

Big data e Organização do Conhecimento: reflexões iniciais a partir de uma proposta classificatória da produção científica

Fabio Orsi Meschini

Doutorando; Universidade de São Paulo, São Paulo, SP, Brasil;
fabioemeschini@usp.br; ORCID: <https://orcid.org/0000-0001-8191-6245>

Marivalde Moacir Francelin

Doutor; Universidade de São Paulo, São Paulo, SP, Brasil;
marivalde@usp.br; ORCID: <https://orcid.org/0000-0002-9576-7743>

Resumo: Considerando os impactos e as complexidades existentes no contexto *big data* e observados na Organização do Conhecimento, tornam-se necessárias reflexões que destaquem as possíveis contribuições dessa área para o desenvolvimento informacional da sociedade. Tendo isso em vista, a presente pesquisa busca compreender o fenômeno *big data* e seus impactos na Organização do Conhecimento por meio de uma proposta classificatória da produção científica. Para tanto, foram coletados artigos relacionados aos temas *big data* e Organização do Conhecimento nas bases de dados Scopus, *Web of Science*, *Library and Information Science Abstracts* e Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação. Os artigos foram classificados por meio da leitura dos títulos, resumos, palavras-chave e textos dos artigos. Os resultados demonstram a predominância de um perfil mais aplicado, tecnológico e orientado a processos relacionados à coleta, curadoria e utilização de dados. Concluiu-se ser necessária a ampliação do número de pesquisas relacionadas às questões sociais, cognitivas, epistemológicas e metodológicas, bem como a elaboração de mais estudos voltados à ética neste intenso contexto de dados. Isso contribui para compreender e ampliar os cenários de pesquisa sobre *big data* na Organização do Conhecimento.

Palavras-chave: *Big data*; Organização do Conhecimento; Produção científica; Classificação; Ciência da Informação

1 Introdução

Os impactos e as complexidades presentes no contexto *big data* e observados na Ciência da Informação e Organização do Conhecimento, demandam reflexões que destaquem as possíveis contribuições dessas áreas para o desenvolvimento social e informacional da sociedade, atualmente pautada em dados. Uma das

formas de análise destas contribuições das áreas em um contexto complexo de dados reside em estudos da literatura científica, uma vez que, conforme destacam Renault (2007) e Arboit, Bufrem e Freitas (2010), esta literatura expressa o comportamento de uma comunidade científica, ressaltando, portanto, seus quesitos teóricos, epistemológicos e metodológicos. Mueller (2007), Macias-Chapula (1998) e Glänzel (2003) ainda destacam que análises do conteúdo da literatura científica propiciam a consecução de indicadores que irão retratar as formas pelas quais determinados campos científicos consolidam seus conhecimentos.

As teorias, epistemologias e metodologias presentes na Ciência da Informação e Organização do Conhecimento possuem correlações com a realidade massiva de dados imposta pelo *big data* e podem contribuir com discussões relacionadas às complexas questões sociais e organizacionais do conhecimento existentes nesse fenômeno de dados (SILVA JÚNIOR; KARPINSKI; DUTRA, 2020). Essa contribuição pode ser identificada nas vivências científicas da área onde estão englobadas análises de tensões geradas entre os elementos humanos, informacionais e tecnológicos, notando-se uma similaridade com o contexto *big data*.

Alvarenga (2003) salienta que as práticas científicas da área de Ciência da Informação e Organização do Conhecimento relacionam-se com o crescimento exponencial da informação e seus desafios de acessibilidade ao longo das últimas décadas. Com efeito, os adventos tecnológicos, são uma espécie de estímulo para essa área repensar suas práticas e métodos relacionados a representação, armazenagem e recuperação de informações. A autora ainda ressalta que as constantes mudanças tecnológicas geram uma pressão positiva na Ciência da Informação, abrindo espaço para sua inovação e evolução, fatores estes importantes, dada a intensa ligação entre a área e os aspectos cognitivos e comunicacionais humanos. Portanto, as relações da Ciência da Informação e Organização do Conhecimento com o contexto *big data* têm diversos potenciais a serem explorados e suscitam os seguintes questionamentos desta pesquisa: a Organização do Conhecimento compreende e desenvolve sob quais formas o

fenômeno *big data*? Como a Organização do Conhecimento pode contribuir para uma sociedade cada vez mais pautada em dados?

