilidade dos mecanismos de busca.

Detalhamento das Etapas:

Etapas 1 (BASE TEÓRICA): As análises e reflexões expostas neste trabalho refletem que a complexidade estrutural da web exige enfoques multidisciplinares. Desta forma, para facilitar o entendimento do quadro teórico, conceitos operacionais são extraídos das áreas de Comunicação, Arquitetura da Informação e Computação. Estes conceitos favorecem pontos teóricos de contato com a Ciência da Informação e são explorados e aplicados na elaboração de dimensões de análise e esquemas explicativos de aspectos do processo de recuperação de informação, usabilidade de ferramentas de pesquisa e, como enfoque principal deste estudo, o usuário e sua percepção sobre os mecanismos de busca.

Visando indicar qual etapa de desenvolvimento do jornalismo na web inserimos a pesquisa, adotamos a classificação de 3ª geração do webjornalismo, referente a produção originalmente em formato multimídia e hipertexto, proposta por Machado, Borges e Miranda (2003), diferenciada das demais gerações de webjornais e seus modelos de produção de conteúdo. Considerando a confiabilidade adquirida pelas versões de jornais on-line, decorrente da forma de apuração e acuidade das notícias dos veículos impressos, destacamos a produção de informação jornalística em meio digital como fonte de pesquisa do usuário profissional, o jornalista (MACHADO, 2003), e do papel da ferramenta (ou sistema) de recuperação de informação oferecido pelo do veículo de comunicação (MANNARINO, 2000).

Para avaliar o comportamento de jornalistas, que muito recentemente lidam com a fusão ou aproximação que os jornais têm provocado entre redações da versão impressa e on-line, coletamos dados que esclarecem a rotina de tarefas do jornalista e, também, foram aplicados para elaboração de questionário on-line descrito na etapa 3. Para isso, utilizamos bibliografia da área de comunicação, avaliação e descrição de informações extraídas de pesquisas de perfil profissional, manuais de práticas jornalísticas, sites de associações profissionais e empresas jornalísticas. Nesse caminho, utilizamos pesquisas recentes que apresentam características concretas de profissionais do jornalismo e as sintonizam com as mudanças provocadas pelo surgimento e evolução da internet.

Sobre os aspectos do comportamento de pesquisa, observamos fundamentos da profissão baseados apenas no processo de produção de notícias que, basicamente, está dividido em 4 etapas: propor uma ideia (pauta), recolher informações que fundamentem a pauta, organizar e hierarquizar estas informações e escrever (PINTO, 2009; MEYER, 1993). Um processo que implica em informações que geram novas informações e cuja constante é a pesquisa. Este aspecto justifica o recorte de análise proposto, delimitado ao momento de produção, por ser essencialmente voltado para checagem e busca por informação de contextualização que, por sua vez, exige do jornalista o uso mais intenso do mecanismo de busca do site para explorar a produção existente. Dentro desse contexto, são consideradas também as novas habilidades experimentadas no jornalismo digital (CALDERÍN; ROJANO, 2006).

Neste sentido, introduzimos também uma abordagem focada no usuário e no processo de pesquisa (informação, consumo e uso), com base teórico-conceitual e análises relacionadas aos estudos da Ciência da Informação, vista pela perspectiva de um modelo (INGWERSEN; JÄRVELIN 2005) que, muito embora esteja focado na ampliação do ponto de vista cognitivo, propõem uma integração de análises dos estudos sobre o processo de busca (proposto em uma dimensão humana pela Ciência da Informação) e recuperação de informação (com suas bases próximas da Ciência da Computação).

Especificamente sobre o mecanismo de busca e a recuperação de Informação na Web, tratamos da anatomia da ferramenta e aspectos da usabilidade, bem como fatores que afetam uso de mecanismos de busca (NIELSEN; LORANGER, 2007) e características de avaliação de buscadores internos, propostos por León, Serrano (2007) e Vicens Arasanz (2008). De forma complementar, a partir da abordagem de Pérez Gutiérrez (2000), refletimos brevemente sobre a expressão de busca, uso de operadores booleanos e linguagem de pesquisa para recuperação de informações.

Para fecharmos a relação usuário-mecanismo de busca-informação, são considerados aspectos da informação jornalística on-line que destacam a estrutura informativa, gênero e atributo de memória do conteúdo de seus repositórios on-line.

Etapas 2 a 3 (BASE EMPÍRICA): como proposta metodológica, este estudo adota investigação de interfaces similares e hábitos de pesquisa do usuário pelo formato de questionário on-line. Como referencial, refletimos sobre o método proposto por MARTINEZ (2002) para análise de requisitos de usabilidade de sites (interface, funcionalidades e necessidades, tarefas e perfil do usuário), que foi aplicado ao desenvolvimento da Agência Universitária de Notícias do Departamento de Jornalismo e Editoração da Universidade de São Paulo. Neste modelo selecionamos pontos para construção da investigação dos hábitos de pesquisa de profissionais de redação de notícias on-line, que utilizam o mecanismo interno de busca do site de notícias, e adotamos as seguintes etapas:

Análise de Interfaces Similares (Etapa 2)

Objetivo: Inspeção e análise (da estrutura de busca, recursos de recuperação de informação e metadados de código fonte) de sites jornalísticos.

Técnica de Análise: observação de páginas de busca e resultado considerando elementos clássicos da interface para recuperação de informação. Elaboração de formulário para análise detalhada dos modelos selecionados, destacando recursos de recuperação de informação e características de busca. Apresentação de quadro geral da inspeção dos sites de jornais selecionados como base para pesquisa.

Para isso, será utilizado o formato de inspeção de usabilidade, que dispensa o uso de um grupo de avaliadores e se constitui como passo inicial para desdobramentos de técnicas e análises dos mecanismos de recuperação de informação na Web. Alguns recortes foram estabelecidos para favorecer esta inspeção:

- Adoção de três sites jornalísticos destacados de três áreas e selecionados segundo os seguintes critérios: circulação (em papel) superior a 250 mil exemplares, com mais de 15 anos de produção on-line, com reconhecida confiabilidade no mercado e representação de extrato geográfico importante (EUA, Europa e Brasil).

- A inspeção baseada nos principais elementos de identificação da produção editorial, considerados aqui como essenciais para organização e recuperação da informação jornalística. Estes elementos estão presentes em cada unidade de informação (texto publicado) produzida: data e período (de publicação da notícia), título, autoria da matéria, área editorial (editoria que produz e agrupa texto de mesma linha temática).
- Como a atividade profissional em redações on-line está concentrada na produção e localização de textos. Apenas unidades de informação em formato textual (da produção editorial) são consideradas. Outros formatos, como áudio e vídeo, são registrados como uma categoria de seleção, mas não são identificados individualmente como elementos de pesquisa.

A inspeção utiliza os elementos de identificação como parâmetros de análise e identificação dos recursos existentes para uma busca simples ou avançada, sem diferença entre elas, permitindo uma visão geral da dinâmica de pesquisa em sites jornalísticos.

Elaboração de Questionário (Etapa 3)

Objetivo: avaliação da interação dos usuários (jornalista) com interfaces de busca, dinâmica de uso, realização de tarefas e necessidades profissionais durante processo de pesquisa.

Técnicas de análise: produção de questionário on-line (formato com questões fechadas: uso justificado pela capacidade e agilidade da ferramenta para aplicação e levantamento de dados sobre hábitos, comportamento e opinião) para composição do perfil do usuário, identificação do comportamento de pesquisa, assim como as impressões sobre a realização de tarefas e a satisfação no uso de buscadores como

ferramentas de trabalho. De abrangência nacional, o levantamento foi realizado em 3 grandes veículos de comunicação (Folha, Estadão e Globo), durante o mês de junho de 2015. Para envio do questionário foram realizados contatos preliminares no primeiro trimestre de 2015 e pré-teste do modelo de questionário realizado no mês de maio. A estratégia de aplicação do questionário, de preenchimento voluntário e anônimo, foi combinada com profissionais dos três jornais, que criaram um canal de comunicação com as redações e enviaram, em duas rodadas, mensagem-convite para participação da pesquisa. Os resultados foram obtidos em amostragem por conveniência (acidental ou casual), de caráter não probabilístico, na qual o pesquisador seleciona membros da população mais acessíveis e permite aplicação rápida e de baixo custo. Embora não permita análises conclusivas, este tipo de amostra, frequentemente utilizada para geração de ideias em pesquisas exploratórias, oferece base para obtenção de *insights* e ajustes em hipóteses.

O questionário foi elaborado com 25 questões fechadas (formato binário, múltipla escolha e escalonadas para registro de opinião) e uma aberta para comentário espontâneo (ver Apêndice A). As questões foram agrupadas em seis partes:

Questões 1 a 6 - Perfil profissional, que inclui dados demográficos;

Questões 7 a 12 - Uso do mecanismo de busca em pesquisas diárias do jornalista, baseado em abordagens do jornalismo de precisão, proposto por Meyer (1993);

Questões 13 a 18 - Conteúdo e categorias de pesquisa, baseado em manuais de prática jornalística, proposto em Pinto (2009);

Questão 19 – Recursos da ferramenta de pesquisa, opções baseadas no quadro de inspeção da etapa 2;

Questão 20 – Resultados de pesquisa e visualização de informação, opções baseadas no quadro de inspeção da etapa 2;

Questões 21 a 26 – Percepção como pesquisador, formuladas para qualificar o usuário, identificar reações no uso do mecanismo de busca e, com base nas duas questões norteadoras da pesquisa, obter opinião sobre atributos de pesquisa, apresentados em questões anteriores, e grau de satisfação no uso da ferramenta.

Além das fontes de apoio mencionadas, as questões foram desenvolvidas considerando alguns modelos de estrutura presentes nas seguintes referências:

- comportamento de pesquisa - baseado em pesquisa do Pew Research Center (2008) sobre consumo de notícias on-line.
- uso de mecanismo de busca como ferramenta de trabalho - pesquisa Harvest Digital (2006) e iProspect (2004), sobre uso de mecanismos de busca.
- satisfação no uso da ferramenta de pesquisa e recursos de usabilidade baseado em modelos reunidos por Perlman (2015), que sugere questionários demográficos com medidas organizadas em fatores específicos da interface. A medida de satisfação geral dos usuários com a interface é por escala de pontos e os questionários são concebidos para configuração de acordo com as necessidades de cada análise de interface, incluindo apenas as seções que são de interesse para o usuário.

Os pontos de observação da pesquisa e referências para ordenação/agrupamento das perguntas estabelecidas no questionário são:

- a) hábitos de pesquisa e satisfação com (interface) mecanismo de busca
 - ✓ frequência e preferência de uso
 - ✓ avaliação de importância de atributos de pesquisa
 - ✓ recursos de busca mais utilizados
 - ✓ visualização de resultados

b) fatores que afetam a severidade do problema de busca

- ✓ frequência > % de usuários com dificuldade
- ✓ impacto > dificuldades e tempo gasto
- ✓ persistência > dificuldade contínua

Etapas Final

Apresentação das análises propostas (interfaces similares, perfil de usuário e questionário dos hábitos de pesquisa), quadro de avaliação e recomendações possíveis (baseado em princípios gerais da avaliação heurística) identificando correspondências entre os hábitos de busca e recuperação de informação dos usuários analisados, as tendências de pesquisa e uso dos mecanismos de busca.

2. PERFIL DO JORNALISTA EM AMBIENTE DIGITAL

Traçar um perfil profissional que exponha características do jornalista em ambientes digitais nos parece um exercício complexo e delicado, já que as competências e habilidades acompanham uma evolução da prática jornalista e sua relação com aparecimento de novas tecnologias de comunicação. O recorte que inserimos neste estudo é de um momento comum da rotina de trabalho, quando o profissional pesquisa o repositório de informações do site jornalístico e, para explorar a produção existente, recorre ao mecanismo interno de busca do site. Esta tarefa rotineira pela busca de informações exige deste usuário (especializado) uma compreensão da interface de busca como ferramenta de trabalho. Neste sentido, esta ação envolve um raciocínio pessoal que incorpora familiaridade com o meio e princípios de investigação para checar, esclarecer, fundamentar ou produzir uma nova informação. Sem ainda entrar na questão de usabilidade desta ferramenta, cabe entendermos brevemente qual o contexto profissional no atual ambiente digital, indicadores de pesquisas que observam características inerentes ao profissional de comunicação e pontuar sua prática focada nesta ferramenta. Para combinar estas abordagens e criar uma imagem de ambiente, profissional e prática, recorreremos às noções expostas pelo experiente jornalista Ryszard Kapuscinski (associados aos cinco sentidos do jornalista: estar, ver, ouvir, compartilhar e pensar) e as características de recentes tecnologias de comunicação e pesquisa. A ideia é criar um ponto de reflexão entre comportamento de pesquisa e a prática profissional aplicada aos novos tempos; apresentar recentes pesquisas e indicadores de perfil e, por fim, análises de pesquisadores atentos à modulação profissional de mídias digitais.

VISÃO GERAL

Sobre o ofício de jornalista, baseados nos relatos de diversas palestras do renomado correspondente internacional Ryszard Kapuscinski (2003), traçamos uma abrangente reflexão sobre características do meio de comunicação escrita (no passado e no presente), perfil profissional do jornalista e apontamentos sobre mudanças na forma de atuar e elementos que prevalecem continuamente na conduta profissional em qualquer dimensão (impresso ou on-line).

Entre 2000 e 2002, várias foram perguntas que emanaram durante as palestras oferecidas por Kapuscinski (2003), entretanto, para o contexto deste trabalho, traduzimos, reelaboramos e destacamos nesta primeira parte duas questões feitas ao jornalista e que posicionam dúvidas do destino profissional frente ao momento tecnológico e seguimos com extratos de suas afirmações profissionais, que nos ajudam criar uma imagem do profissional, bem como associar e relacionar este aspecto ao atual panorama da comunicação digital.

Qual é o principal desafio que enfrenta um jornalista na nova era da informação? É alcançar a excelência na qualidade profissional e na produção de conteúdo ético. Mudaram os meios de averiguar e armazenar informações, de transmitir e comunicar, no entanto, o cerne da profissão continua o mesmo: informar, de maneira que seu trabalho possibilite ao leitor entender o mundo que o rodeia e, desta forma, obtenha informações, aprenda e adquira conhecimento.

Que tendência se vislumbra para o jornalismo atual, levando em consideração as diferenças entre um repórter de rua e o nascente repórter de internet, que pode realizar seu trabalho de casa? A internet oferece um jornalismo de informação imediata, serve muito para acelerar a transmissão de dados, para divulgação rápida pelo mundo. Entretanto, ao acumular uma enorme quantidade de dados não substitui o julgamento, a reflexão, o entendimento. Corremos o perigo de chegar a uma situação na qual os dados sejam abundantes, porém, nossa imaginação não saiba como processar e utilizar estes dados em nossa vida prática. Esta contradição sintetiza o drama de nossa cultura: acumulamos com muita rapidez dados e mais dados, no entanto, esta situação não nos ajuda entender nem melhorar o mundo. Os problemas reais de uma sociedade contemporânea, especialmente em via de desenvolvimento, não se resolvem com a internet.

Estas duas questões e respostas abrem espaço para duas inferências: 1) a mudança no processo de produção exige competências no uso das tecnologias para construir uma

informação e explorar seu formato de transmissão hipermídia, mas é mantido seu efeito informativo, assim como seus desdobramentos (ensinar, educar, lembrar) que agora estão fortalecidos por novas capacidades de interação, compartilhamento e conectividade. 2) o armazenamento de informação relevante e sua recuperação eficaz são também problemas de impacto no âmbito profissional de ambientes digitais.

Lidar com estas questões nos parece uma constante na vida profissional do jornalista, além disso, precisamos associar a elas um novo panorama de trabalho destacado nas seguintes reflexões de Kapuscinski (2003):

Caminhamos da autoria admirada ao anonimato colaborativo: no passado, jornalistas formavam um reduzido e admirado grupo de pessoas reconhecidas pela sociedade. Atualmente é uma pessoa anônima, que compõe um grande grupo de profissionais que, devido à mudança das rotinas de trabalho, atua para criação do produto final dos meios massivos de comunicação, cujo resultado final não apresenta uma autoria, mas um conjunto de pessoas, uma equipe, que participa da construção na notícia.

Associamos mundo virtual e fonte de história (memória da sociedade): o jornalista constrói, por meio da escrita, um mundo virtual que substitui o mundo real, assim como oferece condições de manipulação da opinião pública e registro da história. Saímos de um passado onde conhecíamos nossa história pela escola e relatos de nossas famílias, duas vertentes que formavam parte da memória da sociedade que pertencíamos. No presente, os meios de comunicação adquirem relevância e são agora a principal vertente convertidos em nova fonte da história (memória) e com crescimento muito mais veloz que o saber mais concreto e sólido dos livros.

Sempre um trabalho coletivo: a produção jornalística depende das informações e opiniões de pessoas. Portanto, o ofício do jornalista depende da ajuda, participação e pensamento de outras pessoas. Esta condição é fundamental para o exercício da profissão: ser capaz de atuar em conjunto com outros.

Aprender continuamente: os estudos são permanentes e a experiência não se acumula. Cada novo artigo, reportagem ou crônica se inicia partindo do novo, a partir do zero. O jornalismo diário se ocupa de novos dados, novos fatos e novos problemas.

Tempo de contextos e textos: o contexto editorial de um jornal permite que um texto possa ser bem entendido utilizando um editorial, um texto de análise retrospectiva, informações complementares e interpretações. O tempo altera o valor de um texto, ou seja, uma matéria escrita há 3, 4 ou 5 semanas não tem o mesmo valor de uma informação do dia anterior, por exemplo.

Tecnologias para comunicar: o desenvolvimento das tecnologias de comunicação, principalmente o uso de celulares e email, mudaram de forma radical a relação de jornalistas e seus chefes que, utilizando uma variedade de fontes de informação, direcionam, por exemplo, o trabalho dos correspondentes que são informados à distância por seus chefes sobre os acontecimentos e são consultados para confirmação da imagem que estão construindo na redação do jornal.

Questão (pressão) do tempo: jornalistas possuem pouco tempo para juntar informações e produzir textos e, como qualquer tarefa criativa, é necessário tempo para conversar com mais pessoas, ler mais documentos, observar mais e pensar mais. Resolver as coisas em tão pouco tempo conduz a produção de forma superficial. Provoca simplificações na rotina de trabalho, consequentemente aprisionadas em uma linguagem reduzida, pobre e limitada.

Para cada página escrita, cem foram lidas: a leitura prévia é indispensável, principalmente no campo social e político, para não apenas tocar descobrimentos já destacados por outras pessoas, assim como fortalecer a prosa desenvolvida.

Necessidade de juntar e guardar materiais que importam: na prática jornalística, tudo deve ser documentado: informações, testemunhos e ideias. Um problema básico do trabalho jornalístico é que tudo desaparece no dia seguinte e é esquecido. No entanto, se guardamos os documentos, no futuro podemos revisar este material conservado e decidir se poderá ser explorado em um trabalho mais pessoal.

Este contexto, proposto nestes destaques por Kapuscinski (2003), colabora para formarmos cinco sentidos (ou competências) essenciais ao jornalista (estar, ver, ouvir, compartilhar e pensar) e associarmos uma releitura de forma que:

Estar jornalista é uma condição em concorrência com outros produtores que, mesmo sem uma direção editorial clara promovida por um veículo tradicional de comunicação, foi expandida pelo crescimento exponencial de informação gerada por redes sociais, blogs e microblogs inseridos na cultura digital e criada pelos denominados “*prosumers*”, pessoas que atuam ao mesmo tempo como consumidores e produtores de informação e são capazes de interferir na forma de produção e customização de produtos.

Ver e ouvir agregam elementos de impacto em novos formatos de transmissão e visualização de informações digitais (vídeo, podcast e infográficos) incorporados e explorados na escrita de um hipertexto.

Compartilhar e pensar pressupõe busca e disseminação de informações, na dimensão local e global, por meio de pessoas e comunidades informacionais oriundas de ambientes como Facebook, Twitter, blogs, fóruns e outras plataformas de compartilhamento das chamadas mídias sociais. Espaços que são cada vez mais utilizados pelos jornalistas para captação de fontes e disseminação de informações. Por sua vez, estes novos espaços digitais provocam uma mudança na maneira de pensar a disseminação, identificação e organização da informação.

FATORES ESPECÍFICOS

Partindo dessa visão geral para fatores mais específicos de análise do perfil, abordamos algumas pesquisas realizadas com profissionais da área de comunicação, no Brasil e EUA. Referente ao mapeamento da mídia digital no Brasil (MIZUKAMI; REIA; VARON, 2014), podemos destacar alguns aspectos de percepção do jornalista brasileiro sobre o processo de digitalização⁵ de jornais:

- a digitalização tem causado profundo impacto na forma como jornalistas realizam seu trabalho (consideração constante e reforçada na literatura consultada neste estudo);

⁵ Esclarecemos que a abordagem sobre digitalização não está restrita ao simples processo de conversão de um documento físico em formato digital. No caso da produção jornalística, parte do processo envolve práticas profissionais para integração de todos os elementos digitais, sejam eles originados do conteúdo da versão impressa ou on-line.

- exige abordagem do jornalismo em tempo real;
- a procura por novos modelos de negócios gerou produção de conteúdo em diferentes interfaces, com texto associado a componentes audiovisuais e interativos;
- a convergência de plataformas tem piorado as condições de trabalho, uma vez que jornalistas são obrigados a assumir papéis e funções adicionais, como adaptar conteúdo impresso para o ambiente digital;
- exige cuidados éticos no uso de redes sociais, entretanto, novas tecnologias também favorecem transparência em relação à manipulação de informações.

A mencionada convergência das plataformas (redações de impresso e on-line) e a aceleração das mudanças na tecnologia de novas mídias implicam em novas competências. Esta proposição é evidenciada em recente pesquisa realizada com jornalistas americanos, de acordo com Willnat e Weaver (2014), 68,1% dos profissionais entrevistados assume que necessita de treinamento adicional para uso de recursos digitais dos quais destacamos:

- edição e produção de vídeos (30,5%);
- utilização de registros e documentos (26, 9%);
- promover engajamento por meio de mídias sociais (28,4%), na mesma pesquisa 40% dos jornalistas americanos dizem que utilizam mídias sociais com o propósito de ficar informado e monitorar empresas concorrentes;
- design e codificação para Web (19,4%);
- SEO (17,6%), técnicas de *search engine optimization* que visam potencializar a relação entre informação produzida e sites de busca;
- análise de notícias (15,3%);

Estes novos interesses de capacitação coincidem com outra pesquisa, sobre mudanças no mercado de trabalho do jornalista profissional em São Paulo (FIGARO, 2012), que estabelece uma dependência parcial entre sobrevivência profissional, nos ambientes digitais, e as condições de treinamento e atualização permanente. Além disso, mostra um conflito entre saberes e gerações do mundo analógico e digital, que implicam em empregabilidade, capacidade de uso de ferramentas para aumento da produtividade e domínio dos gêneros jornalísticos da produção on-line. Nestas condições, segundo Scolari⁶ (2008, citado por FIGARO, 2012, p.12), este profissional da era digital se caracteriza como “jornalista polivalente e de multiplataformas”.

