

Ciência da informação em transformação: *Big Data*, nuvens, redes sociais e Web Semântica

Renato Rocha Souza

Pós-Doutorado pela Universidade de South Wales – País de Gales - UK.

Doutor em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG - Brasil.

Professor da Fundação Getúlio Vargas (FGV-RJ) – Rio de Janeiro, RJ - Brasil.

Professor da Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG - Brasil.

<http://lattes.cnpq.br/4726949697973381>

E-mail: renato.souza@fgv.br

Maurício Barcellos Almeida

Pós-Doutorado pela State University of New York at Buffalo – Buffalo, NY - USA.

Doutor em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG - Brasil.

Professor Associado da Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG - Brasil.

<http://lattes.cnpq.br/5218069708058487>

E-mail: mba@eci.ufmg.br

Renata Maria Abrantes Baracho

Doutora em Ciência da Informação pela Universidade Federal de Minas – Belo Horizonte, MG – Brasil.

PDSE pela College of Information Sciences and Technology, Penn State University - State College, PA - USA.

Professora da Universidade Federal de Minas Gerais (UFMG) - Belo Horizonte, MG – Brasil.

<http://lattes.cnpq.br/4218954956709188>

E-mail: renatabaracho@ufmg.br

Recebido em: 15/08/2014. Aprovado em: 23/1/2015. Publicado em: 07/08/2015.

Resumo

Discute, sob uma perspectiva crítica, o panorama atual da ciência da informação e algumas das perspectivas futuras, em face dos fenômenos informacionais contemporâneos. Apresenta considerações sobre processo de esvaziamento temático através de análise de objetos de pesquisa atualmente desenvolvidos na ciência da informação e aponta o processo de absorção de outros destes objetos em diversas áreas do conhecimento. Questiona a primazia e a quantidade de pesquisas interdisciplinares na área e tece comentários sobre o que pode redundar desse processo para o futuro da ciência da informação. Em segundo momento, analisa o impacto para a área da explosão informacional, da consolidação das redes sociais como espaços de interação, da computação em nuvem, bem como de elementos constituintes da Web Semântica.

Palavras-chave: Ciência da informação. Representação do conhecimento. Interdisciplinaridade. *Big Data*. Cloud Computing. Social networks. Web Semântica.

A changing information science: Big Data, clouds and social networks

Abstract

This paper discusses, under a critical point of view, the current Information Science landscape and some future prospects, regarding contemporary information phenomena. It presents thoughts about the process of thematic deflating of Information Science through the analysis of the research objects currently under development in that field, as well as it points out the process of absorption of those and other relevant objects by other fields of knowledge. It challenges the emphasis and the volume of interdisciplinary research within the field and presents some comments about what might be the results of such process for the future of Information Science. Then, it analyses the impact in Information Science field of the phenomena like: information boom, the consolidation of the social networks as interactive spaces, cloud computing, as well as key elements of the semantic web.

Keywords: Information Science. Knowledge Representation. Epistemology. Interdisciplinarity. *Big Data*. Cloud Computing. Social Network. Semantic Web.

Ciencia de la Información en transformación: Big Data, nubes y redes sociales

Resumen

Este artículo discute, bajo un punto de vista crítico, el actual panorama de la Ciencia de la Información y algunas perspectivas de futuro, con respecto a fenómenos de información contemporáneos. Presenta reflexiones sobre el proceso de desinflado temática de Ciencia de la Información a través del análisis de los objetos de investigación actualmente en desarrollo en ese campo, así como señala el proceso de absorción de esos y otros objetos pertinentes por parte de otros campos del conocimiento. Desafía el énfasis y el volumen de la investigación interdisciplinaria en el campo y presenta algunos comentarios sobre lo que podría ser el resultado de dicho proceso para el futuro de la Ciencia de la Información. A continuación, se analiza el impacto en el ámbito de las Ciencias de los fenómenos como: auge de la información, la consolidación de las redes sociales como espacios interactivos, computación en nube, así como elementos clave de la web semántica.

Palabras clave: Ciencia de la Información. Representación del conocimiento. Epistemología. Interdisciplinariedad. Big Data. Computación em nube. Redes Sociales. Web Semántica.

INTRODUÇÃO

Dentre todos os campos científicos, e mesmo dentre aqueles mais recentes, pode-se seguramente apontar a ciência da informação (CI) como das mais introspectivas, no tocante às temáticas de pesquisa.

As questões conceituais subjacentes à área são, por vezes, foco de reflexões tão apaixonadas e profundas, que os seus objetos de estudo ficam por vezes obnubilados, relegados a segundo plano. Tendo em seus textos seminais trabalhos tão amplos como o de Bush (1945) ou Shannon e Weaver (1949), e partindo de um conceito tão multifacetado, como exposto em Barlow (1994), Hofkirchner (1999), e Capurro (2003), é surpreendente que, ao longo do tempo, alguns autores ainda esboçam certa concordância sobre o que se abriga sob sua égide (BORKO, 1968; ZINS, 2009). Sem embargo, não é livre de percalços a sedimentação de uma área relativamente nova, em comparação com as chamadas *hard sciences*.

No tocante a essa relativa novidade, Wersig (1993) nos chama a atenção para a necessidade de uma perspectiva pragmática e evolucionária, aliada às tentativas de construção teórica, e de uma aceitação da complexidade inerente de uma ciência interdisciplinar, de modo que os objetos de pesquisa empírica e as interfaces com outras áreas sejam subsídios para a construção e reconstrução permanentes da área. Em seu texto, descreve o profissional de CI fazendo uma analogia ao *Weaving Bird*, pássaro que constrói seus ninhos fazendo

intrincadas e elaboradas teias, numa referência às “costuras temáticas” comuns à práxis do cientista da informação.

Sendo uma ciência dinâmica, dada a fluidez e ubiquidade de seus objetos de pesquisa, o que se apresenta neste ensaio é uma visão de como podem se dar os caminhos de transformação paradigmática e epistemológica que percorreremos, com o advento de novas tecnologias, construtos e formas de manuseio da informação. Também permeia o texto certa provocação – assumidamente opinativa, e fruto do testemunho pessoal – que aponta um inexorável esvaziamento da área como uma ciência autônoma, o que tem se dado a partir da imbricação crescente dos objetos de pesquisa atualmente desenvolvidos em outras áreas – consolidadas ou não – do conhecimento. Tal fenômeno vem redundando na diminuição substancial, migração – ou esvaziamento – daqueles que poderiam ser considerados objetos legítimos e atavicamente ligados à ciência da informação para outras áreas do conhecimento. Sob esta perspectiva, a reconstrução permanente exaltada por Wersig (2003) não vem redundando no fortalecimento de uma ciência emergente e diferenciada pela abordagem interdisciplinar, mas talvez no surgimento de uma nova CI, quase metaciência, francamente hospitaleira, progressivamente amorfa, e que se ancora em olhares peculiares e idiossincráticos para objetos cada vez mais alheios.

O texto apresenta determinados panoramas: esvaziamento, transformação e perspectivas futuras, a serem desenvolvidos nas seções a seguir.