Após as explicações acima, que ratificam a importância de estudos envolvendo *big data*, Ciência da Informação e Organização do Conhecimento, a presente pesquisa objetiva compreender o fenômeno *big data* e seus impactos na Organização do Conhecimento por meio de uma proposta classificatória da produção científica. Mais especificamente, pretende-se: analisar o contexto *big data* e seus impactos na Organização do Conhecimento, por meio de uma classificação da produção científica, fato que irá propiciar reflexões sobre o papel da área nas práticas envolvendo o contexto *big data*. A seguir, são apresentados os referenciais teóricos da pesquisa.

2 A organização do conhecimento e as complexidades do contexto *big data*

A comunidade de pesquisa em Organização do Conhecimento, conforme observam De Mauro, Greco e Grimaldi (2016) e Ibekwe-Sanjuan e Bowker (2017), teve dificuldades para integrar o crescente contexto tecnológico em suas metodologias. Com o surgimento do *big data*, a pressão para a organização e recuperação dos dados estabeleceu-se em diversos setores científicos e não foi menos sentida pela Organização do Conhecimento. Os autores ainda ressaltam que as padronizações flexíveis classificatórias são importantes para lidarem com o dinamismo e a diversidade impostos pela imensidão de dados, uma vez que a variedade e a multiplicidade desses dados e suas relações e interações impõem um dos principais desafios no tratamento dos dados existentes no contexto *big data*.

A ampla utilização de algoritmos para questões preditivas, segundo Ibekwe-Sanjuan e Bowker (2017), conquistaram um considerável desempenho no que se refere aos gostos e hábitos dos seres humanos. As recomendações de produtos, filmes e livros para os usuários, coletadas sem a percepção destes, obtém resultados satisfatórios que rivalizam com as ações humanas de construção de linguagens documentárias, nicho da Organização do Conhecimento. Os autores ainda observam que a euforia sobre o contexto *big*

data em diversos setores científicos e da sociedade estimulou discussões nas pesquisas sobre Organização do Conhecimento. A adaptabilidade deste campo ao ambiente mutável do contexto *big data* torna-se fundamental, uma vez que diversos setores científicos são impactados por esta realidade massiva de dados (BAGCHI, 2021a, 2021b; CLAVIER, 2019). Esta adaptação torna-se importante, tendo em vista que propicia as ferramentas necessárias para a disputa de recursos financeiros entre os campos científicos, impactando até mesmo em suas sobrevivências.

O contexto massivo de dados exige uma relação interdisciplinar sólida entre a Organização do Conhecimento e as pesquisas que envolvem a Ciência da Computação e a Inteligência Artificial. Os estudos em Biblioteconomia e Ciência da Informação precisam agregar questões relacionadas a “[...] epistemologia, filosofia, estatística e técnicas de análise de dados” (IBEKWE-SANJUAN; BOWKER, 2017, p.196, tradução nossa) e buscar a integração dessas áreas com esta nova realidade de dados, adaptando-se, portanto, ao novo paradigma que o *big data* representa para a ciência (IBEKWE-SANJUAN; BOWKER, 2017).