Por observação, se avaliarmos as propostas para formação de novas competências que incluem o uso de diferentes ferramentas em ambientes digitais, os cursos e publicações oferecidos pelo Centro Knight para o Jornalismo nas Américas podem nos oferecer alguns indicadores da natural adaptação do profissional de comunicação ao ambiente digital que, conseqüentemente, exige administração de um volume maior de informação e mais agilidade para criar potencial de competitividade. Entre os exemplos observados estão cursos e publicações on-line nos seguintes campos: jornalismo de dados, visualização de informação, jornalismo com uso do celular e uso de ferramentas digitais para produção e edição de vídeos e áudio. Associados aos cursos desta organização estão publicações cujo conteúdo reflete também um novo contexto de atuação. Por exemplo, no livro “Ferramentas Digitais para Jornalistas”, da jornalista Sandra Crucianelli, alguns dos temas percorridos são: orientações de busca na Web, uso de marcadores sociais, documentos oficiais e acesso à base de dados⁷,

⁶ SCOLARI, C. **Hipermediaciones**. Elementos para una Teoría de la Comunicación Digital Interactiva. Barcelona: Gedisa, 2008.

⁷ Na opinião de Machado (2006, p.8), o jornalismo digital em base de dados (JDBD) representa uma modalidade jornalística que utiliza a base de dados para oferecer estrutura e organização ao processo de produção jornalística: apuração, composição, edição e circulação. O estudo dessa modalidade, expandido por Barbosa (2007), apresenta categorias de análise e funções operacionais das bases de dados, voltadas para recuperação da informação jornalística, que coincidem com a abordagem desse estudo. Entre as principais funções destacamos: indexar peças informativas; armazenar e preservar arquivos (memória) para facilitar o processo de recuperação de documentos; garantir relacionamento entre conteúdo; armazenar notas semânticas do conteúdo gerado; habilitar o uso de metadados para análise e visualização de informação; regular o sistema de categorização de fontes jornalísticas; transmitir e gerar informação para outros dispositivos digitais (celular, tablet etc.).

redes sociais e Web semântica (CRUCIANELLI, 2013). Estes temas estão integrados as novas práticas jornalísticas que, além dos dados estruturais de produção da notícia (assunto, contatos e abordagem), prevê também um roteiro para composição do material pesquisado que será produzido com uma arquitetura informativa hipertextual (encadeamento de páginas) e multimidiática (textos, fotos, vídeos, áudios, infográficos e animações) e, finalmente, disseminado em diversas plataformas (notebook, smartphones e tablets) (LACHOWSKI; GOMES, 2012).

Por último, segundo Calderín e Rojano (2006), com a reconfiguração dos perfis e áreas de trabalho, o profissional de comunicação atual deve incorporar em suas tarefas algumas técnicas do processo documental para aperfeiçoar atividades de recuperação da informação e planificar estratégias da arquitetura da informação. Isto porque a investigação, tarefa essencial para jornalistas, sofreu as maiores mudanças a partir da incorporação da tecnologia digital nas redações de jornais.

Para os autores, o profissional de informação encontra desafios relacionados à capacidade de formular perguntas neste ambiente e, para superar esta condição, precisa desenvolver e apropriar-se de ferramentas de busca para validação da informação, assim como obter condições que favoreçam a análise do grande volume de informação disponível.

Neste sentido, Codina⁸ (2000, citado por CALDERÍN; ROJANO, 2006) destaca que, assim como para documentalistas, jornalistas compartilham novas tarefas que podem ser classificadas em quatro aspectos:

- Busca e obtenção de informações na Web;
- Avaliação, seleção e descrição de recursos digitais;
- Procedimentos para melhorar a visibilidade na Web;
- Procedimentos característicos da produção e distribuição de informações.

⁸ Codina, L. La documentación en los medios de comunicación: situación actual y perspectiva de futuro. [cd-rom]. In: **Cuadernos de Documentación Multimedia**. Madrid: MULTIDOC. n.10, 2000. Número especial dedicado ao evento Primer Congreso Universitario de Ciencias de la Documentación.

Deste conjunto de tarefas, focamos nosso estudo na busca e recuperação de informação por meio do mecanismo de busca de sites jornalísticos. Entendemos que, para aplicar um método de buscar informação em repositórios de conteúdo jornalístico on-line, o profissional da informação (documentalistas e jornalistas) precisam conhecer características próprias dos documentos (informação) existentes neste suporte, assim como definir alguns princípios básicos da teoria de recuperação de informação que colaborem para entender dificuldades do usuário e permitam avaliar ferramentas de pesquisa e a capacidade informativa de suas interfaces. Abordamos este caminho de entendimento nos próximos tópicos deste estudo.

3. ASPECTOS DA INFORMAÇÃO JORNALÍSTICA ON-LINE

Para entender como a relação entre tecnologia e informação, gerada no ambiente digital, exige ações para resgatar a informação jornalística, frequentes mudanças na mídia digital registram a dinâmica de inovações, propostas para organizar e recuperar informações. Mencionando alguns exemplos, num período curto (de maio/2006 a agosto/2007), o site jornalístico Folha Online desenvolveu duas reformas gráficas no formato de apresentação das notícias, estabelecendo uma nova estrutura de informação que favorece o leitor. Além disso, iniciou a incorporação do arquivo de notícias retrospectivas (posteriores a 1998) e melhorou o tempo de resposta de seu mecanismo de busca, assim como a divulgação e classificação de seu conteúdo, baseada na frequência de acessos: artigos mais lidos, mais curiosos, mais enviados, temas em alta e baixa (FOLHA ONLINE, 2007).

Alterações semelhantes são observadas com mais destaque na imprensa internacional, como é o caso do jornal The New York Times. A versão on-line deste diário americano desenvolveu projetos para incorporar ao site todo conteúdo do jornal, desde seu início, em 1851; investiu na inovação dos recursos de divulgação e recuperação de informações; criou serviço de alerta de notícias (baseado nas palavras-chave

escolhidas pelo leitor) e desenvolveu um sistema para armazenar matérias selecionadas em um arquivo pessoal, localizado na conta do usuário.

Acompanhando estas transformações, o número de leitores de jornais on-line tem apresentado um crescimento significativo e progressivo. Em 2008, no primeiro trimestre do ano, pesquisa da Nielsen Online, solicitada pela Newspaper Association of America, constatou que mais de 66 milhões de pessoas (cerca de 40% de todos os usuários ativos de internet) visitaram Web jornais em busca de informações, número que representa um crescimento de mais de 12% em relação à audiência do ano anterior (NEWSPAPER ASSOCIATION OF AMERICA, 2008); em 2015, números da pesquisa indicam 176 milhões de visitantes no mês de março (NEWSPAPER ASSOCIATION OF AMERICA, 2015).

Em jornais on-line no Brasil⁹, o volume de informação oferecido para consulta também apresenta dimensões relevantes e de crescimento contínuo. Em 2007, o Folha Online, atual Folha.com, por exemplo, divulga que seu conteúdo possuía mais de 2,5 milhões de páginas já publicadas¹⁰. Além disso, anuncia que seu programa de busca permite que qualquer assunto seja consultado com tempo de resposta de décimos de segundo (FOLHA ONLINE, 2007). Outro jornal brasileiro na Web, o Globo Online, registrava que mais de 20% das visitas ao site tem como foco a recuperação de informação por meio do mecanismo de busca (CARVALHO, 2007).

Estes indicadores de audiência e consulta (baseada no consumo de informação), assim como de mudanças (técnicas ou estratégicas) na organização da informação jornalística e de sua plataforma, evidenciam aspectos de um ambiente em constante transformação e, também, em busca de um modelo de negócios que acompanhe evoluções tecnologias para garantir fidelidade e audiência do leitor.

⁹ É o caso dos jornais on-line escolhidos para investigação empírica: Folha.com (www.folha.uol.com.br); Estadão.com (www.estadao.com.br); O Globo Online (www.oglobo.globo.com), que foram destacados pela tradição de suas versões impressas.

¹⁰ Dados sobre a produção, especificamente on-line, indicam que a Folha.com publica um novo texto a cada 3 minutos, o que corresponde a 500 textos novos por dia (FOLHA DE S.PAULO, 2011). Este dado reforça a noção de rapidez de produção e dimensão do conteúdo gerado para Web.

Considerando este panorama, para analisar e refletir sobre questões de organização e recuperação de informação em repositórios de jornais on-line, três aspectos nos parecem importantes para estabelecer parâmetros de avaliação da informação jornalística: **Suporte / Estrutura Informativa / Memória**.

Para facilitar uma visão geral da lógica de raciocínio deste capítulo, antes do desenvolvimento das análises, pontuamos da seguinte forma as dimensões de reflexão:

O primeiro aspecto está relacionado ao tipo de suporte que apóia a estrutura informativa do documento, basicamente originada do formato impresso (mídia tradicional inserida no ambiente digital e narrativa textual com relação linear) e geração on-line (hipermídia com narrativa apoiada em hipertexto e associações interativas com fotos, vídeos, áudio, infográficos e artes);

O segundo aspecto trata da estrutura informativa e envolve a condição dinâmica do conteúdo, a existência de gêneros digitais e características estruturais do documento on-line, ou seja, assim como cada unidade de informação em um jornal impresso se configura (e permite sua localização) com elementos mínimos de identificação da estrutura de organização editorial (data, edição, editoria, página, título e autor), em meios digitais a organização destas unidades dentro de conteúdos maiores se configura da mesma forma, porém, potencializada por uma dinâmica de navegação que acrescenta outros identificadores;

O terceiro e último aspecto considera a noção de memória, sua aproximação com a recuperação e conectividade da informação em ambientes digitais e, de forma implícita, sua relação como atributo do conteúdo jornalístico.

3.1 SUPORTES EM CONVERGÊNCIA

A informação jornalística disponibilizada em sites e seus respectivos repositórios on-line é formada pela reunião de material derivado de dois suportes, impresso e on-line, com dinâmica de produção diferenciada, mas em constante processo de convergência e integração. No entanto, até chegar nesse ponto de integração, o conteúdo jornalístico acompanhou a evolução de três gerações de jornais digitais. Na primeira geração predominava a transposição de informações do jornal impresso para o ambiente digital; na segunda geração conteúdos originais são publicados, mas ainda estão dependentes das estruturas convencionais das organizações jornalísticas e carentes de sistemas digitais para atender características do ciberespaço e ao mesmo tempo facilitar produção e circulação do conteúdo; na terceira geração, a característica determinante é a produção de “conteúdos originais em formato multimídia, constituindo sistemas descentralizados próprios, capazes de incorporar contribuições dos usuários, para apuração, produção e circulação de conteúdos” (MACHADO; BORGES; MIRANDA, 2003). A esta terceira geração, acrescentamos também inovações na capacidade de disseminação de informação favorecidas pelo meio digital.

Para Lara e Conti (2003), “disseminar informação supõe tornar pública a produção de conhecimentos gerados ou organizados por uma instituição”. Neste processo, a disseminação leva em consideração aspectos de uso e linguagem e está associada à noção de transferência, base do processo jornalístico apoiada no tradicional esquema de comunicação: emissor-canal-mensagem-receptor.

No ambiente digital promovido pela internet, o tradicional esquema foi submetido por novos paradigmas e efeitos da tecnologia que, conseqüentemente, provocaram desdobramentos no processo de disseminação: maior volume de dados, velocidade de acesso, autonomia do usuário e dinâmica da relação produção/disseminação/consumo sobre a circulação de informações. Neste sentido, Souto (2010) aponta as principais mudanças ocorridas nos serviços de disseminação seletiva que também configuram a nova fase dos jornais digitais: integração de recursos informacionais em diferentes formatos (vídeos, áudios, fotos); uso de novos canais de disseminação (sites, blogs,

fóruns); interação síncrona e assíncrona com usuários dispersos geograficamente; democratização da ação de disseminar, permitindo que usuários de informação promovam voluntariamente a disseminação; uso de diversas tecnologias para entrega de pacotes informacionais (email, RSS, alertas automáticos, redes sociais); uso de diferentes ferramentas de retroalimentação (enquetes, formulários e questionários on-line).

Esta evolução natural, da forma de produção, linguagem, formato e disseminação da informação jornalística que circula na internet, evidencia um distanciamento de características presentes nos tradicionais meios de comunicação e modificações cada vez mais impulsionadas pela dinâmica do meio digital.

Cabe aqui esclarecer que são diversas, contínuas e complexas as abordagens da área de Comunicação sobre o processo de convergência (impresso e on-line) e sua dinâmica, no entanto, nos interessa apenas situar historicamente que tratamos de uma geração de jornais on-line que transitou da simples reprodução de conteúdo impresso ao atual momento de produção exclusivamente on-line, que explora toda a capacidade de hipermídia e hipertexto, a exemplo do uso de gráficos interativos, criação de páginas especiais e elaboração de dossiês digitais.

3.2 ESTRUTURA INFORMATIVA [CONTEÚDO, GÊNERO E DOCUMENTO ON-LINE]

Para Alzamora (2004), que analisa a informação jornalística pela perspectiva da semiótica linguística (conceito de ação ou influência, proposto por Charles Pierce, que relaciona três elementos: signo, objeto e interpretante no processo de produção de significados), a origem da informação jornalística presente no meio digital é semelhante ao do produto impresso. Porém, ao transferir sua estrutura informativa para este novo meio (ambiente) surgem novas possibilidades de construções e a estrutura é modelada pelo meio. Neste sentido, ocorre uma ampliação na funcionalidade do suporte que, além da característica de transmissão de informação, adquire o atributo de associação de informações.

Se entendermos que o processo de transmissão da informação jornalística implica na relação dinâmica entre signo (texto jornalístico), objeto (fato jornalístico) e intérprete (usuário da informação jornalística que reelabora e absorve o texto a partir de suas experiências), segundo Alzamora (2004), podemos considerar que a informação surge de relações de determinação e representação, nas quais o texto (signo) é mediador entre determinação (transmissão de conteúdos) e representação (associação de conteúdos) com propósito de promover a comunicação.

Este processo de transmissão da informação jornalística permite uma relação dinâmica entre signo (texto jornalístico) e objeto (fato jornalístico), no entanto, é no processo de representação que a noção de informação jornalística adquire sentido de novidade – característica essencial para noção jornalística de notícia. O fato jornalístico (objeto) passa a ser interpretado através da associação com outros textos jornalísticos (signos), apoiado no nexos estabelecido pelo usuário (intérprete).

Esta dinâmica da transmissão e associação se fortalece ainda mais pelas medidas de valor da informação que constituem o relacionamento entre conteúdo jornalístico on-line e seu usuário. Considerando empresas informativas em ambientes digitais, para Tomsen¹¹ (2000, citado por Saad, 2003), “a questão do valor da informação tem relação direta com a potencialização do usuário” e apresenta cinco fatores de valorização da informação digital:

- *credibilidade* da fonte geradora do conteúdo
- *inovação* derivada da exclusividade do conteúdo
- *relevância* capacidade de impacto adquirido pelo conteúdo
- *ineditismo* fortalecido pelo fator tempo, em relação à velocidade de geração do conteúdo
- *utilidade* do conteúdo no cotidiano do usuário

¹¹ TOMSEN, Mai-Lan. **Killer Content**: Strategies for Web Content and E-Commerce. Massachusetts: Addison-Wesley Information Technology Series, 2000.

Alzamora (2004) acrescenta que pela abordagem da notícia o valor da informação jornalística se atrela principalmente ao fator novidade, teoricamente apoiado no modelo de comunicação Shannon-Weaver: *emissor-canal-receptor* para transmissão de informação. Na perspectiva da internet, considerando usos e recursos de linguagem do meio digital, o processo comunicativo vai além da simples transmissão de informações e ocorre na dimensão da hipermídia, isto é, uma linguagem de comunicação que explora e relaciona métodos de transmissão de informação por texto, áudio, vídeo e ainda recebem interferência das comunidades virtuais. Conteúdo composto não apenas por palavras, mas também por “associações visuais, sonoras e gráficas que podem ser feitas em torno dessas palavras que conferem valor à informação.” (SAAD, 2003). Esta estrutura e sua dinâmica permitem a formação de um

[...] hábito associativo que se processa na mente do intérprete, característica do símbolo, deriva de informações prévias do interprete, ou seja, quanto mais o jornalista estiver familiarizado com o assunto de sua reportagem, mais diferenciada e densa tende a ser a notícia por ele produzida. E quanto mais rico for o processo representativo que emerge da notícia, mais complexas e diversificadas tendem a ser as associações [...] (ALZAMORA, 2004).

Esta visão destaca uma nova capacidade da estrutura informativa que, além de adicionar certo valor ao conteúdo, permite múltiplas possibilidades de associações ao usuário leitor, assim como ao usuário produtor (jornalista), provocando uma expansão do processo comunicativo, da relação entre produtor de conteúdo e leitor, do uso e compartilhamento de informações jornalísticas em rede.

GÊNERO DO CONTEÚDO

Na estrutura de um site jornalístico, existe um evidente agrupamento de conteúdo pela respectiva área editorial de um hipertexto (economia, cultura, cotidiano, ciência etc.) e esta organização é cada vez mais auxiliar na seleção de conteúdo. Além disso, complementar a esta categorização, existem gêneros digitais (categorias não

explícitas) de conteúdo jornalístico que oferecem uma tipologia para recuperação de informação. Assim como os tradicionais elementos de identificação, o gênero jornalístico em ambientes digitais permite novas possibilidades para orientação de leitores, jornalistas e pesquisadores. Ressaltamos que esta reflexão não pretende analisar as categorias existentes e suas características, mas apontar este atributo como uma importante funcionalidade linguística para organização e recuperação do conteúdo de repositórios de informação jornalística on-line.

Os gêneros empregam ao conteúdo categorias mais específicas de agrupamentos dos discursos produzidos em ambientes on-line. Segundo Seixas (2003), gêneros se configuram como grupos de textos e estão presentes na “ordem do discurso” e “do processo de interação verbal”. Sua forma e conteúdo são características de classificação e permitem a relação do texto produzido com a realidade (pelo caráter opinativo e informativo), assim como a intencionalidade do produtor do texto, atribuindo funcionalidade ao discurso (opinião, informação, interpretação e entretenimento), preocupação predominante na teoria jornalística em relação aos formatos e seus autores (jornalistas).

Ainda para Seixas (2009), faz parte do aprendizado do jornalismo a produção de gêneros jornalísticos para ação comunicativa. Produzir este gênero se configura como um hábito e um saber profissional que, pela visão de Sánchez e López Pan¹² (1998, citado por SEIXAS, 2009), podem ser observados da seguinte forma:

- (1) o jornalista escreve inevitavelmente em gêneros,
- (2) os gêneros funcionam para o jornalista como balizas/modelos de expressão que facilitam sua tarefa,
- (3) os gêneros cumprem determinadas funções sociais,

¹² SÁNCHEZ, J. F. e LÓPEZ PAN, F. Tipologías de géneros periodísticos em España. Hacia un nuevo paradigma. Comunicación y Estudios Universitarios, **Revista de Ciències de La Informació**, Valencia, CEU San Pablo, n.8, p.15-35, 1998. Disponível em: <<http://dspace.unav.es/dspace/handle/10171/34984>> Acesso em: 10 fev. 2014.

(4) os gêneros são instituições vivas que evoluem para ajustar-se às funções próprias da atividade a que servem. Eles não só evoluem, também desaparecem e surgem outros novos,

(5) a classificação dos gêneros importa, e muito, na medida em que reflete os valores da profissão e seus pressupostos epistemológicos.

Este entendimento da análise de domínio profissional, em certa medida, favorece a compreensão do impacto dos gêneros para agrupamento de informação e, posteriormente, a aplicação de filtros para seleção e recuperação de conteúdo.

No ambiente digital, o gênero é potencializado pelo meio, que permite construir uma narrativa em hipermídia. Além disso, pode ser analisado pelos parâmetros de

[...] tempo (instantaneidade/atualização contínua), número e tipo de interlocutores (interatividade), formato textual e extensão (hipertextualidade), limites impostos a revisão (atualização contínua), grau de automatização das operações, método de armazenamento, busca, gerenciamento de textos (memória) e riqueza e variedade de sinais, ou seja, texto, áudio, imagem (multimedialidade) (SEIXAS, 2008).

Para entender peculiaridades do conteúdo do jornalismo on-line, em uma breve condensação de gêneros, a partir de estudos reunidos por Díaz Noci e Salaverría Aliaga (2003), podemos pontuar as seguintes categorias:

Gêneros Informativos

- *Notícia hipertextual*: no meio digital, sem a limitação de espaço imposta pelo papel, a característica da notícia, de texto curto, sem interpretação e que oferece informação sucinta sobre um fato, se converte em um texto de maior profundidade e enriquecida com links e informações da atualidade, com maior velocidade de produção, correções e atualizações.

Gêneros Interpretativos

- *Reportagem*: relato extenso e minucioso sobre qualquer assunto da atualidade; produzida em ambiente digital, esta narrativa absorve o caráter de multimídia e rompe com elementos como: sequencialidade da leitura (permite amplitude temática, estrutura aberta e formação de dossiês documentais), periodicidade (quando inserida em uma página da Web pode ser continuamente acessada pelo leitor em qualquer momento) e associação de materiais interativos (chats ou fóruns de debate).
- *Crônica*: texto com componente informativo cujo discurso do autor está apoiado na experiência que possui sobre a realidade; mesmo potencializado pelo recurso do ambiente digital está apoiada na credibilidade do autor, mas encontra maiores funcionalidades interativos com o leitor ao migrar para estrutura de blog.

Gêneros Dialógicos

- *Entrevistas*: baseada no diálogo entre pessoas, esta narrativa é potencializada pela capacidade de interatividade do meio digital, seja pelo emprego do diálogo em rede por texto, som ou vídeo, ou pelos canais de comunicação que permitem a construção da entrevista: conversas por email, chat, troca de mensagens por celular (SMS), conversação por Twitter; bem como pela multiplicidade de conexões entre pessoas e amplificação do ciclo comunicativo entre emissor e receptor.

Outros gêneros (editorial, comentários, crítica, cartas do leitor, artigos, colunas, blogs etc.) poderiam ser expostos nesta abordagem, no entanto, o interesse dessa análise é expor a complexidade de organização do conteúdo jornalístico diante de um universo de informações em expansão, de classificação rudimentar e atributos de recuperação elementares. Neste sentido, os gêneros jornalísticos são agrupamentos evidentes e naturais na organização de conteúdo, no entanto, são pouco explorados em interfaces de pesquisa e, conseqüentemente, desconsiderados como atributos de seleção e relevância na recuperação de informação.

Para Bertocchi (2013), que defende a noção de narrativa jornalística digital como “um sistema adaptativo complexo”, o gênero estaria inserido na categoria de formato, componente de um “sistema narrativo” que é composto por dados e metadados. De forma esquemática e simplificada, componente da seguinte formatação narrativa:

[NARRATIVA = DADOS + METADADOS + FORMATOS]

Este conjunto de elementos, integrantes da narrativa digital, estabeleceria uma relação de gênero textual com o termo formato visto que, segundo Ramos¹³ (2012, citado por BERTOCCHI, 2013), no contexto do jornalismo digital a noção de formato utilizada por profissionais de comunicação é empregada “para designar a forma como a notícia será desenvolvida e organizada, ou seja, a sua forma pré-definida.” Este formato, dentro do sistema, permite conexões entre o usuário da informação e a interface narrativa. Por este aspecto, reforçamos ainda mais a análise que concede ao gênero uma significativa importância como mais um atributo de pesquisa capaz de potencializar a recuperação de informação jornalística.

DOCUMENTO ON-LINE

Após analisar características da estrutura informativa do conteúdo jornalístico on-line, pela perspectiva de valor implícito, precisamos também refletir sobre a estrutura informativa de representação do documento digital. Para García Gutiérrez (1990), no âmbito da linguística documentária, a estrutura documental se constrói em função da recuperação de informação e, por isso, precisamos observar cada documento como uma unidade textual composta de enunciados cuja importância impacta na organização dos dados e na "microestrutura" criada pelos produtos, neste caso, informação jornalística on-line. A estrutura dessa documentação, como organização de conteúdos codificáveis e decodificáveis, possui dois sentidos de análise:

¹³ RAMOS, Daniela O. **Formato**: condição para a escrita do jornalismo digital de bases de dados. Uma contribuição da Semiótica da Cultura. Tese de Doutorado. Escola de Comunicações e Artes da Universidade de São Paulo, 2012.