PRIMEIRO PANORAMA: ESVAZIAMENTO

Em sua célebre definição, que tenta capturar epistemologia e práxis, Borko (1968) afirma que a ciência da informação é:

A disciplina que investiga as propriedades e comportamento da informação, as forças que governam seus fluxos e os meios para processá-la, de modo a obter altos graus de usabilidade e acessibilidade. Está preocupada com aquele campo de conhecimento relacionado à origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação. Isto inclui a investigação das representações informacionais em sistemas naturais e artificiais, o uso de códigos para transmissão eficiente de mensagens, e o estudo dos dispositivos e técnicas para processamento de informação, como os computadores e seus sistemas. É uma ciência interdisciplinar derivada de e relacionada a campos do conhecimento como matemática, lógica, lingüística, psicologia, computação, pesquisa operacional, artes gráficas, comunicação, biblioteconomia, administração e outros similares. Contém componentes tanto das ciências puras – na medida em que questiona os objetos de estudo sem uma relação necessária com suas aplicações – quanto das ciências aplicadas, na medida em que desenvolve produtos e serviços. (1968, tradução dos autores)

Não é motivo de celeuma que, com tal definição, virtualmente quaisquer dos objetos de pesquisa acadêmica na atualidade envolvam algum componente ligado à ciência da informação. No entanto, é possível identificar um viés, ainda que amplo, em tais entrelinhas: a ciência da informação preocupa-se primariamente com processos de *representação do conhecimento* e de seus registros associados – a informação. É da natureza desses processos que a informação sofra sucessivas abstrações, modelagens e representações, de modo a ser organizada, transmitida, codificada, consumida. Buckland (1991) captura esses processos e objetos, considerando-os, todos, manifestações do conceito de informação.

Segundo Alvarenga (2003), os processos de representação são de dois tipos: a representação *primária*, que ocorre quando são criados

documentos como objetos que representam o conhecimento, e a representação *secundária*, que visa à organização de documentos com a inclusão em sistemas referenciais, ou seja, através da criação de metadados. De todas as maneiras, a representação de informação e do conhecimento ocupa espaço seminal e fulcral na área, como desenvolvemos a seguir.

Apesar de alguma divergência, há razoável acordo sobre o momento e as causas do nascimento da ciência da informação (PINHEIRO; LOUREIRO, 1995; DIAS, 2002). Ela surge para lidar com os problemas decorrentes do aumento da complexidade na gestão de acervos informacionais, na necessidade de adaptação das metodologias, da práxis; ambas oriundas dos campos da documentação e biblioteconomia; e no surgimento de tecnologias automatizadas para esses fins. Entretanto, o final do século XX assistiu à rápida expansão das tecnologias da informação e a aceleração dos processos de transformação da informação – criação, representação, armazenamento, organização, disseminação e consumo –, fazendo com que as características que originalmente justificaram a criação de uma “ciência da informação”, décadas atrás, sejam sentidos de modo exponencial.

À guisa de ilustração, observem-se os números que refletem a produção atual de registros de informação (SOUZA, 2008): dados estimados sobre a produção mundial de conteúdo digital apontam para a cifra de 281 bilhões de gigabytes gerados apenas no ano de 2007, ou seja, quase 50 gigabytes para cada ser humano vivo. Isto equivale a mais de cinco milhões de vezes o conteúdo de todos os livros já escritos. O número total de páginas na web é calculado como o de alguns trilhões. Cem horas de vídeo são enviadas ao *YouTube* a cada minuto, e mais de seis bilhões de horas de vídeo são assistidas por mês no *YouTube* – quase uma hora para cada pessoa do planeta¹. Comparados com os cerca de 50 milhões de minutos da vida de uma pessoa longaeva, já temos hoje seguramente muito

¹ <https://www.youtube.com/yt/press/pt-BR/statistics.html>

mais conteúdo disponível do que gerações de seres humanos poderiam assistir, mesmo se decidissem dedicar todas as suas vidas para tal.

As redes móveis, os dispositivos celulares, os *tablets* e outros *gadgets*; as bibliotecas digitais e os conceitos emergentes como o de *wearable computing* têm continuamente mudado contextos, encurtado os ciclos, reinventado os suportes materiais e tornado cada vez mais orgânica nossa relação com os registros de informação. Há um desencaixe progressivo, fazendo ruir as permanentes ligações entre informação e seus suportes usuais para registro e consumo, com as tecnologias digitais. A organização de imensas massas de dados necessita de novas e criativas soluções; nunca se precisou tanto de uma ciência da informação para orquestrar estes esforços.

Não obstante, a “ecologia técnica” que pretensamente favoreceria um renovado fôlego para a área – à medida que multiplica os problemas informacionais decorrentes e aumenta as possibilidades para suas soluções – tem, ironicamente, catalisado a migração paulatina de objetos de pesquisa genuínos para outras áreas do conhecimento. Como exemplo da migração temática, artefatos sociotécnicos como as bibliotecas digitais, instrumentos de representação do conhecimento como as ontologias, processos tais como a modelagem de domínios para desenvolvimento de software, o estudo de sistemas de informação e seus variados contextos de uso; se não considerados como propriamente pertencentes, ao menos têm sido ativamente pesquisados em áreas consolidadas – como a linguística, antropologia ou a sociologia – mas também em campos mais recentes, como a ciência da computação, a área de sistemas de informação, a engenharia de *software*, dentre outras.

Em contrapartida, temas como “responsabilidade social”, “interdisciplinaridade e transdisciplinaridade”, “informação e trabalho” e “sociedade da informação” (ANCIB, 2009) – não obstante relevantes, e evidentemente legítimos para uma ciência que é tanto social quanto aplicada – vêm direcionando o foco da área para os aspectos mais sociais (e

possivelmente fronteiriços). Do mesmo modo, a ascensão de pesquisas em temáticas ligadas à administração, como gestão do conhecimento e inteligência competitiva, aproximam a CI de campos como a administração e engenharia de produção, mas contribuem para a difusão das suas fronteiras e a dissolução de sua identidade.

Os cursos de graduação ainda buscam prover bibliotecários e arquivistas com formação instrumental, com pouca ênfase à pesquisa. E a pós-graduação têm recebido egressos de áreas diversificadas, mas que pouco contribuem para a transformação da área. Para distinguir aquilo que é pertinente, e o que se afasta da identidade científica nuclear, há que se delinear as linhas de corte. No caso da administração, por exemplo, uma aproximação legítima se daria ao lidar com os insumos de informação que compõem as atividades de gestão, refletidos nas seguintes questões: Quais os fluxos ou objetos informacionais inerentes aos processos de gestão? Como variam ao longo do tempo? De quais partes são compostos? Quais as suas características? São compostos por elementos generalizáveis ou são exemplos desses elementos? Quais são as relações entre esses insumos que compõem as atividades? Como devem ser representados nos diversos contextos sociais? Como devem ser representados em sistemas de informação? Tais questões, dentre outras, proporcionam entendimento da realidade subjacente, e resultam na possibilidade de organização e representação dos elementos constituintes dessa realidade. Apenas após tais questionamentos podem os administradores definir estratégias e tomar decisões, em seu escopo de trabalho, o qual é diferente daquele em que atua o profissional de CI. O profissional de CI, assim como o bibliotecário, executa trabalho de assessoria a usuários e outros profissionais.

No caso da sociologia, não parece razoável criar uma dicotomia entre informação e o contexto social, visto que são intrinsecamente ligados. Entretanto, cabe refletir em cada abordagem o que se considera um “meio” e o que se considera um

“fim”. Na sociologia, acredita-se, usam-se todos meios disponíveis, metodologias qualitativas e quantitativas, para estudar a estrutura social e a atividade humana. O contexto social, para a sociologia, é portanto um fim. No caso da CI, a informação é o fim, e a sociologia, um meio para estudar aspectos da informação, quando ela se encontra inserida em um contexto específico: a informação disseminada em uma empresa está posicionada em contexto social diferente da informação produzida em uma universidade. Não se trata, absolutamente, conforme já mencionado, de negar a legitimidade das contribuições de outras áreas, e sim de situá-las adequadamente no âmbito das questões historicamente ligadas à CI.