É importante observar que a Organização do Conhecimento possui recursos para colaborar com a interdisciplinaridade exigida neste contexto massivo de dados, aprimorando até mesmo seus métodos. Como exemplo, pode-se citar o tesouro e suas vinculações de dados que podem ampliar seu alcance científico se aliados a pesquisas relacionadas à Inteligência Artificial. Os tesouros têm potencial para “[...] gerenciar efetivamente *big data* através da combinação, alinhamento e criação de redes de vários sistemas de organização do conhecimento, a fim de acessar, indexar, organizar e recuperar *big data*” (SHIRI, 2014, p. 18, tradução nossa). Ao observar a questão da representação dos dados em um domínio científico, Borgman (2015a, p. 37, tradução nossa) destaca que “[...] métodos comuns de representação de dados, metadados, linguagens, tesouros, ontologias etc. facilitam a troca de dados dentro de um campo”. Portanto, Borgman confirma mais uma proximidade contextual do *big data* com a Biblioteconomia e a Ciência da Informação, porém com o desafio de

representar e minimizar barreiras de comunicação com outros campos científicos e suas diversidades conceituais.

Almeida e Crippa (2009) afirmam que a evolução tecnológica, pautada na informação, impacta diferentes dimensões da vida em sociedade, alterando as formas de “[...] pensar, de produzir, de consumir, de negociar, de administrar, de comunicar, de viver, de morrer e de fazer guerra” (ALMEIDA; CRIPPA, 2009, p. 110). Os autores ainda observam que esses impactos possibilitam inferir o surgimento de uma terceira era do capital, na qual os papéis estratégicos são exercidos pela ciência e pelo conhecimento que é mantido e viabilizado através das “escolhas classificatórias e de representação”. Almeida e Crippa (2009) ainda ressaltam que os avanços tecnológicos, dos quais o fenômeno *big data* faz parte, impõem uma reorganização da informação, agora não mais baseada em um espaço físico, mas sim no conteúdo informacional presente em diferentes suportes e espaços digitais, tornando, portanto, a representação da informação executada por profissionais qualificados, um item mais que fundamental na atualidade massiva de dados. Uma vez que, continuam os autores, “[...] a informação, em si mesma, não produz efeitos quando não encontra a forma de ser acessada pelos usuários” (ALMEIDA; CRIPPA, 2009, p. 125), sendo os elementos representativos informacionais primordiais para este acesso, ratifica-se a importância dos profissionais da informação no estabelecimento de diálogos entre os conhecimentos científicos e a sociedade. Como “[...] a informação, em suas múltiplas formas e concepções (científica, artística, mercadológica), tornou-se central na dinâmica social contemporânea” (ALMEIDA; CRIPPA, 2009, p. 126), esse diálogo se torna ainda mais necessário em um ambiente caótico de dados propiciado pelo *big data*. E o papel da Organização do Conhecimento como mediadora entre as fontes de conhecimento e a sociedade, interligando os diversos conhecimentos descentralizados pela disseminação informacional observada nas redes sociais com o conhecimento científico, surge como um de seus principais desafios na contemporaneidade.

Vale salientar que a Organização do Conhecimento e seus importantes métodos analíticos e representativos informacionais indicam a existência de experiências científicas para lidar com o contexto *big data* e suas complexidades

interpretativas. Borgman (2015b) alerta que é urgente evitar a confusão presente no fenômeno de dados entre aquisição (facilitada pelas tecnologias) e análise e interpretação (dificultadas pela imensidão de dados). A Organização do Conhecimento tem possibilidades de agir como uma mediadora entre as questões relacionadas às facilidades de aquisição de dados e sua complexa interpretação, contribuindo para a expansão semântica deste contexto massivo, principalmente no que tange aos dados científicos.