- Estrutura da produção da informação: sua formação, organização e apresentação das ideias do produtor de conteúdo, como corpus para observação e descrição.
- Estrutura da representação do discurso do produtor, vinculadas as formulações metodológicas e modelos de sínteses e traduções do conteúdo produzido.

Desta forma, para situar a estrutura da documentação jornalística on-line no processo comunicativo, precisamos identificar o ponto de interação entre emissor e receptor. Além disso, compreender a divisão dessa estrutura entre categorias cognitivas do produtor (jornalista) e interpretativa do consumidor (leitor), bem como traduzir a não-linearidade de um sistema hipertextual que rege a manipulação de documentos (textos).

Na estrutura do sistema hipertextual, dois componentes básicos adicionam a dinâmica não-linear as condições de organização e associação de conteúdo: **nó** é **vínculo** que, de acordo com Lima (2006), podem ser definidos da seguinte forma:

Nó: Cada uma das unidades de informação em uma base de hipertexto. Cada nó corresponde a uma ou mais exibições de tela. Essas unidades de informação em um hiperdocumento podem conter diferentes tipos de dados, textos, figuras, fotos, sons e são conectadas por vínculos em uma variedade de estruturas. Geralmente, o nó descreve um único conceito ou tópico de modo que pode ser caracterizado como autocontido, não dependendo da leitura prévia de outros nós (LIMA, 2006).

Vale exemplificar que a construção de uma reportagem especial em ambiente digital é construída pela relação de um conjunto de nós (documentos) de diferentes tipos: texto principal e auxiliar, infográficos interativos, galerias de imagens, enquetes, comentários de leitores etc., todos com conteúdo independente, mas estruturalmente relacionados entre si e, por isso, exigem a manutenção da lógica de composição editorial, uma estrutura que precisa ser visível e recuperável.

Vínculo: O conceito mais importante do hipertexto é o vínculo também denominado de *link*, elo, ligação, âncora ou botão, de acordo com o sistema de hipertexto. Vínculos são marcas que conectam um nó ao outro. A ativação desses vínculos implica abertura de uma nova janela contendo o documento

referenciado. Assim é fácil deslocar-se fácil e rapidamente, de um ponto para outro, no conteúdo das páginas. Esses vínculos podem ser representados por palavras ou frases em destaque (negrito, itálico, cores), ou ainda, por figuras (LIMA, 2006).

Pela perspectiva de construção da informação jornalística, no caso dos links, o processo de produção do documento permite vínculos associativos para informações complementares da notícia (imagem, som e ilustração), assim como vínculos associativos de memória (textos já publicados e selecionados para apoio na contextualização da notícia). Características que, de certa forma, estão relacionadas à outra divisão destacada por Lima (2006) ao descrever vínculos referencias (com características organizacionais e auxiliares na navegação e focados na facilidade de leitura) e vínculos semânticos (para associação de conteúdos, por meio do relacionamento de informações adicionais ou similares).

Por encontrar limites nestas operações de análise do conjunto de documentos hipertextuais e na coerência de relacionamento entre nós e vínculos, devido à ausência de elementos estruturados de recuperação de informação em repositórios de jornais on-line, uma das possibilidades de visualização da estrutura documental também pode ser observada pela presença de metadados do documento.

Segundo análise de Alves (2010), metadados são atributos que representam um objeto (entidade) do mundo real e estão presentes no sistema de informação. Dito de outra forma:

[...] são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação (ALVES, 2010, p.47).

Embora seu uso e padronização sejam essenciais, os metadados, entendidos aqui como um conjunto de regras para identificação (etiquetagem) de dados para recuperação de informação e gestão de conteúdo, ainda precisam de ajustes e acordos de aplicação em sistemas de recuperação, sejam eles baseados em texto jornalístico

completo, palavras-chave ou desenvolvidos para construção de Web semântica ou ontologias.

Para Méndez Rodriguez (2002), as características da Web (volume, variedade e volatilidade de informação), assim como a desestruturação técnica das informações de documentos eletrônicos neste universo, intensificou a geração de ferramentas e linguagens para minimizar os problemas de recuperação de informação: qualidade e confiabilidade dos resultados pesquisa e consistência dos buscadores de rede. Entre as ferramentas para evitar a inconsistência na busca de informações relevantes, os metadados aparecem como aplicativos ideais para recuperação de informação na internet. Contudo, os problemas de linguagem computacional não desapareceram.

No caso da informação jornalística¹⁴, esforços de organização e estruturação da informação são percebidos em iniciativas, algumas bem recentes, que acompanham a evolução tecnológica e são desenvolvidas na área de comunicação. Por exemplo, desde 2005 está disponível uma taxonomia para metadados de notícias, denominada NewsCodes, criada e mantida pelo IPTC-International Press Telecommunications Council (IPTC, 2015a), apoiada no modelo SKOS (Simple Knowledge Organization System) para padronização e compartilhamento de dados na Web.

O IPTC também desenvolveu o rNews¹⁵, padrão aprovado para marcação semântica da produção de notícias on-line que, por sua vez, serviu de base para Schema.org¹⁶

¹⁴ Percebemos que a busca por uma melhor solução para lidar com o problema dos dados estruturados está baseada na moderna construção de Web sites que, basicamente, funcionam com uma arquitetura em três camadas: camada de dados, formada pela base de dados do repositório; camada de apresentação dos dados, definida pelo código HTML do documento acessado; camada lógica, definida pelo software que (relaciona) faz a leitura dos elementos da camada de dados e os exibe na camada de apresentação (IPTC, 2015b).

¹⁵ O modelo de dados rNews opera com a associação de duas classes de metadados: item de notícia e conceito. Por exemplo, um item de notícia é formado pelo texto e sua possível relação com outras mídias (áudio, imagem e vídeo). Esta composição compartilha uma série de atributos (como data de criação, título e copyright) que constituem o item de notícia (documento digital) e se relacionam com um conceito que, formado por subclasses do mundo real (personalidade, organização e lugar), qualificam o documento digital (IPTC, 2015b).

expandir, em 2011, as propriedades de marcação do conteúdo noticioso. (SCHEMA, 2011).

Também preocupado com a recuperação de informação jornalística, em 2012, desenvolvedores do Google (2012) propuseram a aplicação da metatag “News Keyword” para ampliar a capacidade de reconhecimento de metadados de notícias e, consequentemente, a visibilidade de conteúdo jornalístico em seus serviços de busca. Outra proposta que também é visível na estrutura de metadados é o protocolo OG (Open Graph)¹⁷ com aplicação em rede social.

Com algumas variações na estruturação de dados, de forma geral, jornais nacionais e estrangeiros adotam recomendações para descrição de metadados (ver Quadro 1). Estes procedimentos facilitam aos mecanismos de busca gerar melhores resultados de busca, ampliam a visibilidade do documento na Web .

Na perspectiva de estratégias futuras, esforços para potencializar estes metadados visam auxiliar que características da informação jornalística possam ser bem exploradas, assim como aproximar usuários de informações. Por esta perspectiva, Gouvêa e Loh (2012) consideram que a noção de “jornalismo semântico”, associação da estrutura do lide¹⁸ aos formatos estruturados (RDFa e microdados) de metadados, ajudaria organizar conteúdos e “aumentar a usabilidade e eficiência no acesso às informações”.

¹⁶ Formado por um grupo que desenvolve recomendações de apoio para incorporação de dados estruturados em páginas da Web e uso de mecanismos de buscas, a Schema.org definiu extensões para marcação de textos jornalísticos que são adotadas pela maioria dos jornais on-line analisados neste estudo (SCHEMA, 2011).

¹⁷ Lançado pelo Facebook em 2010, este protocolo permite que qualquer página da Web tenha funcionalidade semelhante a qualquer objeto de informação do Facebook (OPEN GRAPH PROTOCOL, 2014).

¹⁸ O lide é o primeiro parágrafo do texto e contextualiza um fato de forma clara e objetiva. Contém a informação mais importante para o leitor, a idéia mais significativa do texto, uma revelação, um aspecto da notícia mais curioso ou polêmico. É definido por elementos de estruturação [quando + quem + onde + o que + como + por que] que transmitem informações essenciais logo no início do texto (PINTO, 2009).

Quadro 1 – Ocorrência de metadados em jornais on-line

JORNAL ONLINE	METADADOS (ocorrências identificadas em códigos fonte)				
	RDFa	Microdados	Adota Recomendações		
	aplicação (+) complexa	aplicação (-) complexa	Google	protocolo Open Graph	Schema.org
New York Times	x	x	x	x	x
Washington Post	x	x	X	x	x
Wall Street Journal	x	0	x	x	0
El País	x	x	x	x	x
Le Monde	x	x	0	x	x
Guardian	x	x	x	x	x
Estadão	x	0	0	x	0
Folha	0	x	0	0	x
O Globo	x	x	0	x	x
identifica uso de > marca no código fonte>	propriedade meta property	propriedade itemprop	palavra-chave "news_keywords"	rede social meta property="og:title"	vocabulário Schema.org/NewsArticle

Fonte: Elaboração do autor, 2015.

Nota: O uso de metadados pode ser observado a partir de ocorrências no código fonte de textos jornalísticos recentes (ano de 2015). Durante busca aleatória nos jornais on-line analisados nesta pesquisa, consultamos o código de alguns textos e, apenas para ilustração, identificamos no quadro elementos de metadados e recomendações que definem atributos de recuperação do documento digital e indicam a complexidade técnica e estrutura documental na Web. O uso desses atributos implica em maior visibilidade do documento e favorece a eficiência dos mecanismos de busca para recuperar informações.

Neste contexto, que evidencia uma constante busca por linguagens, padrões e recomendações que melhorem a estruturação dos dados, organização do conteúdo, capacidade semântica e resultados de busca, alguns elementos são exaustivamente estudados. Sem entrar em detalhes explicativos sobre cada um deles, mencionamos apenas que são elementos base para o ambiente Web que recorre, ao mesmo tempo, as propostas estruturantes complexas, de visibilidade dos dados, que utilizam sintaxe comum (XML, HTML5, XHTML), bem como as estruturas simples predefinidas (RDF, RDFa, Microformatos e Microdados)¹⁹ que enriquecem de atributos descritivos um recurso informacional.

¹⁹ RDF (Resource Description Framework) é um modelo de descrição de recursos que usa declarações simples para descrever relações recurso-propriedade-valor / sujeito-predicado-objeto, dito de outra forma, uma tripla RDF é formada por elementos básicos que se relacionam, por exemplo, sujeito (Machado de Assis), predicado (é autor de) e objeto (Dom Casmurro). Considerado de alta complexidade para aplicação, o RDFa (in attributes) provê um conjunto atributos que expressam dados RDF; com média e baixa complexidade de aplicação, soluções alternativas são propostas para ampliar a interoperabilidade de sistemas: os Microdados permitem adicionar marcações estruturadas em qualquer elemento HTML; Microformatos são extensões, um conjunto de padrões simples no formato de dados abertos. Este grupo de formatos é aplicado para dar mais significado as linguagens computacionais de descrição (XML) e exibição (HTML) dos dados em ambientes digitais (FERREIRA; SANTOS, 2013; SPORNY, 2011; MICROFORMAT.ORG, 2014, 2015; W3C, 2014, 2015).

Diante do crescimento contínuo de informação na Web, estes atributos são cada vez mais importantes na tarefa de ampliar a eficiência dos mecanismos de busca no processo de recuperação, assim como acrescentar estrutura de organização semântica que implique em relevância ao conteúdo representado.

Entretanto, são várias as deficiências de recuperação na Web, tais como: sobrecarga da informação, palavras-chave ineficientes, ausência da identificação de autoria e falta de sistemas com adequado processamento de linguagem natural. Além disso, existe a necessidade de lidar com características da Web Semântica que, diferentemente dos mecanismos de usabilidade da web social, possui uma linguagem de controle muito formal e pouco amigável para aplicações em softwares, técnicas automáticas pouco consolidadas para criação de ferramentas (por exemplo, ontologias voltadas para conceituação de um determinado domínio) e duplicidade de propostas com vocabulários de metadados (MOREIRO GONZÁLES, 2011).

Por este quadro, finalizamos destacando dois pontos de observação, sempre presentes na literatura, que parecem essenciais na reflexão sobre a estrutura informativa de um documento on-line:

Tecnológico (representado por metadados) que estabelece recomendações e padrões de linguagem computacional que permita interação entre descrição do documento, funções de recuperação e representação de conteúdo adequado exibido nos resultados.

Humano (representado pelo processo de etiquetar conteúdos, *tagging system*²⁰ para adicionar metadados, realizado pelos produtores da informação) que utiliza a linguagem de um determinado domínio para gerar coerente organização,

²⁰ O sentido de *tagging* que empregamos é de processo para indexação de assuntos. As *tags* são essencialmente metadados de um determinado recurso informacional e, no contexto da informação jornalística, o processo de *tagging* é realizado por um profissional de informação que atua como curador do conteúdo gerado. O modelo básico de um sistema de marcação (*tagging system*) envolve (além da relação entre usuário, recurso informacional e *tag*) a intersecção de três campos de atividade: arquitetura da informação (esquema de organização para a informação ser encontrada), software social (aplicação mediada por computador para compartilhamento de informação on-line entre um grupo de pessoas) e gestão de informação pessoal (prática de organização da informação para posterior recuperação de recursos publicados) (SMITH, 2008; MARCHIORI, 2012; SAAD; BERTOCCHI, 2012)

relacionamento e recuperação de conteúdo e, conseqüentemente, compartilhamento de conteúdo.

Promover uma visão integradora destes dois pontos nos parece a melhor maneira de buscar soluções para representação, organização e recuperação da informação.

3.3 MEMÓRIA (DIGITAL), RECUPERAÇÃO DE INFORMAÇÃO E CONECTIVIDADE

Por ser um conceito transversal e permitir diversas abordagens, analogias e associações importantes, a noção de memória é absorvida por diversas áreas do conhecimento. Pela perspectiva biológica, alimenta discussões sobre a memória humana e seus elementos cognitivos; já por uma perspectiva de caráter físico, que atribui espaço, volume e flexibilidade à sua constituição, pode se apresentar como o dispositivo (de memória) instalado em um computador ou, ainda na mesma área, derivações que a computação oferece como, por exemplo, memória persistente (permanente), memória volátil (desaparece ou apaga), banco de dados (coleções de dados ou memórias) ou, de utilização mais recente, uma memória virtual armazenada e de acesso remoto, nas “nuvens”, como é entendido um sistema de *cloud computing*; quando observada pela perspectiva social, cultural ou filosófica adquire então outras dimensões, mas particularmente no campo da história, oferece uma corrente de pensamento que atribui à memória a condição de patrimônio e, por meio de seu resgate e preservação, permite relações entre passado, presente e futuro de uma sociedade ou pessoa.

Muito embora essa diversidade de aplicações só evidencie a complexidade de abordagens que podem surgir da noção de memória (cognitiva, física ou social), cada uma delas carrega um ponto em comum: informação. Memória é formada por informações e pela capacidade de guardar, preservar, conectar e recuperar estas informações. Partindo dessa proposição, com a intenção de ampliar discussões na área da Ciência da Informação, estabelecemos neste tópico algumas relações entre o atributo da memória, repositórios digitais de conteúdo jornalístico (como exemplo de

fonte de conhecimento na Web), recuperação e conectividade de informação para ativar ou desativar a memória (individual ou social).

RECUPERAÇÃO DE INFORMAÇÃO

Entre as novas teorias da comunicação, o modelo teórico da virtualização, desenvolvido por Pierre Lévy, um dos principais pensadores do ciberespaço, promove uma reflexão sobre a questão da “memória” baseada nas mudanças ocorridas na forma de pensar e se comunicar do homem que, provocadas pela informática, permitiram o aparecimento de “tecnologias da inteligência” (POLISTCHUK, 2003).

Nessa direção, observamos que a característica de registro e preservação da memória humana está também intrínseca ao uso das tecnologias digitais. Portanto, a geração de repositórios de texto existentes na Web se configura em mais uma fonte da memória da sociedade que, situada numa diferente fase na construção dos seus registros de informação, historicamente evoluiu de um processo oral de transferência de conhecimentos e informações, para um registro impresso (físico) deste conhecimento e que, agora, apoiado na linguagem escrita desenvolvida em séculos, adquiriu um formato digital de registro (LÉVY, 2002).

Ao identificar uma evolução e o aparecimento desse novo formato de registro, podemos explorar a metáfora da memória humana (biológica) para representação do que seria a formação de uma memória digital, suas características e limitações. Para Lévy (2002), nossa memória humana em nada se parece com um dispositivo para armazenamento e recuperação fiel das informações. Além disso, para uma nova informação ser memorizada (gravada), precisamos construir uma representação deste saber. Em outras palavras, a fidelidade de recuperação depende da criação de um “documento mental” que represente esta informação. Desta forma, ao informatizarmos o saber, afastamos a necessidade humana do saber “de cor” e passa ser fundamental a operacionalidade (do processo de lembrar) e a velocidade da memória (na recuperação de informação).

Estes pontos (operacionalidade para lembrar e velocidade para recuperar) podem ser introduzidos, por exemplo, na avaliação de um sistema de recuperação de informação na Web. Neste contexto, a noção de memória se configuraria como um novo atributo ou medida para avaliação de sistemas de informação jornalística, regido pela dinâmica da produção de narrativas em “tempo real”, que ultrapassa a velocidade de geração de registros diários para registros por hora ou minutos, permitindo uma recuperação com atualização mais rápida e maior volume de informações para operacionalizar o lembrar de um determinado momento histórico.

No caso da mídia digital, muito embora para empresas jornalísticas a Web tenha sido, logo em seu início, mais um canal de divulgação da informação impressa, com a evolução do jornalismo na Web e a definição de estratégias de convergência dos ambientes (digital e impresso), a criação de redações com produção de material exclusivamente on-line, diferenciado ou complementar, determinou a geração de importantes bancos de dados e repositórios digitais. Este novo aspecto, no sentido que a memória desta produção também precisa ser recuperada (ou ativada) de maneira adequada, contribui para repensarmos as tecnologias de pesquisa desenvolvidas nestes ambientes.

Ao repensar as tecnologias de pesquisa, como um mecanismo de busca, precisamos estabelecer dispositivos que permitam uma memorização eficiente, evitando o esquecimento, um estado relacionado à invisibilidade que uma informação pode adquirir se não for devidamente estruturada e representada. Para isso, a aplicação da linguagem documentária²¹ em ambientes digitais, como mecanismo de representação entre produtor e consumidor de informação de uma determinada área de conhecimento, oferece um caminho contra este esquecimento ao se preocupar em

²¹ Segundo Lara (2004), seja qual for o universo aplicado, sites Web ou bases bibliográficas, a linguagem documentária configura-se como uma linguagem construída para um sistema de organização e comunicação de informação. Para atender essa função, precisa reunir qualidades que permitam funcionar como um código inteligível e fonte para interpretação de sentidos, servir como metalinguagem e incorporar o usuário em seu processo de elaboração e uso.

recuperar e ativar características do documento com um objetivo específico (arquivar, catalogar, pesquisar, traduzir, memorizar, conectar etc.).

Por esta linha de raciocínio, ao abordar o aspecto de um banco de dados como memória, Machado (2006) propõe que talvez tenhamos que fazer uma distinção entre memória como uma simples faculdade de conservação do passado e memória como elemento de ativação do passado por meio de demandas do presente. No âmbito das redes digitais, esta memória, “antes de refletir um passado morto, apresenta parâmetros para aumentar o coeficiente de previsão no fluxo ininterrupto de circulação de notícias” Machado (2006). A qualidade da memória do banco de dados estaria relacionada à sua capacidade de ativar a memória humana durante o processo de recuperação de informação. Desta forma, uma base de dados (ou sistema de recuperação de informação), entendida como forma comum das sociedades em redes, pode assumir a função de memória dos conteúdos jornalísticos na Web.

Este aspecto, de memória como medida de qualidade do ciberjornalismo, é também tratado por Palacios (2008) ao observar a necessidade de incluirmos este atributo na lista de critérios utilizados para avaliar produtos jornalísticos na Web (entre eles: facilidade de acesso, linha editorial, qualidade de notícias e interpretação). Muito embora esta noção exija estudos específicos e desdobramentos mais claros, segundo este autor, seus efeitos “podem e devem ser avaliados” considerando o valor agregado que a memória atribui aos bancos de dados de conteúdo jornalístico e as formas de interação com o usuário, ao permitir criar recursos de investigação dos aspectos históricos deste conteúdo.

No âmbito da Ciência da Informação, Smit e Tálamo (2006) discutem a noção de memória a partir da relação entre documentação e recuperação de informação. As autoras analisam a documentação como indicador de uma memória coletiva e universal que, pela síntese das propostas de Paul Otlet, propõe:

- O documento em si mesmo é registro de memória individual;
- A organização documentária encontra-se materializada em Sistemas de Recuperação de Informação, os quais diferem do componente

informacional homogêneo da mídia massiva [dualidade entre memória x comunicação] porque representam os conteúdos segundo ordens relacionais – diferença, semelhança, etc.;

- Os Sistemas de Recuperação de Informação e de Documentação apresentam-se como memória coletiva pois fundam-se em linguagem normalizada e socializada – a Linguagem Documentária (LD);
- Linguagem e memória encontram-se associadas: a memória expressa-se através da linguagem seja nos documentos seja no próprio Sistema de Recuperação de Informação.

Com esta abordagem, Smit e Tálamo (2006) concluem que na organização de conhecimento é preciso um sistema de recuperação que contribua tanto para sua preservação como para “ativação de recordações que permitem projeções no futuro” e a “socialização” das informações que constroem a memória.

CONECTIVIDADE

Outro aspecto importante pode ser associado à noção de memória digital é a conectividade, limitada nesta reflexão a capacidade tecnológica de criar pontes ou canais de contato entre pessoas e informações. Segundo Pierre Lévy (1999), a conectividade é composta por “todos os aparatos materiais que permitem a interação entre o universo da informação digital e o mundo ordinário”. Para comunicação entre estes dois espaços (universo digital e mundo ordinário) estar na rede nos parece condição primária. Neste contexto, redes sociais digitais adicionam à dinâmica de conectividade fatores como interfaces e arquiteturas que facilitam o empoderamento dos usuários por meio de fluxos de informações relevantes, bem como o livre compartilhamento de informação digital, potencializados por novas condições de mobilidade (celulares, tablets e acesso wireless) que, por sua vez, estabelecem um estado de “presente contínuo” derivado da possibilidade de acesso de qualquer lugar, da atualização permanente de informações e da conexão expandida com tudo e todos que estiverem disponíveis (DI FELICE; TORRES; YANAZE, 2012).

Em um ambiente de rede, por exemplo, a informação jornalística (conectada, ativada e recuperada) possui efeitos que, conseqüentemente, provocam desdobramentos no

processo de disseminação: maior volume de dados, velocidade de acesso, autonomia do usuário e dinâmica da relação produção/disseminação/consumo sobre a circulação de informações. Além disso, surgem também as “comunidades informacionais” que se estabelecem como elementos chave para dinamizar o fluxo de informação na Web, assim como promover a evolução do processo/método de disseminação em rede. Enquanto nas mídias tradicionais o fluxo de informação flui de forma linear, nas “comunidades informacionais”, constituídas em espaços como Facebook, Twitter e blogs, o fluxo ocorre on-line de modo multidirecional e compartilhado, gerando maior intensidade na circulação e propagação de informações (MACKINNON, 2004) e novas conexões de memória.