Em artigo de 2003, Araújo traça o caminho da ciência da informação rumo à configuração de ciência social, hoje dominante, afastando-se talvez de suas origens calcadas em uma abordagem positivista. Tal aproximação possibilitou à nascente ciência diferenciar-se com uma perspectiva pós-moderna (WERSIG, 2003; CARDOSO, 1996), adequada aos paradigmas de uso e usuários que surgiram nas últimas décadas do século XX. Em compensação, a mudança de enfoque se deu em detrimento dos aspectos mais nucleares e aplicados, que ajudariam a conferir uma identidade duradoura a um campo eternamente em construção. Da mesma maneira que se encontra dificuldade em definir a ciência da informação como uma ciência aplicada, discutem-se quais seriam as aplicações naturalmente decorrentes desta condição.

No tocante aos processos de representação da informação e do conhecimento, tão atavicamente ligados às origens da área, corre-se o risco de tornar-se a CI um espectador passivo de um rapidamente cambiante panorama. Apesar de se ter primazia no estudo e construção de instrumentos como as linguagens documentárias e interfaces de sistemas de informação, estes temas são hoje associados às ciências mais técnicas, que com mais propriedade os têm incorporados aos seus fazeres e construtos. No artigo anteriormente citado (1968), Borko analisa

os objetos de pesquisa listados no *Current Research and Development in Scientific Documentation* e lista nove categorias, a saber:

1. necessidades e usos de informação;
2. criação e cópia de documentos;
3. análises linguísticas;
4. tradução automática;
5. produção automática de resumos, classificação, codificação e indexação;
6. design de sistemas;
7. avaliação e análise;
8. reconhecimento de padrões;
9. sistemas adaptativos.

Desses objetos – bastante atuais, diga-se de passagem – apenas 1,2 e parte de 5 continuam sendo reconhecidos como parte do escopo nuclear de ciência da informação. Análises linguísticas (3) têm sido desenvolvidas mais propriamente nos campos da linguística computacional e linguística de corpus; ciência da computação e mesmo matemática aplicada. Sob a égide do título Processamento de Linguagem Natural, são desenvolvidas as categorias 3,4; grande parte de 5 e parte de 7. Design de sistemas (6) foi apropriado pelas áreas de ciência da computação, sistemas de informação e engenharia, dentre outras áreas. Finalmente, o campo crescente da inteligência artificial vem desenvolvendo os temas 8 e 9.

Observem-se os progressos realizados pela ciência da computação no escopo da Web Semântica e no campo específico dos *Knowledge Organization Systems* (W3C, 2009; NKOS, 2009); das tecnologias e metodologias para construção de ontologias (SILVA, 2008); nos sistemas de recuperação de informação (SOUZA, 2006) e nos processos de análise de domínios e modelização (WAND; STOREY; WEBER, 1999). Há muito que o campo da ciência da computação deixou de ser considerado como aquele que provê tecnologias e sistemas; e tomou a frente de áreas nucleares de organização de informação e conhecimento, ensejando novos diálogos e repertórios de parte da CI. Exemplos de áreas relativamente novas e

pujantes não faltam, como as citadas processamento de linguagem natural e a linguística computacional, além de uma recrudescida inteligência artificial, que vêm desenvolvendo pesquisas em classificação e indexação automáticas, construção automática de ontologias, redes neurais, análise do discurso, além de alimentar as pesquisas de ponta em *text mining*, *web mining*, *data mining* e assemelhados.

A pesquisa em ontologias constitui-se em bom exemplo para análise da CI, em seu aspecto original, ligado às suas raízes de organização da informação, e da CI interdisciplinar e multifacetada. Também nesse caso cabe distinção similar àquela aplicada à administração e à sociologia, de modo a verificar em que aspectos as pesquisas em ontologias se aproximam das abordagens caras ao núcleo principal da CI.

Segundo Obrst, Hughes e Ray (2006), a pesquisa em ontologias consiste de miríade de perspectivas provenientes das diferentes comunidades envolvidas: ciência da computação e inteligência artificial, como apontam, Guarino (1998), Gruber (1993); ciência da informação e biblioteconomia, como apontam Vickery (1997), Søerguel (1997); lógica e filosofia, como aponta Smith (2003). A diversidade de campos de pesquisa e a explosão de publicações sobre ontologias verificada a partir da década de 90 têm gerado interpretações equivocadas. Muitas vezes, as ontologias são consideradas aplicações puramente computacionais; em outros casos, considera-se que as contribuições provenientes da filosofia são por demais abstratas para gerar resultados práticos. De fato, uma avaliação sobre pesquisas e publicações sobre ontologias (Smith, 2003) demonstra que, aquilo que tem sido chamado de “ontologia”, muitas vezes consiste apenas em considerações relacionadas à semântica formal, à Teoria dos Modelos e à Teoria dos Conceitos. Tomando-se a Teoria dos Conceitos, vislumbra-se a proximidade com teorias utilizadas na CI (DAHLBERG, 1974). A questão fundamental, entretanto, é que a Teoria do Conceito considera que “conceito” é uma entidade localizada na mente das pessoas, uma proposta similar à da psicologia

(MARGOLIS; LAURENCE, 1999). Essa visão “conceitualista”, apesar de servir a diversos propósitos na CI como, por exemplo, em instrumentos como os vocabulários controlados, não é adequada à realidade atual, caracterizada pelo uso de computadores. Sistemas automatizados e computadores se tornam elementos obrigatórios ante o grande volume de informação disponível. E computadores necessitam de representações formais.

Representações formais podem ser resultado do estudo de ontologias formais, as quais tem raízes em Aristóteles², Kant³, Brentano⁴, Husserl⁵, dentre outros. Os princípios subjacentes à criação de representações formais têm sido utilizados em diversas áreas sobre o rótulo de “ontologia aplicada”. Pode-se imaginar que, uma vez que as ontologias carecem de representações formais, devem ser aplicações da computação. Nada está mais distante da realidade: nesse contexto, a palavra “formal” não diz respeito a formalismos matemáticos ou lógicos, mas a “teoria das formas”, ao estudo dos objetos, de suas partes, suas características, dentre outras (SMITH, 2003). A representação no nível requerido para uso por um computador, o qual é um artefato com limitações evidentes, exige um trabalho de organização e entendimento da realidade característicos da prática em CI.

As pesquisas em ontologias, agregando teorias fronteiriças à CI, podem facilitar a aplicação do conhecimento sobre organização e representação acumulado na área nos últimos cem anos, em insumos baseados em mídias digitais, prevalentes no mundo atual. Sejam ontologias, vocabulários controlados, terminologias ou outros instrumentos, fica evidente a importância da representação do conhecimento no conjunto interdisciplinar que se tornou a CI. Talvez, falar em identidade da área, diz respeito a resgatar tal importância. Cabe

² Aristóteles (384-322 a.C.), filósofo grego.

³ Immanuel Kant (1724 - 1804), filósofo alemão.

⁴ Franz C. H. H. Brentano (1838 - 1917), filósofo alemão.

⁵ Edmund G. A. Husserl (1859 - 1938), filósofo austríaco.

ainda enfatizar que, uma vez mantida a identidade epistemológica, o surgimento de novos instrumentos não requer a desconsideração de outros, conforme nos lembra Morin (2000, p. 2005):

A caminhada pelo pensamento [...] consiste em fazer ir e vir, incessantemente, entre certezas e incertezas [...]. Não se trata de abandonar os princípios da ciência clássica [...] mas de integrá-los em um esquema mais rico.

Nas universidades americanas e europeias a questão parece estar mais bem equacionada, à medida as “escolas de ciência da informação” se transformam em escolas de “estudos de informação”, e grupos interdisciplinares que acolhem os perfis necessários para exercer tal interdisciplinaridade sociotécnica. Nos Estados Unidos, um exemplo dessa discussão é o chamado movimento *i-school*, no qual grande número de universidades adotou novas denominações, posturas e mesmo currículo em escolas e departamentos de pesquisa relacionados à informação (DETLEFSEN, 2008). Embora se estudem temas como *Knowledge Management*, as linhas mestras que definem a área não se distanciam completamente dos aspectos nucleares; realizando, no entanto, uma costura que abrange a periferia temática sob uma perspectiva central (WIKIPÉDIA, 2009).

Os congressos da ISKO (2014) também denotam clara direção que visa à abrangência, mas mantêm o prumo teórico. A incorporação de temáticas que sugeririam um viés social como as *folksonomias*, acontece no esteio dos sistemas de organização de informação.

No Brasil, entretanto, observa-se sob muitos pontos de vista uma deriva conceitual e epistemológica. Como exemplos, temos o trabalho de Maia (2008), que demonstra a impossibilidade de realizar um agrupamento (*cluster*) conceitual de documentos apresentados no Enancib a partir de palavras-chave, tamanha a disparidade temática na área da Ciência da informação; ou o extenso estudo de Silva et al. (2006), que aponta o tema “Informação, Cultura e Sociedade” como o que contém o maior número de linhas de pesquisa, segundo o CNPq, sendo

responsável por quase 20% das linhas de pesquisa da CI no Brasil. O estudo mostra um percentual inferior a 25% para os estudos aplicados da CI.

Uma das agruras de ser a CI tão francamente interdisciplinar é a perspectiva de ver evanescer sua disciplinaridade, tão importante para a construção e preservação de sua identidade. Nas palavras de Francelin (2003), a CI experimenta hoje a necessidade de utilizar a receita que ela mesma propôs à biblioteconomia, ou seja, rever seus princípios e fundamentos de construção científica.

Conclui-se o primeiro panorama com a situação hipotética de pretensão diálogo entre Biblioteconomia e a ciência da informação, em que a primeira diria: “Dê-me um futuro, que lhe confiro um passado”. Há que cuidarmos para que o futuro não seja tornar-se um nicho dentre outras áreas do conhecimento. E que, em vez de um *Weaving Bird* de Wersig, o profissional da informação não se torne uma “Rêmora⁶”, alimentando-se das sobras temáticas daquilo que se desenvolve com maior propriedade em áreas do conhecimento diversificadas e mais pujantes.

SEGUNDO PANORAMA: TRANSFORMAÇÃO

Cabe reforçar que a apresentação de visões específicas, utópicas ou distópicas, sobre a ciência da informação não constituem, absolutamente, fato original. Pesquisas que desenharam a CI e o fenômeno da informação sob uma miríade de aspectos estão disponíveis na literatura, tanto nacional quanto internacional (CAPURRO, 1991; SOUZA, 2008; ODDONE, 2000; WHITTAKER, 2011; apenas para citar alguns). Os panoramas apresentados nesta seção são meramente fruto de observação participante e altamente opinativos.

Optou-se por estudar as imbricações dos temas: *Big Data*, *Cloud Computing*, Redes Sociais e Web Semântica na ciência da informação como possíveis

⁶ Peixe parasita que se acopla à barriga de peixes maiores para aproveitar-lhes os restos de alimentos.

causadores de impactos e mudanças paradigmáticas. Cada um destes temas será apresentado em seguida, assim como as perspectivas futuras a eles associados.

Redes Sociais: dentre os temas, talvez seja o mais decantado, e sobre o qual já se considerou mais amplamente em termos de transformação na área. As redes sociais são estruturas que unem atores – indivíduos e organizações – através de laços (EASLEY et al, 2010). Os laços, por sua vez, podem ser reificados por quaisquer tipos de arranjos, tecnológicos ou sociais. A representação e o estudo das redes incluem, mas não estão limitados, a áreas como a antropologia, biologia, comunicação, economia, geografia, ciência da informação, psicologia social, sociologia e sociolinguística.

Embora as redes sociais tenham contribuído para o fenômeno da globalização e dos novos arranjos sociais e as novas ordens mundiais (LEVY, 1999; CASTELLS, 2013), optamos por focar, principalmente, os fenômenos informacionais derivados, pelas consequências para a área. Limitamo-nos a analisar as redes sociais que se constroem amparadas em arranjos sociotécnicos, como é o caso da *web* (mais notadamente a chamada *web 2.0*), exemplificada pelas redes de relacionamento nascidas neste ambiente, como o Twitter⁷ e o Facebook⁸. Inicialmente descartado como fonte de informação relevante, o Twitter, com suas econômicas mensagens de 140 caracteres, começou a ser coletado⁹ em 2010 pela U.S. Library of Congress, em um acervo que já supera sobremaneira aquele que perfazem suas obras impressas. Segundo o diretor de comunicação da instituição:

O Twitter é um novo tipo de coleção para a Biblioteca do Congresso, mas uma coleção importante para sua missão. Na medida em que a sociedade se volta para as mídias sociais como seus meios primários de comunicação e expressão criativa, esta passa a suplementar e, em alguns

casos a suplantam as cartas, periódicos, e outras fontes rotineiramente coletadas por bibliotecas voltadas à pesquisa” (OSTERBERG, 2014, tradução dos autores).

Foram trocadas mais de 500 milhões de mensagens por dia¹⁰ em 2014. Junto com o Facebook, com cerca de 1.1 bilhão de usuários em 2014 e volume de informação impressionante¹¹, os dois sites obliteraram completamente a informação criada nos meios tradicionais, impressos ou digitais. A principal característica dessa informação é a fluidez, dinamicidade e efemeridade temática, que são exploradas para atividades de monitoramento ambiental, análise de sentimentos (LIU, 2012) e mesmo registro da história humana.

Assim como tomam espaço na produção de informação, esses arranjos engendram seus próprios modos de organização da informação. Muito propaladas em trabalhos da CI, surgiram no final da década de 2000 as folksonomias (PETERS, 2009; VANDER WAL, 2009), que podem ser vistas como taxonomias geradas – muitas vezes de maneira coletiva – pelos próprios usuários da informação. Segundo Guedes e Souza (2008), são utilizadas para categorizar e recuperar páginas, fotos, músicas, bookmarks, notícias, entre outros elementos da web. A Web 2.0 proporciona ao usuário a possibilidade de participar, gerando e organizando as informações, em um processo livre de classificação da informação desejada. Posteriormente, essa informação poderá ser recuperada por ele e pela comunidade que o cerca. O usuário utiliza uma ou mais *tags* próprias para classificar a informação. A folksonomia, neste sentido, é o oposto de um vocabulário controlado, pois a classificação é livre por parte do usuário. Em contraste com as taxonomias desenvolvidas por profissionais e pelos vocabulários controlados, o processo diminui bastante o custo de categorização do conteúdo, pois não existem hierarquias complexas de termos e nomenclaturas as quais o usuário tem que aprender. Além disso, a manutenção desses vocabulários é perene e dinâmica.