Neste contexto digital de redes, hiperlinks e multimídias, tecnologias são meios eficazes para conectar informações (memórias) e pessoas. Como resultado final desse poder das conexões, observamos uma apropriação inovadora de informação pelo usuário e a aplicação da tecnologia como um dos caminhos de transformação social. De acordo com Pariser (2012), as notícias colaboram para moldar nossa “visão do mundo, do que é importante, da escala, tipo e caráter dos problemas que enfrentamos” e sua importância está apoiada no fato de servir como “base das experiências comuns sobre a qual se constrói a democracia”. Se as notícias possuem este grau de importância e impacto como informação, a conectividade agregada ao valor da memória digital adquire também efeito social transformador. Para Torres²² (2008, citado por DI FELICE; TORRES; YANAZE, 2012), a conectividade, baseada em uma nova forma de “tecnologia comunicativa”, produz e fortalece “conexões de competências” (saberes) e compartilhamento de poder, implicando em “uma reelaboração coletiva das questões sociais e uma busca conjunta por soluções a partir de uma interação característica entre indivíduo e espaço”.

²² TORRES, J.C. Cyborgcracia: entre a gestão digital dos territórios e as redes sociais digitais. In: DI FELICE, M. (org.) **Do público para as redes**: a comunicação digital e novas formas de participação social. São Caetano do Sul: Difusão, 2008.

4. MECANISMO DE BUSCA E RECUPERAÇÃO DE INFORMAÇÃO NA WEB

Como ferramenta de mediação do acesso à informação, sistemas de recuperação de informação são a inovação mais interessante introduzida nas versões on-line de jornais diários. Muito embora as análises de mecanismos de busca registrem modificações técnicas, para maior usabilidade e acessibilidade à informação, poucos estudos contemplam uma análise de sua evolução, sua dinâmica de modificações, avaliações de uso e práticas de pesquisa definidas por usuários de conteúdo jornalístico. Trataremos nesta parte do trabalho dos conceitos operacionais que definem melhor a ferramenta e ajudam criar um panorama do processo de recuperação de informação relacionado à sua usabilidade, tema tratado no tópico seguinte.

4.1 CARACTERÍSTICAS GERAIS DE UM MECANISMO DE BUSCA

Um grande número de derivações nomeia a ferramenta de pesquisa escolhida para avaliação deste estudo: mecanismo de busca, motor de busca, sistema de busca, buscador, ferramenta de busca. Sendo assim, no intuito de evitar conflitos com o vocabulário derivado dos estudos sobre recuperação de informação e estabelecer um padrão, desconsiderando a derivação já aplicada e citações, adotaremos a partir de agora o termo mecanismo de busca.

A recuperação de informação na Web é determinada por sua estrutura, linguagem e unidades de informação (objetos). Basicamente, as unidades de informação são representadas por páginas que possuem um endereço (URL) para sua localização e acesso. A maioria das páginas é escrita em uma linguagem (HTML) formada por determinado conjunto de códigos (tags) que permite a definição da aparência e conteúdo de uma página, assim como o uso de uma ligação (link) entre outras páginas, criando conexões entre as unidades de informação. Esses links entre as páginas geram uma estrutura complexa de conexões com as mais diversas ordens (lógica, conceitual, por afinidade ou interesse) que constituem o ambiente Web (WWW, identificado assim por sua estrutura de “teia”) (FERNEDA, 2003, p. 92-93).

O crescimento da Web tornou humanamente impossível acompanhar informações dentro deste ambiente e ampliou necessidade de aplicar processos para recuperação automática de informações. É neste contexto que surgem os primeiros motores de busca que continuam evoluindo e exigindo modificações e ajustes.

Historicamente, foi a partir da década de 90 que surgiram em sequência os primeiros mecanismos de busca (global) da Web, entre eles: Excite, Infoseek e Alta Vista. O desenvolvimento desta tecnologia de pesquisa foi ampliado com o aparecimento do Google, lançado em setembro de 1998, e seu conceito de apresentação dos resultados, obtidos por um algoritmo, que estabeleceu uma relação entre o conteúdo indexado dos sites e a popularidade do conteúdo definida pelo número de citações que determinado link recebe de outros sites.

Antes mesmo da popularidade do Google, Michel Dertouzos, em seu livro “O Que Será: Como o Novo Mundo da Informação Transformará Nossas Vidas.”, fez previsões que continuam pertinentes para abordagem deste tópico. Segundo Dertouzos, o mercado de informação precisava de um guia para percorrer o gigantesco volume de informações oferecido pela Web. No entanto, estes guias não poderiam servir “apenas para juntar palavras e sites” mas, em analogia, precisariam “apresentar uma caixinha de jóias preciosas forrada de veludo” com respostas que pudessem responder perguntas com precisão (DERTOUZOS, p.69). Os mecanismos de busca evoluíram e se estabeleceram como estes guias de informação. Porém, mesmo com inovações, que envolvem arquitetura e recuperação de informação, a relevância e precisão das respostas “preciosas” obtidas (como resultado de uma busca) continuam comprometidas. Esta ineficiência, seja ela causada pelo volume de informações ou por melhores modelos de representação do conhecimento, decorrentes da falta de padrões técnicos e linguagem adequada do ambiente (Web), indica que muito ainda irá mudar e melhorar nos próximos anos.

Para Segaran (2008, p. 51-52), atualmente, o primeiro passo para criar um mecanismo de busca “é desenvolver uma forma de coletar documentos.” Para isso, em alguns casos, a técnica utilizada é denominada de *crawling* (rastejar, processo no qual se

coleta um conjunto de documentos e, em seguida, os links contidos nestes documentos que permitem obter outros documentos). Da mesma forma, este processo também se aplica a um conjunto fixo de documentos que, por exemplo, poderiam representar uma intranet corporativa. O passo seguinte é a indexação destes documentos, que envolve a criação de uma tabela dos documentos e as localizações de diferentes palavras que permitem gerar um índice com as referências (caminhos de um sistema de arquivos ou URL) de onde os documentos se encontram. O passo final é apresentar uma lista de documentos relevantes quando forem realizadas pesquisas. Neste sistema, Segaran (2008, p. 51-52) avalia que a construção de um índice torna bem simples o processo de recuperar um documento por determinado conjunto de palavras, no entanto, “a verdadeira mágica” na recuperação é como os resultados são ordenados e, conseqüentemente, quais as melhores métricas utilizadas para modificar a ordem dos resultados e melhorar um mecanismo de busca.

Definidos pela sua estrutura de organização das informações, os mecanismos de busca podem ser separados em dois tipos: diretórios (ou catálogos) e programas de busca. Diferentemente dos diretórios, que foram a primeira forma de organização de informação na Web (estruturada em hierarquias e categorias de assuntos indexadas manualmente, com a interferência humana), de acordo com Cendón (2001), os mecanismos de busca são desenvolvidos para permitir uma indexação automática do conteúdo da Web e (conforme Figura 1) são formados por quatro componentes:

o **robô** (também denominado *spider* ou *crawler*), programa que localiza e busca documentos na Web por meio de um algoritmo desenvolvido para rastrear links; o **indexador**, que extrai a informação dos documentos obtidos pelo robô e constrói um índice (base de dados); o programa de pesquisa (**mecanismo de busca**) deste índice; a **interface**, que permite a interação entre informação e usuário durante o processo de pesquisa.

Figura 1 - Arquitetura de um mecanismo de busca simples.

Fonte: Esquema traduzido de BAEZA-YATES, 1999.

Basicamente, estes componentes determinam a anatomia de um mecanismo de busca e ajudam entender as fases de funcionamento deste sistema:

captura (robô) > indexação (índice) > recuperação (interface de pesquisa) > resultados

A ordenação e apresentação dos resultados definem algumas categorias dos mecanismos de busca que, conforme Monteiro (2008), podem ser percebidas como programas de:

- Localização/frequência do termo: resultados são apresentados considerando as ocorrências de um termo pesquisado e sua localização (área de maior ou menor importância) na estrutura de um documento para definir sua relevância.
- Análise de links: resultados que consideram análises de citações dos links e suas conexões com outras páginas que agregam valor a determinado site e seu conteúdo.
- Agrupamento: apresenta os resultados agrupados por afinidades definidas por uma palavra em comum, sinônimos ou conceitos.
- Especializados: resultados definidos pela especialidade, setor ou assunto em que opera o mecanismo de busca.
- Personalizados: a partir de um perfil pré-determinado pelo interesse e comportamento do usuário cruza informações e apresenta resultados.

Estabelecidas as características da estrutura e a tipologia mais comum de um mecanismo de busca, cabe aqui esclarecer que o tipo de mecanismo tratado neste estudo é determinado por um programa utilizado para recuperação de informação na Web com característica de aplicação local (interna) e não global (como Google), ou seja, mecanismos de busca existentes e utilizados apenas em sites de jornais on-line. Dada a estrutura do conteúdo jornalístico dos sites, formado por diversos gêneros (entrevista, reportagem, crônica, crítica, entre outros) e setores (editorias como Política, Ciência, Cultura e Esporte), entendemos que esta ferramenta de pesquisa é melhor identificada na categoria de mecanismo de busca especializado. Sendo assim, uma ferramenta de pesquisa de operação local, proposta para recuperação de um conteúdo estruturalmente especializado, determina o objeto de pesquisa deste trabalho: avaliar o comportamento de um usuário especializado (jornalista) no uso do mecanismo de busca para recuperação de conteúdo jornalístico, base para suas atividades profissionais.

4.2 ASPECTOS DA RECUPERAÇÃO DE INFORMAÇÃO NA WEB

A definição do processo de recuperação de informação envolve divergências entre os autores e estudos da Ciência da Informação que desenvolvem suas análises por meio de enfoques baseados em etapas do processo: na atividade de **seleção**; na necessidade dos **usuários** ou no tratamento dos **documentos**. De forma muito simples, podemos descrever o processo de busca e recuperação de informação como uma operação que envolve um conjunto de documentos (objetos ou unidades de informação) selecionados a partir da necessidade de um usuário que utiliza determinada ferramenta de pesquisa para obter uma resposta as suas questões. A efetividade de realização desse processo, de acordo com Belkin²³ (1981, citado por ARAÚJO JÚNIOR, 2007), está caracterizada por: a) o usuário reconhece uma necessidade de informação e recorre a um mecanismo de recuperação da informação,

²³ BELKIN, N.J. Ineffable concepts in information retrieval. In: JONES, K.S. **Information retrieval experiment**. London: Butterworths, 1981. p. 44-58.

por exemplo, pesquisa uma coleção de textos e aguarda que o mecanismo satisfaça sua necessidade; b) o papel do mecanismo de recuperação é atender a demanda registrada pelo usuário; c) o usuário examina os textos recuperados pelo mecanismo e julga se sua necessidade foi satisfeita, parcialmente satisfeita ou não satisfeita. Aproveitando uma representação simplificada do processo de recuperação, utilizada por FERNEDA (2003, p. 15), podemos determinar e relacionar (Figura 2) alguns pontos importantes deste processo no ambiente Web.

Figura 2 Representação do processo de recuperação de informação

Fonte: Adaptado de FERNEDA, 2003, p.15.

A operação de recuperação será determinada pelas características que compõem cada etapa deste processo:

Documentos: Muito embora o estudo baseie-se apenas na recuperação de informação linguística (texto), o objeto que contém informação, neste caso um documento digital, poderia ser também um vídeo, um áudio, ou um infográfico. Neste sentido, a definição de documento segue o conceito de Buckland (1991) de “informação como coisa”, ou seja, um objeto que não denota apenas um registro textual nem uma definição restrita ao seu formato ou meio onde se desenvolve, mas se configura como documento pela abordagem de sua funcionalidade enquanto objeto de informação (ou coisa que contém uma informação).

Representação: No processo de recuperação, a indexação permite a representação do conteúdo dos documentos. Além da indexação temática do conteúdo, a estrutura sintática deste documento, apoiada em uma linguagem de marcação por *tags*, na sua grande maioria codificada em linguagem HTML, colabora para representação deste documento digital e a definição das chaves de acesso do seu conteúdo (por exemplo, componentes como título, autor e resumo). Nesta etapa, a análise do conteúdo e a definição de atributos de acesso permitem a comunicação entre a ferramenta de pesquisa e o conjunto de documentos. Por conta disso, retomando Méndez Rodriguez (2002, p. 229-267), os metadados (meta tags) são essenciais para representação de conteúdo, estruturação e conexão dos documentos em ambientes digitais, consequentemente favorecendo a recuperação de informação relevante.

Função de Busca: O mecanismo de busca, propriamente dito, executa esta função baseada na programação desenvolvida para recuperação de informação. Esta função está fortemente apoiada em um algoritmo. Do ponto de vista computacional, este algoritmo²⁴ representa um conjunto de regras ou passos definidos pela estruturação lógica de um problema (tarefa) que, desenvolvido como um programa, possa ser entendido e executado por um computador (MANZANO; OLIVEIRA, 2008, p. 7). As características técnicas para recuperação de informação são determinadas pelas funções de pesquisa desenvolvidas nesta fase. A funcionalidade do mecanismo dependerá da lógica de pesquisa aplicada (por exemplo, uso de operadores booleanos ou seleção por filtros/campos específicos) assim como das variáveis que permitirão uma classificação de relevância para apresentação dos resultados (por exemplo, definida pela ocorrência de termos em um documento, ou pela proximidade entre termos pesquisados, ou pela densidade de um documento, determinada pelo número de acessos dentro de um documento) (MACKINLEY, 1998, p. 132-151).

²⁴ Entendido como uma série de instruções para resolver um problema, no caso da comunicação digital, o algoritmo opera com o objetivo de eliminar informações indesejáveis e oferecer conteúdo relevante. Um procedimento baseado, em princípio, por critérios de noticiabilidade definidos pela empresa proprietária do algoritmo (SAAD; BERTOCCHI, 2012).

Expressão (Equação) de Busca: Diretamente relacionada à função de busca e integrada à anatomia do mecanismo, a expressão (ou equação) de busca é operada pela construção de uma interface de pesquisa. Steven Johnson (2001), reconhecido pensador do ciberespaço com formação em semiótica e literatura, define que a interface pode ser entendida como o mediador entre dois mundos (real e digital) e seus dois personagens (homem e máquina). A capacidade de tornar estes dois mundos visíveis e sensíveis um ao outro é o grande mérito da interface (JOHNSON, 2001). A partir desta noção, entendemos que a formulação de uma busca será intermediada por uma interface construída para reconhecer a representação técnica do documento e, por meio de recursos estruturais e visuais, permitir a comunicação entre o usuário e o conteúdo mais relevante de um conjunto de documentos. Dada a importância desta fase do processo de recuperação, iniciada pelo primeiro contato do usuário com as características que compõem a interface de pesquisa de um mecanismo de busca, parte da estratégia metodológica deste trabalho inclui a inspeção de alguns modelos de interfaces (Figura 3) desenvolvidas para pesquisa em sites jornalísticos e dos recursos de recuperação (campos de informação) utilizados neste processo.

Figura 3 - Formulário de busca do site Folha.com

busca

Procurar por
brasil

Seção
Site da Folha De até

Buscar

- Site da Folha
- Jornal Impresso
- Site da Folha - Ambiente
- Site da Folha - BBC
- Site da Folha - Ciência
- Site da Folha - Colunas
- Site da Folha - Comida
- Site da Folha - Cotidiano
- Site da Folha - Cotidiano - Ribeirão Preto
- Site da Folha - DW
- Site da Folha - Educação
- Site da Folha - Equilíbrio e Saúde
- Site da Folha - Esporte
- Site da Folha - Folhateen
- Site da Folha - Folhinha
- Site da Folha - Ilustrada
- Site da Folha - Ilustríssima
- Site da Folha - Mercado
- Site da Folha - Mundo
- Site da Folha - Opinião

Fonte: Tela de busca capturada do site Folha.com (<http://search.folha.com.br/search?=online>), 2013.

Para obter benefícios práticos no processo de recuperação de informação, consequentemente na eficiência de uma interface de pesquisa, Pérez Gutiérrez (2000) propõe a aplicação de uma gramática formal que ofereça suporte para estruturação de uma linguagem de pesquisa. Em geral, usuários acostumados com sistemas de

armazenamento e recuperação de informação (SRI) condicionam suas buscas em necessidades informacionais geradas por estados mentais (alguma ordem de curiosidade, insatisfação ou divergência informativa) que são materializados e expressos em linguagem natural. Basicamente, pensamos e construímos nossas buscas a partir de uma expressão ou conjunto de palavras. No entanto, precisamos formular a busca de forma que o SRI possa entender nosso raciocínio.

Para cumprir este objetivo, Pérez Gutiérrez (2000) esclarece que é necessário percorrer por três fases: 1ª - análise conceitual da pesquisa (ponderando os conceitos que se ajustam a necessidade de informação); 2ª - tradução da linguagem natural para uma linguagem de pesquisa, momento no qual o resultado dessa tradução se configura em uma “equação de busca”; 3ª - comparação da equação de busca com a representação dos documentos derivada do processo de indexação do SRI, fase na qual se verifica a sintonia entre equação, resultados e documentos.

Conforme Pérez Gutiérrez (2000), estas fases oferecem contexto para elaboração de uma gramática que, baseada na linguagem de pesquisa e equações de busca, pode ser descrita por duas propriedades:

Sintática - refere-se às regras para construção de equações e combinação das estruturas gramaticais de pesquisa, isto é, conjunto de termos e operadores booleanos (NOT, AND e OR) que, associados e expressos em equações, oferecem inúmeras possibilidades de relacionamento, representam um documento e permitem uma “ação linguística: transmitir uma informação, expressar opinião ou dar uma ordem”.

Semântica - refere-se ao significado das equações construídas, propriedade que cada equação de busca adquire ao expressar ou definir um sentido ao conjunto de documentos recuperados pelo SRI, considerando sua sintonia e adequação com a necessidade de informação original.

Muitas vezes, ao utilizar um SRI nos deparamos com casos de recuperação que indicam ruído (buscas que recuperam documentos não adequados as nossas necessidades de informação) e silêncio (buscas onde ocorre a ausência de documentos que sejam pertinentes a nossa necessidade informacional). Esta estrutura gramatical (sintática e

semântica) de pesquisa é pensada de forma a simplificar e rentabilizar a interação do usuário com SRI, permitindo perceber a qualidade dos resultados, em relação à necessidade de informação, assim como oferecer benefícios práticos como a redução de situações de tentativa e erro.

Em analogia a proposta de Pérez Gutiérrez, encontramos situação similar quando nos deparamos com uma interface de pesquisa que oferece elementos para criarmos uma equação de busca. Para refletir sobre este aspecto, basta utilizar alguns exemplos de possíveis equações de busca e resultados de pesquisa:

Pesquisa com operador AND: [alimentos AND transgênicos] > Resultado: apenas matérias que contenham as duas palavras (pesquisa específica que busca documentos que contenha os dois termos)

Pesquisa com operador OR: [alimentos OR transgênicos] > Resultado: matérias que contenham uma das duas palavras (pesquisa mais abrangente que busca um ou outro termo)

Pesquisa com operador NOT: [alimentos NOT transgênicos] > Resultado: matérias que contenham a primeira palavra e não a outra (pesquisa de exclusão que busca um termo e não o outro)

Pesquisa por “FRASE”: ["raízes do brasil"] > Resultado: recupera a frase, exatamente na ordem de termos como foi aplicada

Pesquisa por CAMPO que, apoiado na estrutura do conteúdo jornalístico e definido pela interface de pesquisa, orienta o agrupamento de informações com mesmo foco temático e produção editorial:

[campo TITULO: racismo] todas as matérias publicadas no jornal cujo título contém a palavra ‘racismo’.

[campo EDITORIA: Economia] AND [Field TITULO: “fome zero”] apenas matérias da editoria Economia cujo título contém a frase “fome zero”.

Considerando estes exemplos básicos, se utilizarmos, especificamente, a interface de pesquisa de um mecanismo de busca de jornais on-line, o exercício de raciocínio também se reflete na elaboração de uma equação de pesquisa adequada às necessidades de informação do usuário, ou seja, na construção de equações que explorem filtros e conexões entre palavras que permitam uma busca mais precisa, sem ruídos ou silêncios. Cabe acrescentar que a noção de equação de busca pode oferecer

uma medida imediata para identificar interfaces (mecanismos de busca) com maiores ou menores níveis de interação com o usuário. Esta medida pode ser definida pela capacidade informativa (atributos) que a interface oferece para criar equações com estruturas de pesquisas mais ou menos complexas.

Usuário: De acordo com Mckinley (1998, p. 151), qualquer pesquisa de textos segue 4 passos básicos: 1º) os termos são digitados e combinados conforme as necessidades do usuário; 2º) os termos podem ser expandidos ou reduzidos; 3º) a pesquisa é executada e os resultados apresentados ao usuário; 4º) o usuário visualiza os documentos recuperados, ou restringe sua consulta baseado na lista de resultados.

Estes são, em termos gerais e extremamente simplificados, os passos desta fase do processo de recuperação de informação, no qual o uso da interface e aplicação do mecanismo de busca é conduzido pelas necessidades e interesses do usuário. Pelas respostas obtidas em sua ação de busca, o usuário sabe se a ferramenta de pesquisa proporciona a resposta mais pertinente e eficaz com respeito a suas necessidades de informação ou se, pelo contrário, o mecanismo de busca devolve apenas uma parte dessa resposta ideal devido ao seu inadequado funcionamento, satisfazendo ou não suas necessidades de informação de forma suficiente (MARTÍNEZ MÉDEZ, 2002). Muito embora esta satisfação (ou insatisfação) seja uma medida importante para avaliar a recuperação de informação, devemos considerar também que esta medida carrega uma considerável dose de subjetividade que, implicitamente, acompanha o conceito de relevância adotado pelo usuário ao avaliar os resultados obtidos.

A ação de busca executada pelo usuário (consequentemente a avaliação dos resultados de pesquisa e documentos recuperados) será determinada por uma série de relações (manifestações de relevância) que podem ser resumidas como: de caráter técnico com origem no sistema (relação entre pergunta e informação obtida pelo sistema de recuperação de informação); temática (relação entre o assunto procurado e a informação recuperada); cognitiva (determinada pela relação entre o estado cognitivo de conhecimento do usuário e a informação desejada); situacional ou

utilitária (definida pela relação entre a situação, tarefa ou problema que precisa ser solucionado e a informação procurada) e finalmente afetiva (relação entre as intenções, metas, emoções e motivações do usuário e a informação desejada) (SARACEVIC, 2007).

Observadas as etapas do processo de recuperação de informação, descritas neste tópico, percebemos os possíveis eixos de análise para este processo e as atuais ferramentas de recuperação de informação na Web. Neste contexto, Martínez Médez (2002) considera que a eficácia destas ferramentas concentram-se principalmente em 4 grupos de análises:

- 1) Centrado na interface e nas possibilidades de recuperar informação
- 2) Centrado na avaliação dos mecanismos de busca desenvolvidos comercialmente
- 3) Baseados na medida de relevância dos documentos
- 4) Baseados em medidas que evitam o uso do conceito de relevância, devido à subjetividade deste conceito.