⁷ <https://twitter.com/>

⁸ <https://pt-br.facebook.com/>

⁹ <http://www.businessinsider.com/library-of-congress-is-archiving-all-of-americas-tweets-2013-1>

¹⁰ <http://www.internetlivestats.com/twitter-statistics/>

¹¹ <http://www.statisticbrain.com/facebook-statistics/>

Outro fenômeno digno de nota é a Wikipédia, um dos mais notáveis construtos coletivos – *crowdsourced* da atualidade. A Wikipédia¹² é uma enciclopédia multilíngue *on-line*, livre e colaborativa, possibilitando a cópia, modificação e ampliação de qualquer informação, desde que os direitos da cópia e modificações sejam preservados; ela está disponível em um *software* próprio, o MediaWiki, desenvolvido por voluntários e sob a licença GNU/ GPL. Distingue-se dos *blogs* por não ser autoral e vai além do resultado oferecido pelos *sites* de busca, já que a resposta é um texto único. Qualquer pessoa com acesso à Internet pode modificar qualquer artigo, e cada leitor é um colaborador em potencial, o que permite que vários autores possam trabalhar em conjunto, editando sucessivamente a mesma página. Um colaborador pode vir a assumir vários níveis de colaboração, como escrever e revisar artigos, corrigir falhas e erros ortográficos, colaborar esporadicamente com o projeto também em outras atividades, a saber, produzir *softwares*, traduzir artigos, e igualmente divulgar ideias ou participar de discussões pertinentes (MENDES; SOUZA, 2007).

Criada em 15 de janeiro de 2001, a enciclopédia é gerida e operada pela Wikimedia Foundation, uma organização sem fins lucrativos. Segundo os dados mais recentes, é composta de 30 milhões de artigos (843 772 em português em setembro de 2014). A maioria de entradas são artigos, mas o número total de entradas inclui imagens, páginas de usuários, páginas de discussão, etc. Graças ao formato digital, as enciclopédias deixaram de ser produtos de elite e passam a integrar o dia a dia de pesquisadores, escritores e alunos, convertendo-se em uma fonte de informação cada vez mais usada.

A filosofia por trás da Wikipédia prega que a informação publicada seja completamente neutra, por isso os artigos não são assinados e os escritores devem buscar a máxima imparcialidade. Para garantir a neutralidade e evitar os erros, a Wikipédia tem um mecanismo interessante que consiste na

possibilidade de qualquer um entrar em um artigo e modificá-lo. A isto se chama “revisão por iguais”, ou seja, cada artigo escrito é corrigido em tempo real por vários colaboradores do mesmo nível do autor. Além disso, um sofisticado esquema de curadoria impede que erros crassos sejam mantidos por muito tempo, fazendo desta a maior, mais atualizada, mais utilizada e mais completa enciclopédia de todos os tempos.

Como reflexão para a área, as redes sociais e seus fenômenos agregados nos apontam para a efemerização tanto dos registros informacionais quanto dos instrumentos necessários para organizá-los. As construções coletivas nos permitem a confecção de produtos de alcance inimaginável (SUROWIECKI, 2004), embora ainda haja grande resistência na adoção desses produtos na área da biblioteconomia e ciência da informação, como auxiliares na produção e organização de informação. Há que se aliar às tradicionais abordagens analítico-sintéticas de construção de *Knowledge Organization Systems* (KOS) a possibilidade de se trabalhar com frequências de palavras, dicionários dinâmicos e (como o *urban dictionary*¹³ ou o *free dictionary*¹⁴) para análise dos léxicos emergentes. Também se faz necessário incorporar nos instrumentos de organização expressões não dicionarizadas (como os memes¹⁵, os *emoticons*¹⁶ e, mais recentemente, os *kaomoji*¹⁷). Quanto ao volume de informações gerado nessas redes, abordaremos as consequências no tópico a seguir.

Big Data: permeando as questões anteriores, o tema da explosão informacional – atualmente tratada como *information / data deluge* – é mais do que o aumento vertiginoso de registros. O termo *Big Data* se tornou banal, visando a descrever quaisquer amontoados de dados que não possam ser processados sem que se tenha à disposição estruturas de computação específicas. Algumas definições

¹² <http://pt.wikipedia.org/wiki/Wikip%C3%A9dia>

¹³ <http://www.urbandictionary.com/>

¹⁴ <http://www.thefreedictionary.com/>

¹⁵ <http://en.wikipedia.org/wiki/Meme>

¹⁶ <http://en.wikipedia.org/wiki/Emoticon>

¹⁷ <http://en.wikipedia.org/wiki/Kaomoji>

atribuem, além de alto volume, as características de “variedade”, “velocidade” e “veracidade” (LANEY, 2012). Fato é que desde o início dos anos 1990 que se manipulam grandes volumes de dados para fins de suporte à gestão ou predição de cenários, mas os anos 2000 viram crescer a produção, nesses contextos, do uso de informação não estruturada, principalmente de cunho textual.

O impacto deste fenômeno na academia, na pesquisa e, mais notadamente, em áreas como a ciência da informação e a ciência da computação, é enorme. Além de experimentarem grande convergência nas últimas décadas, têm buscado subsídios em áreas como a linguística computacional, computação gráfica, matemática aplicada, visualização de informação e bancos de dados para depreender sentidos. As técnicas tradicionais de análise de assunto, análise documentária ou análise temática com fins de indexação, classificação e/ou categorização buscam adaptar-se aos acervos gigantescos e à velocidade com que se produzem sumários, classes e índices, recorrendo cada vez mais a técnicas de visualização de informações (MAGNUSSON; VANHARANTA, 2003; NODUS LABS, 2012; SCHNEIDER, 2014; dentre outros).

Além do fomento às técnicas de manipulação de informação desestruturada, oriundas da ciência da informação – ao menos em seu projeto original – o excesso tem propiciado os modos de se fazer pesquisa. Como observamos em (SOUZA, 2006), ao pensarmos a relação da sociedade com os estoques de informações disponíveis, há inegável democratização do acesso, tanto pela maior disponibilidade dos meios de comunicação e a popularização das tecnologias, quanto pelo alcance a uma gama mais significativa e diversificada da produção cultural da humanidade. Torna-se mais difícil o controle das informações, pois as fontes são tantas e tão variadas que os vieses tornam-se mais explícitos e facilmente contornáveis.

Paradoxalmente, é cada vez mais complexo estabelecer parâmetros para julgamento da qualidade da informação, exatamente porque nenhuma amostra é mais significativa diante do

todo. No caso da academia, o trabalho de revisão bibliográfica, fundamental para a pesquisa, está ficando inviável. São necessários recortes explícitos e muitas vezes arbitrários, pois a quantidade de publicações e fontes disponíveis sobre assuntos específicos é frequentemente intratável. Tanto pior: excetuando estudos em campos de conhecimento mais perenes, como a filosofia, o fenômeno da rápida obsolescência torna o conhecimento produzido cada vez mais datado. Em outra faceta, no campo da ciência, é agora comum o processo de *data driven research* (MOE-BEHRENS, 2012). Nesse paradigma, realizam-se inúmeros testes de hipóteses em grandes massas de dados como preâmbulos exploratórios e, à medida que se encontram caminhos promissores derivados dos dados, busca-se aprofundar quantitativa e qualitativamente através de mudanças de foco iterativas. Cabe às ciências tradicionais incorporar tais dinâmicas em seus cabedais metodológicos.

Cloud Computing: tendo apontado o fenômeno social, materializado pelas redes; e o fenômeno da explosão informacional, representado pelo *hype* em torno do *Big Data*, cumpre finalizar este panorama apresentando o fenômeno da computação em nuvem, que adiciona mais uma camada de abstração ao processo de desmaterialização dos registros de informação. Se antes trocávamos (talvez com receio e muita desconfiança) os suportes físicos pelos digitais, materializados em *tablets* e computadores pessoais; a computação em nuvem abstrai e afasta do usuário os suportes físicos onde ficam armazenados seus dados. Onipresentes, os serviços de armazenamento (*DropBox*¹⁸, *Google Drive*¹⁹, *iCloud*²⁰, dentre outros) fazem parte da vida de centenas de milhares de pessoas que, aparentemente, não perdem o sono com receio de terem perdido ou serem roubadas em seus arquivos e fotos, a despeito de uma dezena de incidentes

¹⁸ <https://www.dropbox.com/>

¹⁹ <https://drive.google.com/>

²⁰ <https://www.icloud.com/>

recentes. A popularização dos smartphones e o recente lançamento do Google Glass²¹ apontam que, literalmente, com as nuvens, o céu é o limite.

As redes sociais romperam limites geográficos de conectividade; a nova explosão de informações rompeu com a possibilidade do tratamento individual e intelectual aos registros; e as nuvens romperam, enfim, com a tangibilidade dos registros. Qual seria a agenda de pesquisa da CI para o século XXI?

Web Semântica: ainda em relação à explosão informacional, é possível verificar o desenvolvimento de um conjunto de iniciativas no sentido de melhorar a recuperação da informação na web, ao qual se convencionou chamar de “Web Semântica”. No âmbito da Web Semântica, desenvolveram-se linguagens declarativas computacionais mais poderosas e, portanto, capazes de expressar “semântica”. Entretanto, uma questão que ainda causa confusão é saber que semântica é essa da Web Semântica? No escopo da Web Semântica, algo é considerado como capaz de expressar semântica quando pode ser entendido e processado via computador. Assim, o termo “semântica” em “Web Semântica” não trata da semântica da linguagem natural, mas sim de um subconjunto desta, denominado semântica formal.

Diversos foram os modelos, para entendimento da linguagem e sua implementação em computadores, propostos a partir da segunda metade do século XX: o modelo estatístico, o modelo sintático, o modelo lógico, o modelo lexical e o modelo neural (SOWA, 2004). Cada um desses modelos é baseado em uma teoria específica (estatística, matemática, gramática, etc) e ignora aspectos da linguagem para os quais a tecnologia não apresenta boas soluções. Os modelos lógicos, baseados em fundamentos filosóficos, têm produzido teorias formais de qualidade superior se comparadas a outras abordagens. As teorias formais têm sido largamente usadas na pesquisa em ontologias no contexto da Web Semântica.

Entretanto, cabe lembrar que os modelos produzidos via semântica formal não são capazes de lidar com linguagem natural utilizada pelas pessoas com propósito de comunicação. Logo, a linguagem da Web Semântica permanece ainda restrita a apenas um subconjunto da linguagem natural, composto por sentenças similares à linguagem natural, adequadas para máquinas, mas sem a sua capacidade de comunicação.

Ainda assim, ontologias têm sido um tema de interesse em ciência da informação, o qual foi destacado pioneiramente por Brian Vickery (VICKERY, 1997). De fato, ontologias são o meio mais utilizado mundialmente para estruturar informação e conhecimento, em campos de pesquisa relevantes para a sociedade humana, como a medicina, o direito, a engenharia, dentre outros. Ao contrário do que se imagina, a pesquisa em ontologias apresenta muitas similaridades reais com pesquisas consagradas em ciência da informação (ALMEIDA, 2013), mesmo que muitas vezes o termo “ontologia” não seja adotado.

Questões enfrentadas por cientistas da computação para a construção de vocabulários legíveis por máquina são questões com as quais os cientistas da informação já têm se deparado por muitos anos (VICKERY, 1997). Existem similaridades e diferenças entre ontologias da computação e os tesouros e taxonomias, dois outros termos amplamente usados na ciência da informação. A possibilidade de restringir a linguagem natural parece ser o ponto de contato entre esses três tipos de estruturas.

O termo também tem sido citado em pesquisas diversas conduzidas no âmbito da ciência da informação. Fox (1983) usa fundamentos filosóficos para explicar sua hipótese que a informação tem um *status* ontológico similar às proposições subjacentes a um texto: ambos são atemporais e não espaciais, além de serem objetos abstratos. As proposições não coincidem com o texto, à medida que várias sentenças podem expressar a mesma proposição. Nessa linha de pensamento, a informação contida em um documento não pode ser identificada com

²¹ <https://www.google.com/glass/start/>

o texto desse documento, mas como seu conteúdo proposicional. Essa abordagem proposicional se originou do estudo da lógica e da filosofia da linguagem, em particular de uma distinção sugerida por Frege, o qual propôs uma teoria representada pelos três vértices de um triângulo, denominados símbolo, significado e referência. A mesma distinção foi apresentada mais tarde por Ogden e Richards (1972) em seu triângulo do significado, bem como por Dahlberg (1978) em seu triângulo do conceito.

No contexto dos sistemas de recuperação da informação, Blair (2006) também adota fundamentos filosófico-ontológicos baseados na filosofia segunda de Wittgenstein, a qual, por sua vez, é baseada na noção de categorização denominada “semelhança de famílias”. A significância dessa visão para o projeto de sistemas de recuperação da informação reside no fato de quais tais sistemas seriam, por natureza, indeterminados. As melhores iniciativas assim seriam no sentido de mitigar essa indeterminação, buscando o equilíbrio entre revocação e precisão.

Existem diversas outras abordagens em ciência da informação que envolvem fundamentos ontológicos, mesmo que nem sempre os autores mencionem o termo “ontologia”. Ainda assim, princípios ontológicos aparecem claramente em trabalhos daquela área, como por exemplo em Ranganathan (1967), Lancaster (1986), Foskett (1985), e Buckland (1991), para citar alguns.

Mesmo assim, ao procurar o termo ontologias em um mecanismo de busca, é possível achar tanta variedade de estudos em diferentes campos científicos, que pode ser difícil entender a que campo de pesquisa pertence e como pode ser útil seu estudo. Entretanto, ontologias são um assunto de pesquisa genuíno e relevante para a ciência da informação, o qual fomenta a interdisciplinaridade característica dessa área (ALMEIDA, 2013). A identificação da interseção entre as estruturas de organização da informação da Web Semântica, as ontologias, e teorias e autores já consagrados na literatura de ciência da informação, é importante para caracterizar o argumento que tem sido desenvolvido no presente

artigo, ou seja, de que pesquisas relevantes da ciência da informação têm sido absorvidas por outras áreas de conhecimento e esvaziadas no contexto da própria ciência da informação.

DISCUSSÃO: PERSPECTIVAS FUTURAS

Tendo perdido a primazia, como área, e talvez mesmo a centralidade, na temática do tratamento de informação, que papel caberia à CI? Retomando a provocação feita ao final do primeiro panorama, o que poderia conter este esvaziamento? Revisitando a definição de Borko (1968) para a CI, como “a disciplina que investiga as propriedades e comportamento da informação, as forças que governam seus fluxos e os meios para processá-la, de modo a obter altos graus de usabilidade e acessibilidade”, podemos conceder que esse programa se tornou amplo demais para a nossa área. Talvez para qualquer área. Entretanto, a proposta de Wersig (1993) se torna cada vez mais presente. Wersig, com o ponto de vista da CI, reconhece o novo papel do conhecimento na sociedade contemporânea, e prega para nossa área um pragmatismo no sentido de lidar com problemas. Assim, menos importância se daria aos estatutos epistemológicos.