Paralelamente, Martinez Médez (2002) também sugere que a confiança no funcionamento dos mecanismos de busca esta concentrada numa série de problemas de recuperação de informação que, para possíveis análises e soluções, podem ser classificados da seguinte forma:

- 1) Formulação adequada da pergunta [busca]
- 2) Interatividade com a interface do usuário
- 3) Indexação inadequada dos documentos
- 4) Atualização dos índices do mecanismo de busca

Este quadro de problemas e linhas de pesquisa (interrelacionados) indicam o quanto é complicado definir um único viés de análise entre o uso da ferramenta e o usuário. Por conta disso, optamos pela análise das características de uso e atributos de busca presentes na interface de um mecanismo de busca de jornais. Desta forma, determinamos que o recorte pretendido neste estudo está identificado com a percepção do usuário sobre a usabilidade da ferramenta e a satisfação de suas necessidades, focadas na recuperação de informação.

4.3 USABILIDADE E AVALIAÇÃO DA QUALIDADE DE MECANISMOS DE BUSCA

Este tópico reúne noções e características gerais de usabilidade com a intenção de apontar atributos de avaliação, como medida da qualidade, para inspeção de mecanismos de busca local em sites de jornais.

Na definição estabelecida pela norma ISO 9241-11, usabilidade é “a medida na qual um produto pode ser usado por usuários específicos com eficácia, eficiência e satisfação num contexto específico de uso”. Tomaél (2008) complementa ainda que a usabilidade é formada por um conjunto de recursos que, desenvolvidos durante o planejamento e desenvolvimento de um site, ajuda “aprimorar e facilitar ao usuário a consulta e uso efetivo da informação”(ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2002).

A norma ISO 9241-11 ainda esclarece que o usuário, o contexto de trabalho (ambiente), a tarefa pretendida e as ferramentas utilizadas para atingir determinado objetivo, fazem parte dos componentes de análise da usabilidade, que será medida pela verificação da eficácia (precisão e abrangência), eficiência (capacidade de render bons resultados) e satisfação (conforto ou desconforto) no processo de interação usuário-ambiente-objetivo.

Em síntese, para Nielsen e Loranger (2007) usabilidade é:

“um atributo de qualidade relacionado à facilidade de uso de algo. Mais especificamente, refere-se à rapidez com que usuários podem aprender alguma coisa, a eficiência deles ao usá-la, o quanto lembram daquilo, seu grau de propensão a erros e o quanto gostam de utilizá-la. Se as pessoas não puderem ou não utilizarem um recurso, ele pode muito bem não existir.”

Para o caso específico deste estudo, da usabilidade voltada para operação de um mecanismo de busca, podemos encerrar as definições incluindo apenas entre elas as considerações de Thurow e Musica (2009), quando constatam que na Web a usabilidade de pesquisa está relacionada à facilidade de um usuário encontrar

determinado conteúdo via recuperação (busca/questão) e navegação, assim como pelo nível de satisfação obtido no processo de localização do conteúdo desejado.

Antes de considerarmos os aspectos de avaliação de usabilidade de um mecanismo de busca, propriamente dito, é preciso entender como ocorre a recuperação de uma informação neste contexto. De forma geral, a classificação da intenção de pesquisa na Web pode ser definida por 3 tipos de busca: navegacional (para localização de um site ou homepage), informacional (desejo de obter mais informação sobre determinado assunto) e transacional (centrada na realização de uma atividade, por exemplo, busca de informação ou algum dado específico dentro do ciclo de compra de algum produto). As buscas de caráter informacional, base deste trabalho, são realizadas por usuários cuja meta é obter uma informação sobre um tópico geral, que pode variar entre uma informação pontual e rápida ou de conteúdo mais analítico e profundo. Este tipo de busca compreende entre 48% e 80% das pesquisas na Web. (THUROW; MUSICA, 2009).

Este dado confirma as opiniões de Nielsen e Loranger (2007) para o declínio do uso de mecanismos de busca apenas para identificar bons sites. Para os autores, atualmente, os usuários procuram respostas e a Web se configurou como um recurso para encontrar páginas com informações relacionadas a uma necessidade específica. Esta mudança na intenção de pesquisa é denominada pelos autores de “teoria da coleta de informações” (ou da “caça de informações”) que, em resumo, propõe que quanto mais fácil for usar recursos de pesquisa, menos tempo o usuário gasta em cada recurso e mais motivado ele fica para realizar novas pesquisas.

Neste processo de busca, Thurow e Musica (2009) esclarecem que existem duas situações para encontrar informações na Web: fora de um site, utilizando um mecanismo de busca global, ou dentro de um site, utilizando o mecanismo de busca interno do site. Visto por esquemas (Figuras 4 e 5), as situações são compreendidas pelos autores da seguinte forma:

Figura 4 - Pesquisa de conteúdo na Web

Fonte: Esquema adaptado de THUROW; MUSICA, 2009.

Figura 5 - Pesquisa de conteúdo interno de um site

Fonte: Esquema adaptado de THUROW; MUSICA, 2009.

Sendo orientada pela execução de tarefas, com eficiência e alto nível de satisfação, a usabilidade do site, em particular do mecanismo de busca interno (local), está apoiada nas medidas de:

- Eficácia: podem os usuários encontrar facilmente a informação desejada?
- Eficiência: o usuário atinge rapidamente seu objetivo? são necessários muitos passos até atingir uma meta?
- Aprendizagem: é fácil e rápido aprender utilizar determinado recurso para localizar o conteúdo desejado?
- Memória: é fácil e rápido lembrar como funciona um recurso? ou repetir uma ação de pesquisa para localizar uma informação?
- Prevenção de erros: existem facilitadores para prevenir erros inesperados que compliquem a busca de uma informação?
- Satisfação: o usuário aprova e recomenda o uso de um site, baseado na sua satisfação de atingir sua meta?

Profissionais de usabilidade utilizam essas medidas para beneficiar um negócio (ampliando a qualidade de localização da informação de um site) e a relação do usuário com a interface (permitindo atingir metas pela execução adequada de tarefas) (THUROW; MUSICA, 2009).

A partir deste ponto, precisamos definir os aspectos necessários para identificar problemas, avaliar e medir a usabilidade de um mecanismo de busca. Por exemplo, Rosenfeld e Morville (1998, p. 99-129) recomendam algumas reflexões antes de iniciar qualquer aplicação destes mecanismos:

- O que está sendo procurado?
- Como o usuário pode formular as questões? Busca simples ou busca avançada?
- Qual o volume de informação desejado? Muito? Pouco? Informações mais relevantes?
- A organização dos resultados da pesquisa será apresentada de que forma? Em ordem cronológica? Em ordem alfabética? Por relevância?

Para Rosenfeld e Morville (1998, p. 99-129), estas são recomendações que deveriam anteceder a construção de um mecanismo de busca. Porém, como a evolução destes mecanismos se apresenta de forma desordenada, estes são pontos importantes para uma avaliação da usabilidade de mecanismos já em funcionamento.

Na literatura encontramos uma variedade de métodos utilizados para avaliação da usabilidade de um produto, entre eles: métodos de inspeção da usabilidade, testes de usabilidade centrados no usuário, experiências controladas realizadas em laboratórios (com controle de variáveis) e métodos interpretativos de avaliação, no qual o usuário participa da coleta e análise de dados. No caso deste estudo, a proposta de avaliação está apoiada em métodos de inspeção e na observação de heurísticas, que não envolvem usuários no processo de avaliação. Pertinente para esta reflexão, o método de inspeção, de acordo com Dias (2003, p. 49), favorece a inspeção de um conjunto de componentes, características ou módulos de um sistema (neste caso uma ferramenta de pesquisa) envolvido na realização de uma tarefa (de recuperação de informação). O foco da análise é checar a disponibilidade e utilidade de cada componente.

As heurísticas mencionadas podem ser entendidas como a expressão de conhecimento humano decorrente da vivência ou observação de algum tipo de situação ou problema. O uso correto deste conhecimento permite lidar de forma mais eficiente com uma dada situação. Sendo as heurísticas experiências do passado eles podem ser utilizadas para obter sucesso em ações no futuro. Por indução, se uma determina situação se repete várias vezes e gera sempre um mesmo resultado, diante de uma situação semelhante o resultado novamente se repetirá.

Jakob Nielsen (1994), reconhecido especialista em usabilidade, elaborou 10 heurísticas, referências para as análises deste estudo, que permitem identificar violações de práticas que melhoram e ampliam a usabilidade de um produto. São elas:

- 1) Linguagem simples e natural: apresentar exatamente a informação que o usuário precisa no momento, nem mais nem menos. A intenção é compatibilizar a interação (acesso aos recursos e operações) com o modo pelo qual o usuário realiza suas tarefas.

2) Falar a linguagem do usuário: a terminologia deve ser baseada na linguagem do usuário e não orientada pelo sistema. O modelo mental do usuário deve ser guia para organizar as informações.

3) Minimizar a sobrecarga de memória do usuário: o sistema deve mostrar elementos de diálogo e permitir que o usuário faça suas escolhas, sem a necessidade de lembrar um recurso específico.

4) Ser consistente: uma ação deve sempre ter o mesmo efeito; uma operação deve sempre ter a mesma localização e ser apresentada da mesma maneira para facilitar o reconhecimento.

5) Feedback: o sistema deve sempre informar ao usuário sobre o que ele está fazendo. 10 segundos é o limite para prender a atenção do usuário.

6) Saídas bem indicadas: o usuário que controla o sistema pode, a qualquer momento, abortar uma tarefa, ou desfazer uma operação e retornar ao estado anterior.

7) Atalhos: Possibilitar que usuários mais experientes executem as operações mais rapidamente. Atalhos servem para recuperar informações que estão distantes, na profundidade da estrutura navegacional, da interface principal.

8) Boas mensagens de erro: linguagem clara e sem códigos ajudam o usuário a entender e resolver o problema; não devem culpar ou intimidar o usuário.

9) Prevenir erros: evitar situações de erro conhecendo as situações que mais provocam estes erros e modificar a interface para que eles não ocorram.

10) Ajuda e documentação: o ideal é que um software seja tão intuitivo (fácil de usar) que não necessite de ajuda ou documentação; se for preciso a ajuda deve estar facilmente acessível on-line.

Este conjunto de recomendações são componentes relevantes para análise de usabilidade de um site e seus componentes. No entanto, o recorte de análise deste estudo está delimitado ao mecanismo de busca interna de sites jornalísticos. Para García de León e Caldera Serrano (2007), chamamos de mecanismos de busca interna os programas que examinam a informação existente no repositório de um único site e, a partir do conteúdo indexado e armazenado, podem ser pesquisados por meio de palavra(s) para recuperação de informação. A partir dessa noção, esclarecemos que este estudo analisa ferramentas com esta característica de aplicação e estritamente relacionadas ao conteúdo de um site jornalístico e, por isso, identificadas como mecanismo de busca interna ou local. Além disso, focamos a análise do dispositivo interno pela relação com usuários especialistas (jornalistas), familiarizados com o conteúdo noticioso, o que implica na busca por informações estruturadas no âmbito da produção jornalística.

Para ampliar a visão de análise da usabilidade desta tecnologia de pesquisa, proposta por García de León e Caldera Serrano (2007), incluímos e adaptamos uma lista de considerações sobre os principais fatores para avaliação da operação de um mecanismo de busca interna:

1. Permite a criação de interfaces de busca simples ou avançada;
2. Possibilita pesquisa cruzada e utilizar operadores booleanos, de proximidade ou outros operadores especiais;
3. Possibilita eliminar sinonímias e polissemias;
4. Permite uso de termos compostos (utilizando aspas e frases literais);
5. Reconhece caracteres em outros idiomas;
6. Permite recuperação baseada em vocabulário controlado e operações com conceitos relacionados (*clustering*);
7. Permite na recuperação diferenciar letras maiúsculas e minúsculas;
8. Permite omitir “palavras vazias”, contém lista de *stopwords* incorporada ou pode criar uma lista ou adaptar a existente;

9. Recupera indexação de outros tipos de arquivos (texto, vídeo, áudio)
10. Permite busca restrita por áreas do site (editorias, por exemplo);
11. Permite buscas restritas por campos (autor e título) da unidade de informação;
12. Permite realizar buscas definidas por data específica ou período de produção dos documentos.

Muitos dos itens da lista dependem de testes mais complexos e conhecimento da tecnologia aplicada. Porém, salientamos que estes referenciais foram adotados e dimensionados, como princípio de análise, para observação e elaboração de um quadro de inspeção de sites jornalísticos, descrito no capítulo 5 deste estudo.

Para concluir esta abordagem sobre usabilidade, que se apoia nos estudos de Jacob Nielsen, finalizamos com duas são conclusões de acordo com Vicens Arasanz (2008), sobre a importância de uso dos mecanismos de busca interna de sites:

1. provocam uma influência decisiva na primeira impressão do usuário do site pois, na presença desse elemento, provocam um impacto emocional baseado em fatores que sugerem um sistema seguro, profissional e amigável.
2. ao lado da arquitetura da informação e sistemas de navegação, mecanismos de busca são ferramentas cruciais para que usuários localizem a informação que desejam em um site de conteúdo mais amplo.

Por esta visão, a avaliação de qualidade está determinada pela relação entre interface, necessidade do usuário e elementos aplicados no processo de busca. Desta forma, a análise de um mecanismo de busca interna pode contribuir para que os criadores do site, assim como os geradores de conteúdo, estabeleçam medidas para avaliação do dispositivo de pesquisa e obtenham maior conhecimento sobre sua capacidade de recuperar informações. Consequentemente, esta avaliação pode se converter em um ponto de partida para modificações no arranjo do conteúdo e nas análises que indicam valor aos métodos e equações de pesquisas disponíveis em mecanismos de busca interna.

5. ANÁLISE DOS RESULTADOS

Este capítulo do estudo inclui resultados e interpretações das investigações empíricas que, desenvolvidas em duas etapas, apoiam a avaliação de mecanismos de busca interna de jornais on-line por meio das técnicas: inspeção de interfaces similares e questionário on-line.

5.1 INSPEÇÃO DE INTERFACES SIMILARES: MECANISMO DE BUSCA

Para esta etapa da pesquisa, realizamos um estudo exploratório desenvolvido por meio de simulação de busca e definição de “método de inspeção” em exemplos selecionados de buscadores de sites jornalísticos. Na inspeção dos mecanismos, proposta neste estudo, o objetivo foi avaliar especificamente elementos de informação que permitissem maior ou menor usabilidade dos mecanismos de busca de webjornais, desconsiderando a experiência do usuário de informação jornalística.

Para inspeção dos recursos de pesquisa mais comuns em mecanismos de busca de sites jornalísticos, escolhemos nove jornais on-line com reconhecida audiência:

The New York Times (<http://www.nytimes.com/>),
Wall Street Journal (<http://online.wsj.com>),
Washington Post (<http://www.washingtonpost.com/>),
El País (<http://www.elpais.com>),
Le Monde (<http://www.lemonde.fr/>),
The Guardian (<http://www.guardian.co.uk/>),
Folha.com (<http://www.folha.uol.com.br/>),
Estadão.com (<http://www.estadao.com.br/>),
O Globo (<http://oglobo.globo.com/>)

A partir do resultado desta avaliação, realizada entre a terceira e quarta semana de abril de 2013 e revisada em junho de 2014, foi elaborado um quadro de inspeção (Quadro 2) com os principais atributos de pesquisa presentes nestes importantes jornais digitais, apontando diferenças e semelhanças que permitem reforçar a necessidade de propor padrões e indicadores para análise destas ferramentas.

Várias são as práticas utilizadas em testes de usabilidade. Este exemplo utiliza uma das etapas de avaliação que, de acordo com Andrade (2007, p. 46), dispensa o uso de um grupo de avaliadores e se constitui como passo inicial para desdobramentos de técnicas e análises dos mecanismos de recuperação de informação na Web, já observadas na pesquisa de Mannarino (2000).

Alguns recortes, já mencionados nos procedimentos metodológicos, foram estabelecidos para favorecer esta inspeção:

- Escolha de três jornais on-line representativos de três regiões (EUA, Europa e Brasil) com destacada audiência Web, sites com mais de 15 anos de funcionamento e, apoiados na tradição impressa, com reconhecida confiabilidade no mercado e circulação (em papel) superior a 250 mil exemplares.²⁵
- Não são pontos considerados nesta avaliação: a lógica comercial dos sites selecionados, sua estrutura tecnológica e o perfil do usuário da ferramenta. Muito embora estes fatores contribuam nas decisões de construção dos mecanismos, este recorte foi necessário para evitar implicações diversas e desconhecidas.
- A inspeção está baseada nos principais elementos de identificação da informação jornalística e presentes em cada unidade de informação produzida: data e período (de publicação da notícia), título, autor e área editorial (editoria).
- São consideradas apenas as unidades de informação em formato textual. Demais formatos, como áudio e vídeo, não fazem parte da análise, mas foi considerada a existência de recursos para seleção destes formatos durante o processo de busca.
- A inspeção utiliza os elementos de informação (ver Figura 6) como parâmetros de análise e identificação dos recursos existentes para uma busca simples ou avançada, sem diferença entre elas, permitindo uma visão geral da dinâmica de pesquisa em sites jornalísticos.

²⁵ Critérios amparados em informações institucionais divulgadas nos sites jornalísticos selecionados.

Figura 6 - Exemplos de áreas de inspeção em mecanismos de busca

The figure displays four examples of search engine interfaces, each with specific areas highlighted for inspection:

- The New York Times:**
 - Selecção por datas específicas e períodos:** Points to the 'Date Range' sidebar filter.
 - Selecção por formato:** Points to the 'Format' dropdown menu.
 - Filtros por editoria e assunto (clustering):** Points to the 'By Editor' and 'By Topic' filters.
 - Selecção por autoria:** Points to the 'Author' filter.
- Estadão:**
 - Filtros por editoria e assunto (clustering):** Points to the 'By Editor' and 'By Topic' filters.
- El País:**
 - Tags e textos relacionados:** Points to the 'Related tags and texts' section at the bottom.
- Le Monde:**
 - Tamanho do texto:** Points to the 'Text size' filter.
 - Operadores booleanos:** Points to the 'Boolean operators' (et, ou) filter.
 - Selecção por partes do texto: título:** Points to the 'Title' filter.

Fonte: Elaboração do autor. Telas capturadas das páginas de busca dos jornais online The New York Times, El País, Le Monde e Estadão, 2014.

Quadro 2 - INSPEÇÃO DE MECANISMOS DE BUSCA DE SITES JORNALÍSTICOS

FILTROS DE PESQUISA (campos ou recursos aplicados)	NYTimes	WSJ	WP	Folha	Estadão	O Globo	El País	Le Monde	Guardian
Data	X	X	X	X	-	-	-	X	-
Período > seleção livre	X	X	-	X	-	-	-	X	-
Período > seleção determinada pelo sistema	X	X	X	-	-	-	-	X	-
Título	-	-	-	-	-	-	-	X	-
Autor	X	-	-	-	-	-	-	X	-
Editoria (área)	X	-	-	X	X	-	-	-	-
Seleção por fonte (impressa ou on-line)	-	X	-	X	X	-	-	-	-
Seleção por formato (artigos,vídeo, áudio, blog)	X	X	X	-	-	X	-	-	-
Operadores booleanos > seleção livre	X	-	-	-	-	-	-	X	X
Operadores booleanos > seleção determinada pelo sistema	X	X	X	X	X	X	X	X	X
Recurso automático de filtros (clustering)	X	-	X	-	X	X	-	-	-
RESULTADOS PESQUISA (organização e exibição do conteúdo)									
Indica nº de textos localizados	X	X	X	X	X	-	X	X	X
Indica tipo de fonte ou formato	X	X	X	X	X	X	X	X	-
Indica tamanho do arquivo	-	-	-	-	-	-	-	X	-
Organização por relevância	X	X	X	-	-	-	X	X	X
Organização cronológica	X	X	X	-	X	X	X	X	X
Tags de assuntos ou textos recomendados	X	-	X	-	X	-	X	X	X
Links patrocinados	X	X	-	X	-	-	-	-	X
Instruções de pesquisa (help)	X	-	-	-	-	-	-	-	-
Contato para reportar erros ou inconsistência de pesquisa	X	-	-	-	-	-	-	-	-
LOCALIZAÇÃO DO MECANISMO DE BUSCA NO SITE									
Busca simples na 1ª pag	X	X	X	X	X	X	X	X	X
Localização no site lado esquerdo superior	X	-	-	-	-	-	-	X	-
Localização no site lado direito superior	-	X	X	X	X	X	X	-	X
Localização no site centro superior	-	-	-	-	-	-	-	-	-
Seleção por tipo de busca (simples e avançada)	unificada	unificada	unificada	distintas	unificada	unificada	unificada	unificada	unificada
Banco de dados para pesquisa de matérias antigas	X	X	X	X	X	-	X	X	-

Fonte: Elaboração do autor, 2014.

A inspeção permitiu a elaboração do Quadro 2, que contém dados gerais dos recursos de pesquisa, elementos de visualização dos resultados e posição da ferramenta de pesquisa no site. Para detalhamento da análise de inspeção foram elaboradas três tabelas reduzidas (1 a 3) com as seguintes considerações:

Os recursos indicados no Quadro 2 e Tabelas 1 e 2 possuem duas formas de ocorrência:

- seleção livre determinada pelo usuário, que realiza a pesquisa e define os parâmetros da busca;
- seleção determinada pelo sistema, que oferece opções de refinamento de busca antes ou após apresentação dos resultados.

A seleção determinada pela idade do documento é favorecida pela opção de pesquisa por período, determinada pelo usuário ou definida pelo sistema (por exemplo: edição do dia, da semana, do mês, ultimo ano). No caso da seleção por data, os jornais brasileiros avaliados não apresentam possibilidades de recuperação do conjunto de matérias publicadas em uma mesma data. O usuário que procura apenas as notícias do dia anterior precisaria percorrer a lista de resultados, exibida em ordem cronológica decrescente, para localização de documentos recentes. Esta condição impede que o usuário reduza os resultados de pesquisa, ou então, que consiga monitorar todos os textos que formaram a cobertura do dia.

Tabela 1 - Filtros de Pesquisa: Avaliação Geral

FILTROS DE PESQUISA (campos ou recursos aplicados)	% de ocorrência do recurso no mecanismo
Data	56
Período > com seleção livre	44
Período > com seleção determinada pelo sistema	44
Título	11
Autor	22
Editoria (área)	33
Seleção por fonte (impressa ou online)	33
Seleção por formato (artigos, vídeo, áudio, blog)	44
Operadores booleanos > com seleção livre	33
Operadores booleanos > com seleção determinada pelo sistema	100
Recurso automático de filtros (clustering)	44

Fonte: Elaboração do autor, 2014.

Analisando a Tabela 1, percebemos que a minoria dos jornais (33%) propõe uma seleção entre a versão impressa do jornal e a produção jornalística on-line. A distinção entre as duas mídias (papel e on-line) implica em decisão do usuário sobre a equivalência e relevância do conteúdo durante a busca.

Devido a pluralidade de formatos de comunicação que um site de notícias reúne, parece pouco representativa (44%) a seleção de conteúdo em outros formatos (áudio, vídeo, infográfico, blog etc). Na ausência de filtros para diferentes formatos, a busca e recuperação de informação nestes canais são conduzidas pela navegação retrospectiva, sem possibilidade de recuperação do conteúdo descrito, que identifica cada elemento de informação.

A Tabela 1 também evidencia que ainda é muito baixa a presença de elementos considerados como itens clássicos para recuperação de informação. Esta condição reduz o fator de usabilidade do mecanismo e a facilidade na recuperação da informação. Os elementos estabelecidos como filtros de pesquisa (data, título, autor e editoria) não apresentam padrão ou presença predominante nos três jornais de cada área geográfica inspecionada (ver também Tabela 2 abaixo). Os jornais brasileiros não possuem elementos de informação suficientes para uma busca refinada, o que implica imprecisão no processo de busca e maior esforço do usuário para recuperação de informações.

Tabela 2 - Filtros de Pesquisa: Avaliação por Região

FILTROS DE PESQUISA (campos ou recursos aplicados)	RESULTADO POR REGIÃO ocorrências: jornais que utilizam recurso		
	USA	Brasil	Europa
Data	3	1	1
Período > com seleção livre	2	1	1
Período > com seleção determinada pelo sistema	3	0	1
Título	0	0	1
Autor	1	0	1
Editoria (área)	1	2	0
Seleção por fonte (impressa ou online)	1	2	0
Seleção por formato (artigos,vídeo, áudio, blog)	3	1	0
Operadores booleanos > com seleção livre	1	0	2
Operadores booleanos > com seleção determinada pelo sistema	3	3	3
Recurso automático de filtros (clustering)	2	2	0

Fonte: Elaboração do autor, 2014.

Mesmo relacionados aos demais itens, não é relevante a porcentagem/ocorrências de recursos para refinamento de pesquisa que foram observados na inspeção e apresentados nas duas tabelas (1 e 2): seleção por fonte e formato, uso de operadores booleanos e presença filtros automáticos/clustering²⁶ (definidos pelo sistema de busca para apresentar categorias especiais de filtros).

Tabela 3 - Resultados de Pesquisa: Avaliação Geral

RESULTADOS PESQUISA (organização e visualização do conteúdo)	% de ocorrência
Indica nº de textos localizados	89
Indica tipo de fonte ou formato	89
Indica tamanho do arquivo	11
Organização por relevância	67
Organização cronológica	89
Tags de assuntos ou textos recomendados	67
Instruções de pesquisa (help)	11
Contato para reportar erros ou inconsistência de pesquisa	11

Fonte: Elaboração do autor, 2014.

A apresentação dos resultados de pesquisa indicados na Tabela 3 também reflete estruturas de usabilidade (arquitetura da informação, localização, suporte a tarefas) deficientes para organização dos registros recuperados e, conseqüentemente, para apoio ao processo cognitivo de visualização²⁷ da informação, pesquisa e orientação do usuário.

²⁶ O agrupamento de documentos ou clustering é um recurso utilizado no processo de recuperação da informação e auxilia a organização dos resultados decorrentes da pesquisa de um usuário. Em pesquisas simples, com dois ou três termos empregados, os resultados são amplos e diversificados. Uma solução para este caso é a aplicação de técnicas automáticas de clustering que permitem categorizar por temas os resultados obtidos. A aplicação desta técnica de agrupamento pode ocorrer antes ou depois da tarefa de recuperação realizada pelo motor de busca (MATEOS SÁNCHEZ; GARCÍA-FIGUEROLA PANIAGUA, 2009).

²⁷ Assim como García Santiago (2003), entendemos que a “visualização de informação facilita as tarefas de busca, recuperação e tomada de decisões”. Neste sentido, consideramos importante o uso de diferentes elementos/estratégias para visualização de informação que, de forma geral, apresentem dados já organizados, de forma simples e dinâmica, e que representem a informação recuperada expandindo a capacidade do usuário para interação e análise do repositório de informação jornalística.

5.2 QUESTIONÁRIO: RECUPERAÇÃO DE INFORMAÇÃO EM JORNAIS ON-LINE

Em decorrência da etapa de inspeção de interfaces similares, visando ampliação da análise exploratória sobre a capacidade de pesquisa de mecanismos de busca interna de jornais on-line, elaboramos um questionário para investigação empírica da percepção profissional sobre a usabilidade do mecanismo de busca dos três jornais brasileiros adotados na fase de inspeção.

Produzido em formulário Google (ver Apêndice A), para facilitar aplicação e organização das questões, o questionário foi estruturado de maneira a criar uma transição de análise: inicia com um conjunto de reflexões sobre a aplicação e uso do dispositivo de pesquisa e termina com a inserção das duas questões, norteadoras do estudo, que estão relacionadas à adequação dos atributos de pesquisa e satisfação no uso profissional dos mecanismos de busca investigados.

Para análise dos resultados, e maior relacionamento das partes que ordenam as questões, propostas na metodologia, agrupamos os dados em três blocos de informações:

Quadro 3 - Esquema de análise do questionário on-line

Fonte: Elaboração do autor, 2015.

APLICAÇÃO NA PRÁTICA PROFISSIONAL

Os dados demográficos que caracterizam o perfil dos respondentes (Tabela 4) indicam que, dos 33 profissionais da comunicação participaram da pesquisa, 60,6% são do sexo masculino e formados principalmente por dois grupos de pessoas na faixa etária de 25 a 30 anos (45,5%) e 34 a 44 anos (36,4%). A maioria dos respondentes (84,8%) reforça o panorama de convergência e atua nas áreas de impresso e on-line. Esta condição predominante é impulsionada pelas estratégias de fusão e compartilhamento de tarefas, entre redações de conteúdo on-line e impresso, praticadas por grandes empresas de comunicação.

Tabela 4

PERFIL (Questões 1 a 6) %			
Sexo		Veículo de comunicação onde atua	
Masculino	60,6	Estadão	45,5 (15 respondentes)
Feminino	39,4	Folha	51,5 (17 respondentes)
		O Globo	3 (01 respondente)
Idade		Tempo de experiência profissional	
18 a 24	9,1	1 a 5 anos	39,4
25 a 34	45,5	6 a 10 anos	12,1
35 a 44	36,4	11 a 15 anos	21,2
45 a 54	9	16 a 20 anos	21,2
acima 55	0	mais de 20	6,1
Área de atuação na atividade jornalística			
on-line	9,1		
Impresso	6,1		
on-line e impresso	84,8		
Em que medida utiliza o mecanismo de busca do jornal on-line onde você trabalha			
uso bastante	66,7		
mais ou menos	21,2		
uso pouco	12,1		

Fonte: Elaboração do autor, 2015.

Os dados obtidos pelo questionário (Tabela 4) são percepções geradas por profissionais dos jornais Folha (51,5%) e Estadão (45,5%). O tempo de experiência desses profissionais apresentou uma variação positiva que observamos em dois agrupamentos: grupo com pouca experiência profissional (39,4% com 1 a 5 anos de

atuação) e grupo com alto nível de experiência profissional (42,4% de 11 a 20 anos de atuação) que, em sua maioria (66,7%), são profissionais com elevado comportamento de uso do mecanismo de busca do jornal on-line onde atuam.

Com relação ao uso do mecanismo de busca em pesquisas diárias dos jornalistas, esclarecemos que, para este bloco de perguntas, consideramos dois pontos importantes:

O primeiro está relacionado ao fluxo de informação de um jornal. Conforme Mena e Pliger (2011), nas engrenagens de um jornal a informação percorre cinco fases principais: Pauta (chegada das informações no jornal por várias entradas: agência de notícias, correspondentes nacionais e internacionais, reportagem, colunistas etc.); Avaliação e Checagem (da veracidade e relevância da informação; informações recebidas são avaliadas pela redação do jornal e definido interesse de publicação); Produção e Redação (momento em que a informação analisada e selecionada é transformada em reportagem pela redação do jornal); Edição (momento em que são incorporados ao texto produzido outros elementos de acabamento final: fotografias, infográficos, textos complementares, diagramação das unidades de informação); Publicação (momento de disseminação da informação nas diversas plataformas: redes sociais, celular, site, versão impressa e digital)

Desse fluxo de informação²⁸, escolhemos como recorte de análise o momento de Produção e Redação das matérias que, embora complexo e entrelaçado com a fase de Edição, possui na atividade de pesquisa sua principal tarefa.

²⁸ Em recente mudança na abordagem de notícias, com destaque para plataformas digitais, o jornal O Globo divulgou em abril de 2014 um quadro, denominado “o ciclo de conteúdo”, que reforça o recorte adotado e descreve a dinâmica para início da edição on-line do dia. A partir da escolha de um fato até a transformação em reportagem, a notícia é disseminada nas plataformas digitais e redes sociais em: 1) primeiros flashes (informações básicas) e 2) ampliação da informação (levantamento de dados que complementam os flashes); a partir deste ponto, a cobertura do fato ganha profundidade e o conteúdo é publicado no tablet, no site e preparado para versão impressa, finalizando o ciclo em: 3) aprofundamento (busca de fatos correlatos para melhor compreensão do assunto) e 4) enriquecimento (produção de materiais que permitam uma visão analítica da notícia) (GLOBO, 2014).

O segundo ponto está relacionado aos princípios do jornalismo investigativo, propostos por Philip Meyer (1998), que contêm etapas do processo de pesquisa jornalística aplicável nas questões que envolvem o jornalista como filtro, transmissor, organizador e intérprete de informações. Por isso, além de colocar um texto on-line e saber como preparar conteúdo relevante para o leitor, um jornalista também precisa ser um processador e analista de dados. Neste contexto, entre as competências do corpo jornalístico de conhecimentos, destacamos os seguintes passos de investigação: como encontrar a informação; como avaliar e analisar a informação encontrada e como comunicar a quem precisa desta informação.

Baseados nestas referências perceptuais do processo de pesquisa jornalística, recorreremos ao modelo que Kuhlthau (2004) elabora para identificar estágios do processo de busca por informação (Iniciação, Seleção, Exploração, Formulação, Coleta e Apresentação) e consideramos, especificamente, as fases definidas pela autora como apropriadas para tarefa de pesquisa: reconhecimento da necessidade de informação; identificação de um tópico geral de pesquisa; investigação de informação sobre este tópico; formulação de um foco de pesquisa; reunião de informações pertinente ao foco; por fim, completar a busca por informação. A partir dessas considerações, definimos as seguintes tarefas como pontos de análise para o estudo: reconhecimento da necessidade de informação a partir da relação entre rotinas do trabalho jornalístico e tarefas de pesquisa; definição do principal tópico dentro de algumas categorias de pesquisa²⁹; identificação dos principais recursos oferecidos pelo mecanismo de busca para investigar/formular a pesquisa e reunir informações; por fim, completar a pesquisa pela apresentação dos resultados e satisfação no uso do mecanismo de busca.

Por essas diretrizes, elaboramos questões para identificar os principais comportamentos de pesquisa e exploração do conteúdo jornalístico, e obtivemos (Tabela 5) os seguintes resultados:

²⁹ As categorias de pesquisa adotadas foram definidas pela experiência profissional do autor, com base nas dúvidas mais comuns de jornalistas que foram registradas por Pinto (2009).

69,7% usam com maior frequência o mecanismo de busca para coletar informações que apoiam a produção de uma pauta (entrevistas, perfis, textos de memória, matérias retrospectivas), ou seja, que permitem reunir referenciais ou conjunto de informações relevantes para elaboração e consolidação de uma nova proposta de matéria.

60,6% utilizam para analisar uma ideia de pauta (verificar o ineditismo da proposta, qual a cobertura já realizada ou ângulos de abordagem já publicados), isto implica que o mecanismo de busca é utilizado, já na primeira fase de pesquisa, para elaboração de uma hipótese de trabalho.

57,6% usam com frequência para checar informações (números, grafias, datas) utilizadas na produção de matérias, neste sentido, utilizam o conteúdo já publicado para manter um padrão de qualidade e evitar equívocos que exijam posteriores correções ou notificação de erros (nota de erramos).

54,6% utilizam para relacionar conteúdo (associar matérias recentes aos links de textos históricos ou de apoio, que ajudem ampliar e explicar o tema abordado; construção de links no formato "saiba mais" ou "leia também") e, portanto, contextualizam o tema tratado ao leitor e promovem conexões de informação a partir da produção existente.

51,5% buscam recuperar diferentes formatos de informação (mapa, ilustração, infográfico interativo, tabela, gráfico, vídeo, áudio e outros formatos anexados ao texto); este indicador entra em conflito com a interface de busca dos jornais brasileiros inspecionados; neste caso, embora a busca por diferentes formatos (não textuais) seja elevada, os mecanismos de busca não possuem atributos suficientes para filtrar e recuperar todos os recursos informacionais existentes em seus repositórios; desta forma, fica comprometida a associação integral (relação parte-todo) de todas as unidades de informação que compõem um documento com estrutura hipermídia.

15,2% utilizam o mecanismo para aplicar e controlar tags (pesquisar tags/palavras-chave para identificação de novos textos, pensar palavras-chave que evitem ruído de informação e promovam relacionamento com outros textos), fica evidente o baixo uso do mecanismo de busca nessa tarefa, em contraste com a indicação de que 42,4% não

utilizam para esta finalidade. Desconsiderando aspectos tecnológicos que determinem a importância ou nulidade dessa tarefa, a partir da reflexão de Saad e Bertocchi (2012), levamos em conta na análise o papel de comunicador e seu domínio sobre o processo de *tagging* do conteúdo gerado, bem como suas competências para tornar o conteúdo visível ao mecanismo de busca e promover aproximação entre produtor e usuário da informação, por meio de um processo comunicativo que incorpora elementos (tags) com sentido semântico, essenciais ao desenvolvimento da Web semântica. Em síntese, é uma tarefa não incorporada na rotina profissional e, pela perspectiva de García Gutiérrez (1990), para investigações da análise documentária, com função pouco relevante para servir como interprete entre o produtor de informação e o analista documentador preocupado em extrair características do documento com o objetivo específico de recuperação da informação.

Tabela 5

USO DO MECANISMO DE BUSCA EM PESQUISAS DIÁRIAS %					
CONSIDERANDO SUA ROTINA DE TRABALHO, COMO CLASSIFICA O USO DO MECANISMO DE BUSCA INTERNO DO SITE JORNALÍSTICO PARA REALIZAR AS SEGUINTE TAREFAS:					
Questões 7 a 12	não uso				uso sempre
	1	2	3	4	5
CHECAR Informações	6	15,2	21,2	9,1	48,5
COLETAR informações	0	12,1	18,2	24,2	45,5
Recuperar OUTROS FORMATOS	21,2	12,1	15,2	33,3	18,2
Analisar uma IDEIA DE PAUTA	9	15,2	15,2	18,2	42,4
RELACIONAR conteúdo	9	21,2	15,2	9,1	45,5
Controlar e aplicar TAGs	42,4	24,2	18,2	9,1	6,1

Fonte: Elaboração do autor, 2015.

Ao tratar do processo de transformação de uma pauta em reportagem, nas observações de Pinto (2009), quando um jornalista recebe uma pauta, o processo de pesquisa é baseado no levantamento de informações publicadas que permitam recolher dados sobre o assunto e, portanto, não repetir enfoques; descobrir pessoas (fontes) que entendem do tema; e compreender aspectos técnicos de algum assunto para entrevistar melhor uma fonte especializada. Este contexto de trabalho cria, naturalmente, classes comuns de pesquisa que aplicamos no questionário.

Percebemos que a busca por categorias de pesquisa possui uma grande variação na escala de frequência, entretanto, alguns pontos (Tabela 6) podem ser destacados:

51,5% buscam sempre por conteúdo retrospectivo (categoria Memória) para recordar um fato, realizar relacionamento histórico, esclarecer ou acompanhar eventos. Este resultado reforça considerações (tratadas no tópico 3.3) sobre o atributo de memória presente na informação jornalística que, segundo Palácios (2010), reflete o efeito da digitalização no processo de conexão de informações, visto que tornou mais fácil para o jornalista incorporar elementos de memória na produção e estruturação do texto, favorecendo a relação hipertextual de conteúdo do passado e do presente, assim como a produção de reportagens especiais de “cunho memorialístico”.

existe maior frequência de busca (somatória das escalas 4 e 5) nas seguintes categorias: Pessoas (buscam por conteúdo biográfico, fontes ou referências de especialistas sobre determinado tema); Significados e Explicações (buscam por informações que esclareçam ou expliquem determinado conceito, teoria, tema, caso, assim como técnica, uso ou funcionamento de algo);

indicam menor frequência de busca (somatória das escalas 2 e 3) as seguintes categorias: Números (informações estatísticas ou indicadores sobre determinado tema) e Geografia (informações sobre determinado lugar: bairro, cidade, país ou região); Legislação (referenciais legais sobre determinado tema apoiado em leis, decretos ou códigos).

Tabela 6

CONTEÚDO E CATEGORIAS DE PESQUISA %					
CONSIDERANDO SUA BUSCA POR CONTEÚDO, COMO VOCÊ AVALIA AS SEGUINTE CATEGORIAS DE PESQUISA EM SUA ROTINA:					
Questões 13 a 18	não uso				busco sempre este conteúdo
	1	2	3	4	5
PESSOAS	15,2	24,2	18,2	18,2	24,2
MEMÓRIA	3	18,2	15,2	12,1	51,5
LEGISLAÇÃO	39,4	15,2	24,2	9,1	12,1
NÚMEROS	24,2	18,2	21,2	9,1	27,3
GEOGRAFIA	21,2	27,3	18,2	9,1	24,2
SIGNIFICADOS E EXPLICAÇÕES	15,2	18,2	18,2	12	36,4

Fonte: Elaboração do autor, 2015.

RECURSOS DO MECANISMO DE BUSCA

Em avaliação geral (Tabela 7), existe um elevado grau de importância para maioria dos filtros de pesquisa, no entanto, alguns indicadores podem ser destacados:

Filtros para data ou período (94%) e título (76%) são indicados como importantes elementos de busca, porém, de acordo com o quadro de inspeção das interfaces de pesquisa, não são um padrão presente nos jornais brasileiros analisados. Na mesma condição estão tags / palavras-chave (73%); outros formatos (55%), por exemplo, vídeo, áudio ou infográfico e (43%) relevância da notícia, definida na pesquisa por valores como texto de grande repercussão, matéria premiada ou reportagem especial.

52% avaliam como moderado o grau de importância do gênero jornalístico (entrevista, crítica, reportagem) em estratégias de busca, assim como (49%) o uso de operadores especiais (por exemplo, para exclusão de termos). Estes filtros não estão presentes nas atuais interfaces de pesquisa e dependem de uma aplicação específica como recurso, no entanto, são tipos de filtros considerados pelos respondentes como necessários para refinamento das estratégias de pesquisa.

Tabela 7

RECURSOS DA FERRAMENTA DE PESQUISA %			
CONSIDERANDO SEU PROCESSO DE PESQUISA PARA FILTRAR INFORMAÇÕES JORNALÍSTICAS, COMO VOCÊ AVALIA A IMPORTÂNCIA DAS SEGUINTE FUNÇÕES EM UM MECANISMO DE BUSCA DE JORNAL ON-LINE:			
Questão 19	sem importância	mais ou menos	Importante
AUTOR	3	33	64
TÍTULO	3	21	76
DATA OU PERÍODO	3	3	94
EDITORIA	3	33	64
GÊNERO JORNALÍSTICO	15	52	33
TAGS / PALAVRAS-CHAVES	3	24	73
FORMATOS	9	36	55
OPERADORES	12	49	39
RELEVÂNCIA	24	33	43

Fonte: Elaboração do autor, 2015.

Esta parte do questionário (Tabela 8) levou em consideração os principais elementos de identificação das unidades de informação textual (autor, título, editoria e data da matéria) e indicam que data (78,8%) e título (87,9%) são os elementos de visualização mais importantes. Este aspecto demonstra uma pequena inversão de importância em relação aos mesmos recursos (data e título) aplicados para filtrar informações durante o processo de pesquisa. No caso da visualização dos resultados, os títulos (palavras em ordem lógica que esclarecem o objeto da notícia e atraem o leitor) e datas (indicador cronológico de fatos do passado e do presente) são os principais elementos de influência para decidir qual texto consultar. Com impacto menor, (66,7%) trechos de textos exibidos nos resultados entram também na mesma escala de importância. Para opção de resultados com exibição de tags/palavras chave, geradas para síntese e relacionamento de conteúdo, não percebemos na escala de importância um indicador predominante que destaque a visualização de tags como um elemento influenciador para seleção de textos relevantes. Além disso, o tamanho do texto (pequeno, médio ou grande) que, em princípio, dimensiona visualmente o conteúdo recuperado e sugere maior qualidade aos textos com tratamento em profundidade, não possui destacado grau de importância para influenciar escolhas de leitura.

Tabela 8

RESULTADOS DE PESQUISA E VISUALIZAÇÃO DA INFORMAÇÃO %			
APÓS REALIZAR UMA PESQUISA, QUAIS INFORMAÇÕES EXIBIDAS NO RESULTADO DE BUSCA AUXILIAM NA SUA DECISÃO DE ABRIR UM TEXTO ON-LINE PARA LEITURA:			
Questão 20	sem importância	mais ou menos	Importante
AUTOR	24,2	48,5	27,3
TÍTULO	3	9,1	87,9
EDITORIA	9	45,5	45,5
DATA	3	18,2	78,8
TAMANHO DO TEXTO	57,6	30,3	12,1
TAGS / PALAVRAS-CHAVES	33,3	30,3	36,4
TRECHO DO TEXTO	12,1	21,2	66,7

Fonte: Elaboração do autor, 2015.

PERCEPÇÃO COMO PESQUISADOR (SATISFAÇÃO DO USUÁRIO)

O perfil de pesquisador é traçado por afirmações dos respondentes que definem, em traços gerais, um profissional que utiliza bastante o mecanismo de busca do jornal onde atua. Considerando que o fator tempo está associado ao processo de pesquisa e exerce forte pressão sobre os jornalistas, 45,5% indicam realizar pesquisas com velocidade moderada (para entender temas e produzir textos de apoio) e 39,4% pesquisas rápidas (para checar dados). Além disso, a maior parte dos respondentes qualifica seu grau de experiência para uso do mecanismo de busca como usuário intermediário (39,4% realizam pesquisas mais elaboradas com alguns recursos avançados). Nesse contexto, ao utilizar o mecanismo de busca interna e deparar-se com uma situação de dificuldade para recuperação de informação, duas são as principais reações evidenciadas e indicadores de dificuldade, fuga e rejeição: substituição do mecanismo de busca interno por um mecanismo de busca global (87,9% usa o Google) e a necessidade de refazer várias vezes a estratégia de pesquisa (54,5%).

Tabela 9

PERCEPÇÃO COMO PESQUISADOR (Questões 21 a 23) %	
Se você procura um texto publicado no conteúdo on-line do jornal, mas o resultado de busca não recupera a informação. Qual sua reação mais frequente nessa situação?	
54,5	refaz várias vezes as estratégias de pesquisa no buscador do site jornalístico
87,9	usa um mecanismo de busca global (procura no Google)
15,2	procura a informação que precisa navegando pelo site (por exemplo, recuando datas ou consultando páginas especiais)
Em geral, como você qualifica sua velocidade de pesquisa?	
39,4	rápida (realiza pesquisas para checar dados)
45,5	moderada (realiza pesquisas para entender temas e produzir textos de apoio)
15,2	demorada (realiza pesquisa de estudo em profundidade ampliando conhecimento sobre algum tema)
Em geral, como você qualifica seu grau de utilização do mecanismo de busca do site jornalístico:	
27,3	básico (realiza pesquisas simples sem optar por filtros ou outros recursos avançados)
39,4	intermediário (realiza pesquisas mais elaboradas com alguns recursos avançados)
33,3	avançado (domina o uso dos recursos avançados disponíveis e elabora pesquisas com estratégias mais complexas)

Fonte: Elaboração do autor, 2015.

Por fim, apoiados nas indagações norteadoras da pesquisa, as duas questões finais (Tabela 10 e 11) indicam que:

É negativa a percepção do jornalista sobre a suficiência e adequação dos atributos de pesquisa disponíveis nos mecanismos de busca pesquisados; 57% indicam algum grau de discordância sobre a eficiência das estruturas e atributos de pesquisa do mecanismo de busca interno.

Tabela 10 – Questão 24

Na sua avaliação, os mecanismos de busca utilizados para recuperação de informações em jornais on-line possuem estruturas e atributos de pesquisa adequados e suficientes?