Na interpretação de Freire e Araújo (2001), o terceiro modelo de Wersig sugere para a CI uma estrutura teórica que considere menos a formulação de leis gerais e mais a de estratégias de ação, mediante uma abordagem de entrelaçamento de conceitos científicos. Neste modelo, os conceitos fundamentais “... se constituem semelhantemente a ímãs, ou “atratores”, atraindo os materiais [teóricos ou empíricos] para fora [dos seus respectivos campos científicos] e reestruturando-os dentro da estrutura científica da informação. ...” Dessa maneira, seria tecida uma protorrede de conceitos básicos em ciência da informação, a partir da qual outros indivíduos ou grupos poderiam encontrar e entretecer outros fios soltos, fazendo a rede ainda mais inclusiva e mais apertada, de modo a aumentar seu caráter científico.

Muitos programas de graduação e pós-graduação em CI têm buscado intercâmbios maiores com outras áreas. O contato com a ciência da computação aparece

no sentido de se desenvolverem melhores concepções de sistemas de informação; ou, juntamente com a engenharia, a estatística e a matemática aplicada, ajudando a prover modelos que permitam a evolução das tarefas de indexação, classificação/*clustering* e os *pipelines* analíticos necessários ao tratamento da informação. Os corpora e os sistemas de organização de conhecimento (SOC/KOS) têm sido estudados conjuntamente com a linguística de corpus, com a área de processamento de linguagem natural e mesmo a área da Ontologia, em tarefas de *Information Mining* e *Knowledge Extraction*, além de *Knowledge Discovery in Databases*²². A área de estudos de usuários tem incorporado conceitos de ergonomia cognitiva e usabilidade, e mesmo tentativas de diminuir as lacunas entre busca de informação e sistemas de informação apareceram recentemente com viés de CI (INGWERSEN; KALERVO, 2005). Os estudos de usuários em rede buscam e se beneficiam dos estudos de redes complexas e análise de redes sociais. E, influenciados pelas exortações à codificação em linguagens de programação²³, muitos programas de pós-graduação em CI consideram introduzir mais este letramento digital em seus currículos.

Retomando o ponto, mas sob outra perspectiva, a interdisciplinaridade nos enfraquece enquanto área. Em contrapartida, traz a possibilidade – e até mesmo a prerrogativa – de mediação dos diálogos disciplinares. Essa essência interdisciplinar exorta o cientista da informação a navegar nos espaços teóricos, adaptar-se aos contextos tecnológicos e reinventar-se continuamente. Ou assim deveríamos ser. Ou adotamos a perspectiva pós-moderna de Wersig, ou mesmo a Modernidade de Berman, quando diz que:

Outros acreditam que as formas realmente distintas da arte e do pensamento contemporâneos deram um salto quantitativo para além de todas as diversas sensibilidades do modernismo e ganharam o direito de se chamar a si próprias de “pós-modernas”. (...) Ser moderno, eu dizia, é experimentar a vida pessoal e social como um maelstrom, encontrar o próprio mundo e a si próprio em perpétua desintegração e renovação, agitação e angústia, ambiguidade e contradição: ser parte de um universo no qual tudo o que é sólido se desmancha no ar.

²² http://en.wikipedia.org/wiki/Knowledge_extraction

²³ <https://learn.code.org/>; <http://csedweek.org/>

REFERÊNCIAS

- ALMEIDA, M.B. Revisiting ontologies: a necessary clarification. *Journal of the American Society of Information Science and Technology*. v.64, n.8, p.1682–1693, 2003. DOI: 10.1002/asi.22861.
- ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. 2003. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/97>>. Acesso em: 21/12/2009
- ANCIB. Histórico do Enancib. Disponível em: <<http://www.ancib.org.br/enancib/historico-do-enancib/>>. Acesso em: 24/01/2013
- ARAÚJO, C. A. A. A ciência da informação como ciência social. *Ciência da Informação*, v.32. n.3, p.21-27, 2003.
- BARLOW, J. P. A taxonomy of information. *Bulletin of the American Society for Information Science*, v. 13, n.17, 1994.. Disponível em: <http://findarticles.com/p/articles/mi_qa3633/is_199406/ai_n8713489/?tag=content;col1>. Acesso em: 10/08/2004
- BERMAN, Marshall. *Tudo o que é sólido desmancha no ar*. São Paulo: Companhia das Letras, 2001.
- BORKO, H. (1968). Information science: what is this?. *American Documentation*, v.19, p.03-05.
- BUCKLAND, M. Information as thing. *Journal of the American Society for Information Science* (1986-1998). Jun. 1991.
- BUSH, Vannevar. As we may think, in the atlantic monthly. Jul., 1945 Disponível em: <<http://ccat.sas.upenn.edu/jod/texts/vannevar.bush.html>>. Acesso em: 29/08/2001
- CAPURRO, R. The concept of information. *Annual Review of Information Science and Technology* (ARIST), v.37, n.8, p.343-411, 2003.
- CAPURRO, R. What is information science for?: a philosophical reflection was published In:
- VAKKARI, Pertti; CRONIN, Blaise (Ed.). *Conceptions of Library and Information Science. Historical, empirical and theoretical perspectives*. London: Taylor Graham, 1992, p. 82-98.
- CARDOSO, A. M. P. Pós-modernidade e informação: conceitos complementares? *Perspectivas em Ciência da Informação*, v.1, n.1, p.63-79, 1996.
- CASTELLS, Manuel. *Redes de indignação e esperança: movimentos sociais na era da internet*. Rio de Janeiro: Zahar, 2013.
- DALBERG, I. Towards a theory of concept. *International Classification*, v.1, n.1, p.12-18, 1974.
- DAHLBERG, Ingetraut. *A referent-oriented, analytical concept theory for Interconcept*. *International classification*, n. 5, 1978. p. 142-151.