%	
0	Concordo
24	Concordo parcialmente
19	Nem concordo nem discordo
33	Discordo parcialmente
24	Discordo

Fonte: Elaboração do autor, 2015.

É também negativa a percepção de satisfação para uso do mecanismo de busca, 70% indicam algum grau de insatisfação no uso ferramenta de pesquisa em sua rotina de trabalho.

Tabela 11 – Questão 25

Em que medida os buscadores de sites jornalísticos satisfazem jornalistas e outros profissionais de comunicação em sua rotina de trabalho e sua dinâmica de pesquisa? A ferramenta de pesquisa facilita sua rotina de trabalho?

%	
18	nada satisfeito
52	pouco satisfeito
27	Satisfeito
3	muito satisfeito

Fonte: Elaboração do autor, 2015.

Para associar aos resultados obtidos uma perspectiva qualitativa e complementar as opiniões e atitudes dos respondentes, obtidas em questões fechadas, adotamos uma questão aberta proposta da seguinte forma:

Questão 26 - Comentário aberto: gostaria de acrescentar algo? indique sua principal dificuldade no uso de um mecanismo de busca de site jornalístico.

A partir dos comentários coletados, destacados abaixo, aspectos não explícitos nas questões fechadas confirmam a insatisfação no uso dos mecanismos de busca de sites jornalísticos, assim como a ineficiência dos recursos de pesquisa disponíveis e a incapacidade do sistema de recuperação oferecer uma busca ampla e integrada, com todos os recursos informacionais e áreas disponíveis no repositório. De forma espontânea, emergiram pontos comuns de consciência sobre o preocupante o volume de informação gerada (não recuperável) e o reconhecimento de que mecanismos de busca são relevantes interfaces de pesquisa na rotina profissional.

Para contornar a condição de ineficiência da ferramenta, que adquiriu reputação de inoperante, o Google é aceito como substituto imediato das interfaces de pesquisa dos sites jornalísticos e interpretada como uma ação rotineira; outro caminho apresentado é a utilização dos sistemas de gerenciamento de conteúdo do site (CMS) como recurso de pesquisa.

A ausência de *tags* nas matérias é mencionada como indicador da dificuldade para relacionamento de conteúdo, consequentemente, implica em falhas na indexação e desorganização de conteúdo.

Por fim, existe o reconhecimento que o controle algorítmico impõe aos usuários de mecanismos de busca restrições na liberdade de pesquisa, ou na definição de relevância, e não refletem com precisão as necessidades de informação compostas por formulações mentais e construídas em uma equação de busca. Além disso, uma interface como Google ou simplificadas, como as existentes nos sites jornalísticos, condicionam hábitos de pesquisa e estabelecem modelos mentais padrão que são adquiridos pelo uso.

Comentários obtidos (identificação de sexo, medida de uso do mecanismo de busca e tempo de experiência profissional):

1. *“Muitos sites, especialmente os sites de notícias locais, não têm mecanismos próprios de busca, mas usam o sistema do Google. Isso dificulta muito a checagem de informações e a utilização de determinado veículo como referência numa reportagem. Os veículos de busca dos sites dos grandes jornais, mesmo que próprios, são pouco personalizáveis, na minha opinião. Poucos dão opção para ordenar/filtrar as reportagens por data, ou relevância, ou editoria, ou autor. Ainda assim, por força de hábito, me acostumei a lidar com as poucas ferramentas que tenho em mãos. Hoje, esses mecanismos de busca são uma de minhas principais fontes para checagem de dados e preparação para reportagens e/ou entrevistas.”* [F, usa bastante o mecanismo, possui entre 6 a 10 anos de experiência profissional]

2. *“Dificuldade de encontrar palavras-chave; falta de tags marcando matérias de temática semelhante; ausência de links de suporte dentro das matérias, com uma lista de outros textos relacionados ao mesmo texto.”* [M, usa bastante o mecanismo, possui entre 1 a 5 anos de experiência profissional]

3. *“A maioria dos sites de busca têm algoritmos "viciados" (baseados nas pesquisas de outros usuários), que dificultam a pesquisa de informações muito específicas ou particulares (como mudança em lei, particularidades de procedimentos, e até cronologia dos fatos). No caso das buscas dos jornais, quando o acervo fica grande, fica cada vez mais difícil localizar esses dados.”* [M, usa bastante o mecanismo, mais de 20 anos de experiência profissional]

4. *“As buscas internas aos sites, geralmente, são inócuas. As buscas no Google, usando a pesquisa restrita ao site (do jornal, por exemplo) são mais proveitosas, sempre.”* [M, usa bastante o mecanismo, possui entre 16 a 20 anos de experiência profissional]

5. *“Com um grande número de formas de publicar o conteúdo do jornal (fac-símile da versão impressa, matérias dos sites, vídeos, áudios, páginas interativas, etc.) as vezes é necessário fazer buscas em mais de um local. Além disso, pelo menos na Folha, os vídeos não são indexados no sistema de busca do jornal. Achá-los é uma tarefa muito difícil. É mais fácil achar pelo google.”* [M, usa bastante o mecanismo, possui entre 1 a 5 anos de experiência profissional]

6. *“Como utilizo programa interno de gerenciamento de conteúdo, as buscas diretas no site do jornal são mais pontuais, para conteúdos que eu tenha lido e queira recuperar no ambiente do site.”* [F, usa mais ou menos o mecanismo, possui entre 16 a 20 anos de experiência profissional]

7. *“O melhor mecanismo de busca é o Google, sempre minha primeira opção. Lá, pesquisando com as mesmas palavras usadas na busca dentro do site jornalístico, você acha matérias que você não acha no buscador interno do jornal.”* [F, usa mais ou menos o mecanismo, possui entre 1 a 5 anos de experiência profissional]

8. *“O sistema de busca da Folha é ruim, incompleto e não recupera todo o conteúdo de um determinado tema. No do Estado, a busca é mais eficiente pelo nome do autor/ repórter que fez a matéria.”* [F, usa bastante o mecanismo, possui entre 16 a 20 anos de experiência profissional]

9. *“A busca dentro de um site jornalístico retorna com uma certa poluição - eu não preciso ver fotos, trechos dos textos em letras grandes, etc. Partindo do princípio que em uma busca na internet eu desejo relevância aliado à agilidade, listar o título e um pequeno trecho, da mesma forma que o Google faz, tornaria a pesquisa mais atraente para mim. Pode ser também uma questão de costume - veja, usamos o Google há quantos anos? Nos habituamos a tê-lo como exemplo de caso bem sucedido. Assim, se não nos for apresentado algo que seja melhor do que o que a gente já usa, não vejo motivos para trocar meu sistema de busca. É mais fácil você achar algo dentro do site do Estadão pelo Google do que pelo buscador do Estadão...”* [M, usa pouco o mecanismo, possui entre 16 a 20 anos de experiência profissional]

10. *“Eles ainda precisam melhorar MUITO. Mas para isso os CMSs também precisam melhorar MUITO para dar conta dos diversos tipos de informação que podem ser úteis numa busca.”* [M, usa mais ou menos o mecanismo, possui entre 11 a 15 anos de experiência profissional]

11. *“Os mecanismos de busca dos principais jornais brasileiros são muito ruins. O Google é bem mais eficiente. Costumo usar o mecanismo do próprio jornal apenas por desengano de consciência, mas é sempre frustrante.”* [M, usa mais ou menos o mecanismo, possui entre 16 a 20 anos de experiência profissional]

OBS.: Para atender a solicitação dos respondentes que manifestaram interesse em acompanhar os resultados da pesquisa, indicando contato por email no final do questionário, uma versão reduzida e interativa dos resultados foi desenvolvida no formato de infográfico e está disponível no endereço:

https://infogr.am/recuperacao_de_informacao_em_jornais_on_line

6. CONSIDERAÇÕES FINAIS

Empresas de informação, que geram e utilizam documentos em ambientes on-line, começam a formar grandes repositórios de informação híbrida, de estrutura complexa e conteúdo com dinâmica associativa e modelado para atingir redes sociais.

O documento digital adquiriu nova forma (formatos) e é mais denso, orgânico, com simultâneas e múltiplas conexões que, conseqüentemente, implicam em maior complexidade no processo de busca por informações e ampliam o papel que os atuais sistemas de recuperação de informação, desenvolvidos para este meio Web, adquirem ao intermediarem a memória de seus bancos de dados e a memória da sociedade.

Além disso, tecnologias de informação e comunicação, geradas pela internet, trouxeram novas perspectivas de ordem digital, dimensões físicas e temporais (espaço e tempo) no tratamento da informação e transformações paradigmáticas baseadas na volatilidade de dados, velocidade de ações e interatividade de pessoas e ideias.

Os resultados iniciais desta pesquisa indicam que os modelos de mecanismo de busca selecionados não operam com padrões comuns e absorvem as mudanças provocadas pela tecnologia disponível e também pela lógica comercial de empresas na Web. O volume de informação gerado pelo meio é outro fator que implica em uma série de indagações: Toda a informação contida no repositório de textos produzidos é recuperável? O mecanismo de busca (interno) dos sites selecionados oferece soluções ideais para pesquisa? Como recuperar informações de áreas restritas (páginas especiais, blogs de comentaristas, infográficos interativos, serviços de vídeo e podcast)? Como facilitar a recuperação de informações perdidas em sistemas desestruturados? Quais linguagens de recuperação podem ser aplicadas? Que estruturas de representação deste conhecimento precisam ser padronizadas ou desenvolvidas?

Diante deste panorama, percebe-se que a necessidade de ampliar estudos em tecnologia de pesquisa e acesso à informação, aplicados na avaliação do tráfego de informação jornalística em rede, exige do profissional em Ciência da Informação uma maior compreensão sobre a mudança da modelagem conceitual e do ambiente digital.

Para isso, novas abordagens precisam fornecer reflexões sobre parâmetros fundamentais na avaliação de sistemas de informação em hipermídia, linguagem documentária de hipertextos, bem como análises do comportamento de pesquisa dos usuários on-line e indicadores de limitações ou modificações das práticas utilizadas na organização e recuperação de informação digital.

No caso deste estudo, as reflexões estão norteadas por duas indagações, a primeira diz respeito à estrutura dos atributos de pesquisa e a segunda sobre a satisfação de uso da interface de busca, ambas as questões focadas na recuperação de informação e na relação informação-mecanismo de busca-usuário. A partir desse caminho de investigação, avançamos na busca dos parâmetros mencionados e, agora, verificamos que os resultados obtidos demonstram que:

1) Baseados na estrutura da informação jornalística, atributos de pesquisa dos mecanismos de busca de jornais on-line apresentam recursos não padronizados, reduzidos e pouco explorados. Em conflito com a dinâmica de evolução tecnológica aplicada no ambiente Web, construídos em narrativa de hipertexto e hipermídias, buscadores de notícias possuem capacidade limitada para filtrar informação relevante, ou mesmo criar relações semânticas por meio de conexão entre todas as unidades de informação (texto, imagem, vídeo, infográfico etc.). Esta condição implica em desagregação do conteúdo gerado em crescimento constante e, conseqüentemente, invisibilidade, ruído na recuperação e uso desarticulado da linguagem jornalística. Além de impedimentos para inovação na representação e organização de conhecimento, por exemplo, uso de ontologias, ou construção de modelos elaborados para visualização dos resultados de pesquisa.

2) Ocorre insatisfação do profissional jornalista ao utilizar a ferramenta de busca em sua rotina de pesquisa, derivada de fatores estruturais que afetam a severidade do problema de busca, frequência de usuários que reconhecem comprometimento da pesquisa, indicadores de impacto (tempo gasto) e persistência (dificuldade contínua). Em relação direta com a geração de atributos de pesquisa, o grau de satisfação apresentado justifica investimentos no processo de recuperação de informação, uso avançado e consistente de padrões de metadados, assim como maior conexão entre o

ciclo de produção de notícias, aplicação de meta tags, interface de busca, necessidades do informacionais do usuário, visualização e resultados da pesquisa.

Como proposta final deste estudo, em sintonia com a necessidade do usuário profissional, apresentamos algumas percepções com indicações de categorias e funções de análise que podem contribuir para boas práticas de otimização da interface de busca e recomendações de estruturação da informação jornalística.

Considerando o panorama revelado durante a pesquisa, parece urgente pensar um modelo de análise que implique na identificação da capacidade informativa de interfaces de pesquisa. No caso dos mecanismos de busca interna dos sites jornalísticos, esta capacidade informativa implica em relacionar o ciclo de produção da informação, a estrutura informativa da notícia, a organização do conteúdo gerado e os recursos de pesquisa para recuperação da informação. Desta forma, a interface mediadora, que permite acesso às informações, precisa identificar níveis de “informatividade” do dispositivo: insuficiência de dados e recursos de pesquisa, técnicas de identificação e arranjo da informação, problemas de gênero e contexto temático, sempre considerando em que medida os recursos de pesquisa são (ou não) esperados, conhecidos e visualizados pelo usuário. Para isso, alguns pontos exigem observação:

- A dinâmica do sistema de recuperação de informação, considerando um ambiente com conteúdo renovado constantemente, a aplicação de um mecanismo de busca avançada deve considerar a análise da evolução estrutural de um documento digital, bem como sua utilização em outras plataformas de disseminação adotadas pela empresa de informação;
- A evolução tecnológica e o reconhecimento de tendências que configuram novas gerações do jornalismo digital;
- A diversidade temática do repositório, acompanhando o aparecimento de novos campos semânticos decorrentes de mudanças editoriais, bem como a exclusão de temas, esta análise observa redefinições da relação entre assuntos e adequações no tratamento da informação;

- A conectividade informativa que, apoiada na indexação definida por metadados/tags, visa avaliar a relação temática entre unidades informativas (por exemplo, um dossiê digital) e todos os recursos informacionais que o integram (texto, vídeo, áudio) e estão presentes no repositório;
- A visualização da informação recuperada, visando análise da visibilidade do conteúdo como recurso de pesquisa e associado ao processo de recuperação de informação com foco na linguagem, aspectos semânticos e correspondência de conteúdos;

Observados estes pontos, propomos também algumas funções operativas que, acreditamos, são necessárias para avaliar a capacidade informativa dos mecanismos de busca e checar sua eficiência como recurso de pesquisa:

- a) Relacionar unidades informativas (documento) e recursos informacionais (texto, áudio, vídeo, infográfico etc.);
- b) Integrar processo de produção, organização e gestão de conteúdo;
- c) Integrar novos padrões de identificação dos recursos informacionais;
- d) Permitir a condição de memória considerando preservação, armazenamento, recuperação e visualização de informação;
- e) Estabelecer padrões de diálogo entre usuário e o sistema de recuperação por meio de tradução da linguagem natural para linguagem de pesquisa;
- f) Permitir novas configurações de visualização do conteúdo recuperado, visando produzir padrões comparativos mais eficientes para avaliação de equações de busca e resultados.

Por fim, no âmbito da Ciência da Informação, acreditamos que a linguagem documentária oferece caminhos em duas dimensões essenciais para organização do conhecimento e validação das funções operativas propostas:

Dimensão tecnológica - que estabelece recomendações e padrões de linguagem computacional que permitem interação entre descrição, representação do documento e funções de recuperação do conteúdo adequado exibido nos resultados. Para esta pesquisa em particular, destacamos como referencia a IPTC NewsCode que propõe um vocabulário controlado ou uma taxonomia para metadados de notícias. Considerando diversos formatos (texto, fotografias, gráficos, áudio, vídeos etc), a codificação NewsCode mantém um conjunto de conceitos que permite uma aplicação consistente de metadados de produtos jornalísticos e, por ser formado de códigos com definição explícita e semântica para compartilhamento entre usuários, pode ser utilizado em diferentes línguas. Não obstante a esta condição, acreditamos ser importante o desenvolvimento de estudos nacionais que analisem sua aplicação e definam melhores padrões para jornais brasileiros.

Dimensão humana – que, destacada pelos profissionais da Ciência da Informação, utiliza a linguagem de um determinado domínio para gerar coerente organização, relacionamento e recuperação de informação, por consequência, amplia a memória e o compartilhamento de conteúdo em sociedade. Neste sentido, como recurso de síntese contextual e captação de conhecimento, a linguagem documentária oferece um canal de comunicação entre as unidades de informação que compõem um documento digital e o usuário que, por sua vez, busca sintonia com seu pensamento e recorre ao discurso contido no documento digital, resultado da produção de determinada comunidade discursiva e apoiado em modelos sociais, culturais e científicos.

REFERÊNCIAS

ALMEIDA, Daniela dos Reis de et al. Paradigmas contemporâneos da ciência da informação: a recuperação da informação como ponto focal. **Revista Eletrônica Informação e Cognição**, v.6, n.1, p. 16-27, 2007.

ALVES, Rahel Cristina Vezú. **Web semântica: uma análise focada no uso de metadados**. Marília, 2010. 132 f. Tese (Doutorado) – Universidade Estadual Paulista, Brasil.

ALZAMORA, Geane. A semiose da informação webjornalística. In: BRASIL, André; FALCI, Carlos Henrique; JESUS, Eduardo de; ALZAMORA, Geane. **Cultura em fluxo: novas mediações em rede**. Belo Horizonte: PUC Minas, 2004. p.110-125.

ANDRADE, Antonio Luis Lordelo. **Usabilidade de interfaces web: avaliação heurística no jornalismo on-line**. Rio de Janeiro : E-papers, 2007.

ARAÚJO JÚNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 9241-11: requisitos ergonômicos para trabalho de escritórios com computadores: parte 11 - orientações sobre usabilidade**. Rio de Janeiro: ABNT, 2002.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York : ACM Press, 1999.

BARBOSA, Suzana. **Jornalismo digital em base de dados (JDBD): um paradigma para produtos jornalísticos digitais**. Bahia, 2007. 329f. Tese (Doutorado) – Universidade Federal da Bahia. Faculdade de Comunicação.

BERTOCCHI, Daniela. **Dos dados aos formatos**: um modelo teórico para o design do sistema narrativo no jornalismo digital. São Paulo, 2013. 250p. Tese (Doutorado) – Escola de Comunicações e Artes. Universidade de São Paulo, Brasil.

BUCKLAND, M. **Information as thing**. Journal of the American Society for Information Science, v.42, n.5, p. 351-360, 1991.

CALDERÍN, Mabel; ROJANO, Miladys. (2006). El periodista y la documentación digital: nuevo espacio para la investigación. In: ROJANO, Miladys (coord.), **Diez años del periodismo digital en Venezuela**. Caracas: Publicaciones UCAB, 2006. p. 273-292

CAPURRO, Rafael. Epistemologia e ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 5., 2003, Belo Horizonte. **Anais...** Belo Horizonte: ECI/UFMG, 2003. Disponível em: <http://www.capurro.de/enancib_p.htm> Acesso em: 14 abr. 2013

CAPURRO, Rafael; HJØRLAND, Birger. O conceito de informação. **Perspectivas em Ciência da Informação**, v.12, n.1, p. 148-207, jan./abr. 2007. Disponível em: <<http://www.eci.ufmg.br/pcionline/index.php/pci/article/view/54/47>> Acesso em: 15 jan 2013.

CARVALHO, Cleide. Jornalismo online: audiência maior e mais voz ao leitor. **Globo Online**. 13 nov. 2007. Disponível em: <<http://oglobo.globo.com/tecnologia/mat/2007/11/13/327156752.asp>>. Acesso em: 21 out. 2013.

CASTELLS, Manuel. **A sociedade em rede**. São Paulo : Paz e Terra, 1999. (A era da informação: economia, sociedade e cultura, v.1)

CENDÓN, Beatriz Valadares. Ferramentas de busca na Web. **Ciência da Informação**, Brasília, v.30, n.1, jan./abr. 2001.

CRUCIANELLI, Sandra. **Herramientas digitales para periodistas**. Centro Knight para el Periodismo en las Américas de la Universidad de Texas, 2013.

DERTOUZOS, Michael. **O que será** : como o novo mundo da informação transformará nossas vidas. São Paulo : Companhia das Letras, 1997.

DI FELICE, Massimo; TORRES, Julliana Cutolo; YANAZE, Leandro Key Higuchi. **Redes digitais de sustentabilidade**: as interações com o meio ambiente na era da informação. São Paulo: Annablume, 2012.

DIAS, Cláudia. **Usabilidade na web** : criando portais mais acessíveis. Rio de Janeiro : Alta Books, 2003.

DÍAZ NOCI, Javier; SALAVERRÍA ALIAGA, Ramón. (coords.) **Manual de redacción ciberperiodística**. Barcelona: Editorial Ariel, 2003. p.425-524.

FERNEDA, Edberto. **Recuperação de informação** : análise sobre a contribuição da ciência da computação para a ciência da informação. São Paulo, 2003. 137p. Tese (Doutorado) – Escola de Comunicações e Artes. Universidade de São Paulo, Brasil.

FERREIRA, Jaider Andrade; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade**: Estudos, João Pessoa, v. 23, n. 1, 2013. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/15436>> Acesso em: 05 jun. 2015.

FIGARO, Cláudia (coord.). **O perfil do jornalista e os discursos sobre o jornalismo**: um estudo das mudanças no mundo do trabalho do jornalista profissional em São Paulo [Relatório Final da Pesquisa]. São Paulo: Centro de Pesquisa em Comunicação e Trabalho; ECA/USP, 2012. Disponível em: <http://www.eca.usp.br/comunicacaoetrabalho/wp/wp-content/uploads/relatorio_final_2012.pdf> Acesso em: 20 ago. 2013.

FOLHA DE S.PAULO. A Folha em números: um instantâneo, em cifras, do cotidiano do maior jornal do país. São Paulo: **Folha de S.Paulo**, 19 fev. 2011. Suplemento: Folha 90 anos. p.44. Disponível em: <<http://acervo.folha.com.br/fsp/2011/02/19/556//556544>> Acesso em 20 ago. 2014.

FOLHA ONLINE. Folha Online estréia novo projeto gráfico; veja o que mudou no site. **Folha Online**. 20 ago. 2007. Disponível em: <<http://www1.folha.uol.com.br/folha/brasil/ult96u321402.shtml>>. Acesso em: 20 out. 2013.

GARCIA DE LEÓN, Alicia; CALDERA SERRANO, Jorge. Buscadores internos: una aproximación a los criterios de calidad [on line]. **Hipertext.net**, n. 5, 2007. Disponível em: <<http://www.upf.edu/hipertextnet/numero-5/buscadores-internos.html>> Acesso em: 20 out. 2013.

GARCÍA GUTIÉRREZ, A. Estructura, documentación y lenguaje. In: _____. **Estructura lingüística de la documentación: teoría y método**. Murcia: Universidad, Secretariado de Publicaciones, 1990. p.17-47.

GARCÍA SANTIAGO, Lola. **Extraer y visualizar información en internet: el web mining**. Gijón: Trea, 2003.

GLOBO, O. A informação na hora em sintonia com o leitor: jornal renova a abordagem da notícia e aposta na agilidade das plataformas digitais. **O Globo**, Rio de Janeiro, 06 abr. 2014. Caderno País, p.10

GOOGLE. **A newly hatched way to tag your news articles**. 19 set. 2012. Disponível em: <<http://googlenewsblog.blogspot.com.br/2012/09/a-newly-hatched-way-to-tag-your-news.html>> Acesso em: 04 ago. 2014.

GOUVÊA, Cleber; LOH, Stanley. Jornalismo semântico: uma visão em direção do futuro do jornalismo online. In: LONGHI, Raquel; D'ANDREA, Carlos (orgs.). **Jornalismo convergente: reflexões, apropriações, experiências**. Florianópolis: Insular, 2012. p.73-86.

HARVEST DIGITAL. **User attitudes to search**: How do experienced Internet users feel about search engines? London: Harvest digital, Metro Research, 2006.

INGWERSEN, Peter; JÄRVELIN, Kalervo . **The turn** : integration of information seeking and retrieval in context . Dordrecht : Springer, 2005.

IPROSPECT. **Search Engine User Attitudes Survey**. [U.S.]: iProspect, 2004.

IPTC. **NewsCodes: Controlled Vocabularies for the Media**. 2015a. Disponível em: <<https://iptc.org/standards/newscodes/>> Acesso em: 12. Abr. 2015.

IPTC. **rNews: Embedded Metadata For The News Industry**. 2015b. Disponível em: <<http://dev.iptc.org/rNews-1-Introduction-to-rNews>> Acesso em: 12 abr. 2015.

JOHNSON, Steven. **Cultura da interface** : como o computador transforma nossa maneira de criar e comunicar. Rio de Janeiro, Jorge Zahar Editora, 2001.

KAPUSCINSKI, Ryszard. **Los cinco sentidos del periodista**: estar, ver, oír, compartir, pensar. México : Fundación Nuevo Periodismo Iberoamericano; Fundación PROA; FCE, 2003.

KUHLTHAU, Carol Collier. **Seeking meaning** : a process approach to library and information services. Westport : Libraries Unlimited, 2004.

LACHOWSKI, Gibran Luis; GOMES, Iuri Barbosa. A prática jornalística na era da convergência tecnológica: reflexões sobre os fundamentos do jornalismo aplicados no âmbito digital. **Revista Funec Científica – Multidisciplinar**, Santa Fé do Sul (SP), v. 2, n. 3, jul./dez. 2012.

LARA, M. L. G. Linguagem documentária e terminologia. **Transinformação**, Campinas, v.16, n.3, 2004.

LARA, M. L. G. & CONTI, V. Disseminação da informação e usuários. **São Paulo em Perspectiva**, Fundação Seade, v.17, n.3/4, 2003.

LEITE, Edmundo. **11 de setembro e amnésia digital**. São Paulo, 11 set. 2011. Disponível em <<http://brasil.estadao.com.br/blogs/edmundo-leite/11-de-setembro-e-amnesia-digital/>> Acesso em 18 maio 2015.

LÉVY, Pierre. **A inteligência coletiva**: por uma antropologia do ciberespaço. São Paulo, Loyola, 1999.

LÉVY, Pierre. **As tecnologias da inteligência** : o futuro do pensamento na era da informática. Rio de Janeiro : Ed. 34, 2002.

LIMA, Gercina Ângela Borém. Organização da informação para sistemas de hipertextos. In: NAVES, Madalena Martins Lopes; KURAMOTO, Hélio. (orgs.) **Organização da informação**: princípios e tendências. Brasília, DF: Briquet de Lemos, 2006. p.99-116.

MACHADO, Elias. **O ciberespaço como fonte para jornalistas**. Salvador, BA : Calandra, 2003. (Coleção Biblioteca J)

_____. **O jornalismo digital em base de dados**. Florianópolis : Calandra, 2006. (Coleção Biblioteca J)

MACHADO, Elias. BORGES, Clarissa. MIRANDA, Milena. Modelos de produção de conteúdo no jornalismo digital baiano. In: MACHADO, Elias. PALACIOS, Marcos. (orgs.) **Modelos de jornalismo digital**. Salvador : Edições GJOL, Calandra, 2003.

MACKINNON, Rebecca. **The World Wide Conversation**: Online Participatory Media and International News. The Joan Shorenstein Center on Press, Politics and Public Policy, Working Papers Series. n. 2, 2004.

MACKINLEY, Tony. **Do papel até web** : como tornar as informações acessíveis instantaneamente. São Paulo : Quark Book, 1998.

MANNARINO, Marcus Vinicius Rodrigues. **O papel do web jornal** : veículo de comunicação e sistema de informação. Porto Alegre : EDIPUCRS, 2000. (Coleção Comunicação, 5)

MANZANO, José Augusto N.G.; OLIVEIRA, Jayr Figueiredo de. **Algoritmos** : lógica para desenvolvimento de programação de computadores. São Paulo : Érica, 2008

MARCHIORI, Patrícia Zeni. Informação e conhecimento nos processos de tagging (marcação) de conteúdo colaborativo. In: TOMAÉL, Maria Inês (org.) **Compartilhamento da informação**. Londrina: Eduel, 2012. p.41-71.

MARTINEZ, M.L. **Um método de webdesign baseado em usabilidade**. Tese (Doutorado) - Escola Politécnica. Universidade de São Paulo, Brasil. São Paulo, 2002. 301p.

MARTÍNEZ MÉDEZ, Javier Francisco. **Propuesta y desarrollo de um modelo para la evaluación de la recuperación de información em Internet**. Murcia, 2002. 283p. Tese (Doutorado) – Facultad de Ciencias de la Documentación. Universidad de Murcia, Espanha.

MATEOS SÁNCHEZ, Montserrat; GARCÍA-FIGUEROLA PANIAGUA, Carlos. **Aplicación de técnicas de clustering en la recuperación de información web**. Gijón: Trea, 2009.

MENA, Fernanda; PLIGER, Marcelo. Siga o caminho das notícias na era das multiplataformas: o fluxo da informação na Folha, da captação de notícias à publicação do jornal impresso e nos vários meios digitais. São Paulo: **Folha de S.Paulo**, 19 fev. 2011. Suplemento Comemorativo: Folha 90 anos. p.38. Disponível em: <<http://acervo.folha.com.br/fsp/2011/02/19/556//5565434>> Acesso em 20 ago. 2014.

MÉNDEZ RODRÍGUEZ, Eva. **Metadados y recuperación de información**: estándares, problemas y aplicabilidad en bibliotecas digitales. Gijón: Trea, 2002.

MEYER, Philip. **Periodismo de precisión** : nuevas fronteras para la investigación periodística. Barcelona: Boch, 1993.

MICROFORMAT.ORG. **About microformats**. 19 jun.2015 (última versão). Disponível em <<http://microformats.org/wiki/about>> Acesso em: 20 jun.2015.

MICROFORMAT.ORG. **Microdata**. 16 out. 2014 (última versão). Disponível em <<http://microformats.org/wiki/microdata>> Acesso em: 20 jun. 2015.

MOREIRO GONZÁLEZ, José Antonio. **Linguagens documentárias e vocabulários semânticos para web**: elementos conceituais. Salvador: EDUFBA, 2011.

MIZUKAMI, Pedro; REIA, Jhessica; VARON, Joana. **Mapeamento da mídia digital no Brasil**: um relatório da Open Society Foundations. Rio de Janeiro : Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas : Centro de Tecnologia e Sociedade, 2014. 340p. Disponível em: <<https://monitorando.files.wordpress.com/2014/09/mapeamento-da-mc3addia-digital-no-brasil.pdf>> Acesso em: 10 jan. 2015.

MONTEIRO, Silvana Drumond. Os mecanismos de busca : à guisa de uma tipologia das múltiplas sintaxes. In: TOMAÉL, Maria Inês. (org.) **Fontes de informação na Internet**. Londrina : Eduel, 2008.

NEWMAN, Nic; LEVY, David A. L (orgs.) **Reuters Intitute Digital News Report 2014: Tracking the Future of News**. Reuters Institute for the Study of Journalism, University of Oxford, 2014. Disponível em: <<http://bit.ly/1s6j2W7>> Acesso em: 01 jul. 2014.

NEWSPAPER ASSOCIATION OF AMERICA. **Newspaper web sites attract record audiences in first quarter**. NAA. Disponível em: <<http://www.naa.org/PressCenter/SearchPressReleases/2008/NEWSPAPER-WEB-SITES-ATTRACT-RECORD-AUDIENCES-IN-FIRST-QUARTER.aspx>>. Acesso em: 15 set. 2013.

_____. **Newspaper web audiences**. NAA. Disponível em: <<http://www.naa.org/~media/NAACorp/Public%20Files/TopicsAndTools/Digital/Mar2015DigitalAudience.ashx>>. Acesso em: 10 jun. 2015.

NIELSEN, Jacob . Ten usability heuristics. In: NIELSEN,J.; MACK,R. (eds.) **Usability inspection methods**. New York: John Wiley & Sons, 1994. Disponível em: <http://www.useit.com/papers/heuristic/heuristic_list.htm> Acesso em: 01 out 2013.

NIELSEN, Jakob; LORANGER, Hoa. **Usabilidade na web** : projetando websites com qualidade. Rio de Janeiro : Elsevier, 2007

OPEN GRAPH PROTOCOL. 20 out. 2014 (última versão). Disponível em: <<http://ogp.me/>> Acesso em 20 mar. 2015.

PALACIOS, Marcos. A memória como critério de aferição de qualidade no ciberjornalismo: alguns apontamentos. **Revista FAMECOS**, n.37, p. 91-100, dez. 2008.

_____. Convergência e memória: jornalismo, contexto e história. **Revista Matrizes**, v. 4, n. 1, p.123–144, jul./dez. 2010 Disponível em: <<http://www.matrizes.usp.br/index.php/matrizes/article/view/173/290>> Acesso em 08 jan. 2014.

PARISER, Eli. **O filtro invisível**: o que a internet está escondendo de você. Rio de Janeiro : Zahar, 2012.

PERLMAN, Gary. **User Interface Usability Evaluation with Web-Based Questionnaires**. 18 maio 2015 (última versão). Disponível em: <<http://garyperلمان.com/quest/>> Acesso em: 03 jun. 2015.

PÉREZ GUTIÉRREZ, Mario. El lenguaje de interrogación: una gramática formal para recuperación de información. Revista **Española de Documentación Científica**, v. 23, n. 3, p. 246-266, 2000.

PEW RESEARCH CENTER. **Key News Audiences Now Blend Online and Traditional Sources**: Audience Segments in a Changing News Environment: Pew Research Center Biennial News Consumption Survey. Washington, DC: Pew Research Center, 2008.

PINTO, Ana Estela de Sousa. **Jornalismo diário**: reflexões, recomendações, dicas e exercícios. São Paulo : Publifolha, 2009.

POLISTCHUK, Ilana; TRINTA, Aluizio Ramos. **Teorias da comunicação** : o pensamento e a prática da comunicação social. Rio de Janeiro : Elsevier, 2003.

QUADROS, Claudia; CAETANO, Kati; LARANJEIRA, Álvaro (orgs.). **Jornalismo e convergência**: ensino e práticas profissionais. Covilhã: LabCom Books, 2011

ROSENFELD, Louis; MORVILLE, Peter. **Information architecture for the word wide web**. California : O'Reilly, 1998.

SAAD, Beth. **Estratégias para a mídia digital** : internet, informação e comunicação. São Paulo : SENAC, 2003.

SAAD, Beth; BERTOCCHI, Daniela. A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria. **Revista Matrizes**, v. 5, n. 2, p.123-144, jan./jun. 2012 Disponível em: <<http://www.matrizes.usp.br/index.php/matrizes/article/view/340>> Acesso em 01 jan. 2014.

SAN SEGUNDO, Rosa. Panorama da investigación en organización de conocimiento em su dimensión epistemológica. In: DODEBEI, Vera; GUIMARÃES, José Augusto Chaves (orgs.) **Complexidade e organização do conhecimento: desafios de nosso século**. [recurso eletrônico] Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013.

SARACEVIC, Tefko. Relevance: A Review of Literature and a Framework for Thinking on the Notion in Information Science. Part II: Nature and Manifestation of Relevance. **Journal of the American Society for Information Science and Technology**, v.58, n.13, p. 1915-1933, 2007.

SCHEMA. **Extended schema.org news support**. 20. set. 2011. Disponível em: <<http://blog.schema.org/2011/09/extended-schemaorg-news-support.html>> Acesso em: 18 nov. 2014.

SEGARAN, Toby. **Programando a inteligência coletiva** : desenvolvendo aplicativos Web 2.0 inteligentes. Rio de Janeiro : Alta Books, 2008.

SEIXAS, Lia. Gêneros jornalísticos digitais: uma proposta de critérios para definir os produtos do webjornalismo. In: DÍAZ NOCI, Javier. PALACIOS, Marcos. (orgs.) **Metodologia para estudo dos cibermeios**: estado da arte e perspectivas. Salvador : EDUFBA, 2008. p.195-214.

_____. Metodologia de análise de cibergêneros no jornalismo brasileiro. In: MACHADO, Elias. PALACIOS, Marcos. (orgs.) **Modelos de jornalismo digital**. Salvador : Edições GJOL, Calandra, 2003.

_____. **Redefinindo os gêneros jornalísticos**: propostas de novos critérios de classificação. Covilhã: LabCom, 2009.

SILVA, Armando Malheiro da. **A informação**: da compreensão do fenômeno e construção do objecto científico. Porto: Edições Afrontamento, 2006.

SOUTO, Leonardo Fernandes. **Informação seletiva, mediação e tecnologia**: a evolução dos serviços de disseminação seletiva das informações. Rio de Janeiro: Interciência, 2010.

SMIT, Johanna W.; TÁLAMO, M. F. G. M. Sistemas de Recuperação de Informação e Memória. In: IIV ENANCIB, 2006, Marília. **VII ENANCIB: A Dimensão Epistemológica da Ciência da Informação e suas Interfaces Técnicas, Políticas e Institucionais**, 2006.

SMITH, Gene. **Tagging**: people-powered metadata for the social web. Berkeley, CA: New Riders, 2008.

SPORNY, Manu. **An Uber-comparison of RDFa, Microdata and Microformats**. The Beautiful, Tormented Machine [blog] 25 jun. 2011. Disponível em: <<http://manu.sporny.org/2011/uber-comparison-rdfa-md-uf/>> Acesso em 12 jun. 2015.

THUROW, Shari; MUSICA, Nick. **When search meets web usability**. Berkeley, EUA : New Riders, 2009.

TOMAÉL, Maria Inês; ALCARÁ, Adriana Rosecler; SILVA, Terezinha Elisabeth da. Fontes de informação na Internet: critérios de qualidade. In: TOMAÉL, Maria Inês. (org.) **Fontes de informação na Internet**. Londrina : Eduel, 2008.

VICENS ARASANZ, Toni. **Normas básicas de usabilidad para buscadores internos**. [online] 21 jul.2008. Disponível em: <<http://elfosdesign.blogspot.com.br/2008/07/normas-bsicas-de-usabilidad-para.html>> Acesso em: 20 jan. 2013

WILLNAT, Lars; WEAVER, David H.. **The American Journalist in the Digital Age**. Bloomington, USA: School of Journalism, Indiana University, 2014. Disponível em: <<http://news.indiana.edu/releases/iu/2014/05/2013-american-journalist-key-findings.pdf>> Acesso em: 01 jul. 2014.

W3C. **RDF**. 15 mar.2014 (última versão). Disponível em: <<http://www.w3.org/RDF/>>
Acesso em: 12 jun. 2015.

W3C. **RDFa 1.1 Primer - Third Edition Rich Structured Data Markup for Web Documents**. 17 mar. 2015 (última versão). Disponível em: <<http://www.w3.org/TR/xhtml-rdfa-primer/>> Acesso em: 12. jun. 2015.

APENDICE A – QUESTIONÁRIO ON-LINE

Telas capturadas da versão on-line do formulário.

RECUPERAÇÃO DE INFORMAÇÃO EM JORNAIS ON-LINE

Caro profissional de comunicação,

Este questionário integra estudo acadêmico, apoiado pelo Programa de Ciência da Informação da Universidade de São Paulo, que analisa questões de organização e recuperação de informação em jornais on-line. Observamos que a convergência de plataformas, o crescimento constante da produção exclusivamente on-line e a dinâmica de atuação em ambientes digitais têm causado profundo impacto na maneira como jornalistas realizam seu trabalho. Considerando este panorama, investigamos potenciais dificuldades de pesquisa no uso de mecanismos de busca de sites jornalísticos brasileiros. Vale ressaltar que analisamos apenas o uso de buscadores internos dos sites: Estadão, Folha e O Globo. *Não avaliamos mecanismos de busca global, como Google.

Acreditamos que a importância da pesquisa encontra-se na sua percepção sobre esta ferramenta e aspectos do processo de busca por informações em sua rotina de trabalho, por isso, convidamos você para responder as questões abaixo. A coleta de dados é anônima e o tempo médio de resposta é 8 minutos.

Obrigado pelo apoio e participação na pesquisa!

***Obrigatório**

PERFIL PROFISSIONAL

Sexo *

☐ masculino

☐ feminino

Idade *

☐ 18 a 24

☐ 25 a 34

☐ 35 a 44

☐ 45 a 54

☐ acima de 55

Área de atuação na atividade jornalística *

- ☐ on-line
☐ impresso
☐ on-line e impresso

Veículo de comunicação onde atua: *

- ☐ Estadão
☐ Folha
☐ O Globo

Tempo de experiência profissional *

- ☐ 1 a 5 anos
☐ 6 a 10 anos
☐ 11 a 15 anos
☐ 16 a 20 anos
☐ mais de 20 anos

Em que medida utiliza o mecanismo de busca do jornal on-line onde você trabalha? *

- ☐ uso bastante
☐ mais ou menos
☐ uso pouco

USO DO MECANISMO DE BUSCA EM PESQUISAS DIÁRIAS

CONSIDERANDO SUA ROTINA DE TRABALHO, COMO CLASSIFICA O USO DO MECANISMO DE BUSCA INTERNO DO SITE JORNALÍSTICO PARA REALIZAR AS SEGUINTE TAREFAS:

Checar informações (números, grafias, datas) *

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

Coletar informações de apoio para produção de uma pauta *

(entrevistas, perfis, textos de memória, matérias retrospectivas)

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

Recuperar diferentes formatos de informação *

(mapa, ilustração, infográfico interativo, tabela, gráfico, vídeo, áudio e outros formatos anexados ao texto)

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

Analisar uma ideia de pauta *

(verificar o ineditismo da proposta, qual a cobertura já realizada ou ângulos de abordagem já publicados)

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

Relacionar conteúdo *

(associar matérias recentes aos links de textos históricos ou de apoio, que ajudem ampliar e explicar o tema abordado; construção de links no formato "saiba mais" ou "leia também")

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

Aplicar e controlar tags *

(pesquisar tags/palavras-chave para identificação de novos textos, pensar palavras-chave que evitem ruído de informação e promovam relacionamento com outros textos)

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ uso sempre nesta tarefa

CONTEÚDO E CATEGORIAS DE PESQUISA

CONSIDERANDO SUA BUSCA POR CONTEÚDO, COMO VOCÊ AVALIA AS SEGUINTE CATEGORIAS DE PESQUISA EM SUA ROTINA:

Pessoas *

busca por conteúdo biográfico, fontes ou referências de especialistas sobre determinado tema

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

Memória *

busca por conteúdo retrospectivo para recordar um fato, realizar relacionamento histórico, esclarecer ou acompanhar eventos

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

Legislação *

busca por referenciais legais sobre determinado tema apoiado em leis, decretos ou códigos

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

Números *

busca por estatísticas ou indicadores sobre determinado tema

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

Geografia *

busca de informações sobre determinado local: bairro, cidade, país ou região

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

Significados e Explicações *

busca de informações que esclareçam ou expliquem determinado conceito, teoria, tema, caso, assim como técnica, uso ou funcionamento de algo

1 2 3 4 5

não uso ☐ ☐ ☐ ☐ ☐ busco sempre este conteúdo

RECURSOS DA FERRAMENTA DE PESQUISA

CONSIDERANDO SEU PROCESSO DE PESQUISA PARA FILTRAR INFORMAÇÕES JORNALÍSTICAS, COMO VOCÊ AVALIA A IMPORTÂNCIA DAS SEGUINTE FUNÇÕES EM UM MECANISMO DE BUSCA DE JORNAL ON-LINE:

*

	sem importância	mais ou menos	importante
busca por autor	☐	☐	☐
busca por título	☐	☐	☐
busca por data ou período	☐	☐	☐
busca por editoria	☐	☐	☐
busca por gênero jornalístico (entrevista, crítica, reportagem)	☐	☐	☐
busca por tags / palavras-chave	☐	☐	☐
busca por formatos diferentes (texto, vídeo, áudio, infográfico)	☐	☐	☐
busca com operadores especiais (por exemplo, para exclusão de termos)	☐	☐	☐
busca por relevância da notícia (texto de grande repercussão, matéria premiada, reportagem especial)	☐	☐	☐

RESULTADOS DE PESQUISA E VISUALIZAÇÃO DA INFORMAÇÃO

APÓS REALIZAR UMA PESQUISA, QUAIS INFORMAÇÕES EXIBIDAS NO RESULTADO DE BUSCA AUXILIAM NA SUA DECISÃO DE ABRIR UM TEXTO ON-LINE PARA LEITURA.

*

	sem importância	mais ou menos	importante
nome do autor do texto	☐	☐	☐
título da matéria	☐	☐	☐
editoria	☐	☐	☐
data de publicação	☐	☐	☐
tamanho do texto (pequeno, médio ou grande)	☐	☐	☐
tag ou palavra-chave destacada	☐	☐	☐
trecho do texto exibido	☐	☐	☐

PERCEPÇÃO COMO PESQUISADOR

Se você procura um texto publicado no conteúdo on-line do jornal, mas o resultado de busca não recupera a informação. Qual sua reação mais frequente nessa situação? *

Escolha apenas duas (2) opções.

- ☐ refaz várias vezes as estratégias de pesquisa no buscador do site jornalístico
- ☐ usa um mecanismo de busca global (procura no Google)
- ☐ procura a informação que precisa navegando pelo site (por exemplo, recuando datas ou consultando páginas especiais)

Em geral, como você qualifica sua velocidade de pesquisa? *

- ☐ rápida (realiza pesquisas para checar dados)
- ☐ moderada (realiza pesquisas para entender temas e produzir textos de apoio)
- ☐ demorada (realiza pesquisa de estudo em profundidade, ampliando conhecimento sobre algum tema que auxilia na produção de textos mais elaborados, projetos especiais ou preparação para entrevistas)

Em geral, como você qualifica seu grau de utilização do mecanismo de busca do site jornalístico: *

- ☐ básico (realiza pesquisas simples sem optar por filtros ou outros recursos avançados)
- ☐ intermediário (realiza pesquisas mais elaboradas com alguns recursos avançados)
- ☐ avançado (domina o uso dos recursos avançados disponíveis e elabora pesquisas com estratégias mais complexas)

Na sua avaliação, os mecanismos de busca utilizados para recuperação de informações em jornais on-line possuem estruturas e atributos de pesquisa adequados e suficientes?

o conjunto de recursos são suficientes? considere a interface de pesquisa e possibilidades existentes para filtrar informações e apresentar resultados mais precisos

concordo	☐
concordo parcialmente	☐
nem concordo nem discordo	☐
discordo parcialmente	☐
discordo	☐

Em que medida os buscadores de sites jornalísticos satisfazem jornalistas e outros profissionais de comunicação em sua rotina de trabalho e sua dinâmica de pesquisa?

a ferramenta de pesquisa facilita sua rotina de trabalho?

nada satisfeito	☐
pouco satisfeito	☐
satisfeito	☐
muito satisfeito	☐

Comentário aberto: gostaria de acrescentar algo? indique sua principal dificuldade no uso de um mecanismo de busca de site jornalístico

Interessado no resultado da pesquisa? indique seu e-mail para contato:

Enviar

Nunca envie senhas em Formulários Google.

100% concluído.

Powered by
Google Forms

Este conteúdo não foi criado nem aprovado pelo Google.
[Denunciar abuso](#) - [Termos de Serviço](#) - [Termos Adicionais](#)