- DETLEFSEN, E. G. What's is a name?: the I-Schools project. *Medical Library Association News*, v.28, n.3, p.1-17, 2008.
- DIAS, E. J. W. O específico da ciência da informação. In: AQUINO, M. (Ed.). *O campo da ciência da informação: gênese, conexões e especificidades*. João Pessoa: Editora Universitária, 2002. p. 87-99.
- EASLEY, David; KLEINBERG, Jon. "Overview". *Networks, Crowds, and Markets: reasoning about a highly connected world*. Cambridge University Press, 2010. p.1-20.
- FOSKETT, D.J. Classification and integrative levels. In: CHAN, L.M.; RICHMOND, P.A.; SVENONIUS, E. (Ed.), *Theory of subject analysis* Littleton. CO: Libraries Unlimited, 1985. p. 212-220.
- FRANCELIN, M.M. A epistemologia da complexidade e a Ciência da Informação. *Ciência da Informação*. v.32, n.2, p. 64-68, 2003.
- FREIRE, Isa Maria; ARAÚJO, Vânia M.R H. Tecendo a rede de Wersig com os indícios de Ginzburg. *DataGramaZero - Revista de Ciência da Informação*, v.2, n.4, Ago. 2001.
- FOX, Edward et al. *Building digital libraries made easy: toward open digital libraries*.2002. Disponível em: <<http://www.springerlink.com/content/7x87gky9k5ggt5yq/>>. Acesso em: 13/03/2009
- GUARINO, N. Formal ontology and information systems. In: GUARINO, N. (Ed.). *Formal ontology in information systems*. Amsterdam: IOS Press, 1998. p. 3-15.
- GRUBER, T. *What is an ontology?*. 1993. Disponível em: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: 31/08/2014
- GUEDES, Roger Miranda; SOUZA, Renato Rocha. Navegando entre nuvens de etiquetas: uma proposta de utilização da tag cloud em catálogos eletrônicos de bibliotecas. *Ponto de Acesso*, v.2, n.3, p.2-13, 2008.
- HOFKIRCHNER, W. The quest for a unified theory of information. In: HOFKIRCHNER, W. (Ed.). *Proceedings of the Second International Conference on the Foundations of Information Science*. Amsterdam: Gordon and Breach, 1999. p. 9-30.
- INGWERSEN, Peter; JÄRVELIN, Kalervo. *The turn: integration of information seeking and retrieval in context*. Springer, v. 18. 2005, XIV, 448 p. (Series: The Information Retrieval Series).
- INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION. ISKO Chapters. Disponível em: <<http://www.isko.org/chapters.html>>. Acesso em: 22/08/2014.
- LANEY, Douglas. *The Importance of 'Big Data': a definition*. Gartner. 2012. Disponível em: <<https://www.gartner.com/doc/2057415/importance-big-data-definition>>. Acesso em: 06/10/2014.
- LANCASTER, F.W. *Vocabulary control for information retrieval*. Arlington, VA: Information Resources Press,1986.
- LÉVY, Pierre. *Cibercultura*. São Paulo: Editora 34, 1999.
- LIU, Bing. *Sentiment analysis and opinion mining*. Californi: Editora San Rafael, 2012.
- MAGNUSSON, Camilla; VANHARANTA, Hannu. Visualizing sequences of texts using collocational networks. In: Third International Conference, MLDM, 2003, Leipzig, Germany. *Machine Learning and Data Mining in Pattern Recognition.*, 2003, v.2734, p. 276-283. (Lecture Notes in Computer Science)
- MAIA, L.C.G. Uso de sintagmas nominais na classificação de documentos eletrônicos. 2008. PhD Thesis. Escola de Ciência da Informação. Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2008.
- MARGOLIS, E.; LAURENCE, S. *Concepts: core readings*. Cambridge: MIT Press,1999.
- MENDES, K.I.; SOUZA, R.R.A enciclopédia digital e o fenômeno wikipédia: mais sobre o fenômeno da desmaterialização dos registros de informação. In: ENCONTRO NACIONAL DE ESTUDANTES DE BIBLIOTECOMIA, DOCUMENTAÇÃO, CIÊNCIA E GESTÃO DA INFORMAÇÃO. 2007, São Carlos. *Anais...* São Carlos: Enebd. Disponível em: <<http://www.ufscar.br/enebd2007/index1.html>> Acesso em: 16/09/2014.
- MOE-BEHRENS, Gerd. *From hypothesis to data driven research*. 2012. Disponível em: <<http://www.sciencetogrok.com/2012/12/from-hypothesis-to-data-driven-research.html>>. Acesso em: 22/09/2014.
- MORIN, E.; LE-MOIGNE, J.L. *A inteligência da complexidade*. São Paulo: Editora Petrópolis, 2000.
- NKOS (2009). Simple Networked Knowledge Organization Systems/Services. Retrieved. mar. 21, 2009. Disponível em: <<http://nkos.slis.kent.edu/>>. Acesso em: 15/11/2011
- NODUS LABS. Texttexture: visualize any text as a network. 2012. Disponível em: <<http://texttexture.com/>>. Acesso em: 15/09/2014
- OBRST, L.; HUGHES, T.; RAY, S. Prospects and possibilities for ontology evaluation: the view from NCOR. In: Proceeding of the Fourth International Workshop on Evaluation of Ontologies for the Web at the 15th International World Wide Web Conference. Edinburgh, UK. 2006
- ODDONE, Nanci Elizabeth. A ciência da informação à luz das redes sociotécnicas, das controvérsias científicas e da circulação do saber. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 4., 2000, Brasília. *Anais...* Brasília: UnB, 2000.
- OGDEN, Charles K.; RICHARDS, Ivor A. *O significado de significado: um estudo da influência da linguagem sobre o pensamento e sobre a ciência do simbolismo*. Rio de Janeiro: Zahar, 1972.

- OSTERBERG, G. *Update on the Twitter archive at the Library of Congress*. Disponível em: <<http://blogs.loc.gov/loc/2013/01/update-on-the-twitter-archive-at-the-library-of-congress/>>. Acesso em: 29/03/2014
- PETERS, Isabella. *Folksonomies. indexing and retrieval in Web 2.0*. Berlin: De Gruyter Saur, 2009.
- PINHEIRO, L.; LOUREIRO, J. Traçados e limites da ciência da informação. *Ciência da informação*, v.24, n.1, p.42-53, 1995.
- RANGANATHAN, S.R. *Prolegomena to library classification*. London: Asia Publishing House, 1967.
- SHANNON, C.E.; WEAVER, W. *The Mathematical theory of communication*. Urbana: University of Illinois Press, 1949.
- SCHNEIDER, Bruno. *Visualização em multirresolução do fluxo de tópicos em coleções de texto*. 2014. 88 f. Dissertação (mestrado) – Fundação Getúlio Vargas, Escola de Matemática Aplicada.
- SILVA, D.L. Uma proposta metodológica para construção de ontologias: uma perspectiva interdisciplinar entre as Ciências da Informação e da Computação. 2008. Dissertação Mestrado em Ciência da Informação - Universidade Federal de Minas Gerais.
- Silva, E.L. et al. (2006). Panorama da pesquisa em Ciência da Informação no Brasil. *Informação & Sociedade: estudos*, v.16, n.1, p. 159-177, 2006.
- SOWA, John. *A dynamic theory of ontology*. 2004. Retrieved. 23 Jun., 2008. Disponível em: <<http://www.jfsowa.com/pubs/dynonto.htm>>. Acesso em: 16/09/2014.
- SMITH, B. *Ontology and information systems*. 2003. Disponível em: <[http://www.ontology.buffalo.edu/ontology\(PIC\).pdf](http://www.ontology.buffalo.edu/ontology(PIC).pdf)>. Acesso em: 16/09/2014.
- SØERGUEL, D. *Functions of a thesaurus, classification and ontological knowledge bases*. 1997. Disponível em: <<http://www.clis.umd.edu/faculty/soergel/soergelfctclass.pdf>>. Acesso em: 09/07/2010
- SOUZA, R.R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. *Perspectivas em Ciência da Informação*, v.11, n. 2, 2006. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/issue/view/25>>. Acesso em: 23/11/2002
- SOUZA, R.R. Quando ignorar é preciso!. *Diversa*, v.11, n.2, p.25-28, 2008. Disponível em: <<https://www.ufmg.br/diversa/16/index.php/tendencia/quando-ignorar-e-preciso>>. Acesso em: 14/04/2010
- SUROWIECKI, James. *The Wisdom of Crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. Little Brown. XXI, 2004.
- VANDER WAL, Thomas. Folksonomy Coinage and Definition. Disponível em: <<http://vanderwal.net/folksonomy.html>>. Acesso em: 14/01/2014
- VICKERY, B.C. Ontologies. *Journal of Information Science*, v.23, n.4, p.227-286, 1997.
- W3C. Simple Knowledge Organization System. 2003. Disponível em: <<http://www.w3.org/2004/02/skos/>>. Acesso em: 03/11/2008
- WAND, Y.; STOREY, V.C.; WEBER, R. An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems*, v.24, n.4, p.494-528, 1999.
- WERSIG, G. Information science: the study of postmodern knowledge usage. *Information Processing & Management*, v.29, n.2, p.229-239, 1993.
- WHITTAKER, S. Personal information management: from information consumption to curation. *Annual Review of Information Science and Technology*. v.45, n.1, p. 1-62, 2011.
- WIKIPEDIA. *Information Science: related fields and sub-fields*. 2009. Disponível em: <http://en.wikipedia.org/wiki/Information_science>. Acesso em: 14/12/2013
- ZINS, C. *Knowledge map of information science: issues, principles, implications*. 2009. Disponível em: <<http://www.success.co.il/is/index.html>>. Acesso em: 03/05/2013