Nesse sentido, a proposta classificatória da produção científica presente nesta pesquisa representa uma das possibilidades de compreensão deste fenômeno de dados e seus impactos na Organização do Conhecimento. Uma vez que a referida classificação possui relações com o Tratamento Temático da Informação (TTI), tendo em vista a sua importância para a “[...] mediação entre a produção e o uso da informação, entre elas tecendo a mais sólida ponte: a que dá acesso ao conteúdo informacional” (GUIMARÃES, 2008, p. 78). Acredita-se que esta proposta classificatória auxilie na ampliação da compreensão sobre *big data* e Organização do Conhecimento, pautando-se no conteúdo informacional da produção científica relacionada a esses temas. Guimarães (2008) ainda ressalta a importância do objeto registrado e social da Organização do Conhecimento, já que sua “[...] organização e representação serão desenvolvidas de modo que, a partir dele, possa ser gerado novo conhecimento” (GUIMARÃES, 2008, p. 85). Logo, observa-se que as ligações da área com o contexto *big data* são nítidas, uma vez que a geração de novos conhecimentos e o acesso a eles em uma realidade massiva de dados representa um complexo desafio para a sociedade e a Organização do Conhecimento tem potencial para contribuir com esse cenário. Guimarães (2008) também salienta a dimensão estratégica da TTI para a extração de conteúdo relevante para determinado campo científico, propiciando quesitos avaliativos da ciência, utilizando-se, para isso, da “[...] identificação de elementos que permitem analisar um contexto científico, tais como objetos, pessoas, processos e relacionamentos” (GUIMARÃES, 2008, p. 93).

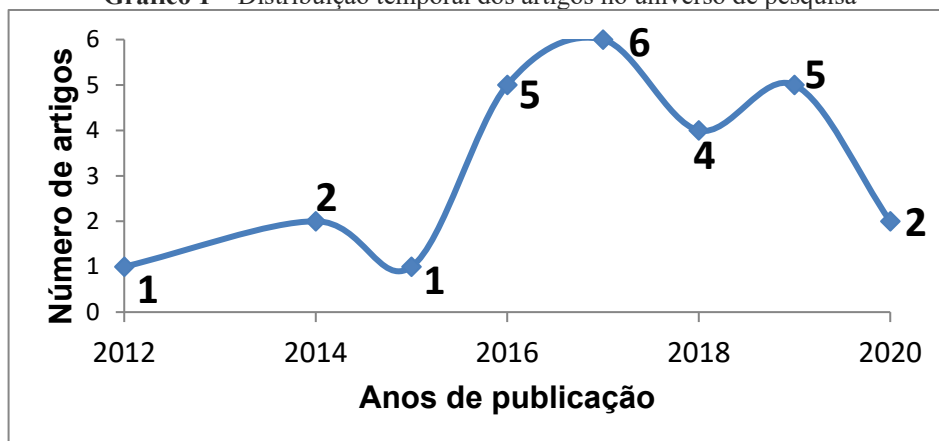
Observa-se, ainda que a Organização do Conhecimento e sua nuclearidade relacionada ao TTI propiciam elementos organizativos e

representativos capazes de extrair conteúdo informacional, gerando novos conhecimentos à sociedade. Esta perspectiva social, conforme observam Bräscher e Guimarães (2018), é bem alicerçada nas áreas mencionadas acima, uma vez que a complexidade observada em seu objeto científico está mais relacionada aos elementos humanos e sociais do que aos quesitos tecnológicos. Portanto, classificações, como a proposta nesta pesquisa, tornam-se necessárias, na medida em que propiciam reflexões sociais neste complexo ambiente massivo de dados ocasionado pelo fenômeno *big data*. Na próxima seção, as decisões metodológicas são explicitadas.

3 Procedimentos metodológicos

O universo desta pesquisa foi constituído a partir da coleta de artigos nas bases de dados Scopus, *Web of Science*, LISA e BRAPCI, por meio de buscas realizadas entre novembro e dezembro de 2020, com os seguintes termos: “*big data*” AND “*knowledge organization*” OR “organização do conhecimento” OR “*information organization*” OR “organização da informação”. Ressalta-se que as buscas foram realizadas sem delimitação de tempo nas bases de dados, almejando-se uma recuperação temporal mais abrangente. Quanto à distribuição temporal dos artigos do universo de pesquisa (gráfico 1), nota-se que a primeira publicação foi realizada somente em 2012, com o título “*Big data era*”, sob autoria de Shuichi Iwata (2012) e publicada pelo periódico “*Journal of Information Processing and Management*”. Esse artigo aborda ações proativas de descoberta de conhecimento orientadas por dados na gestão de crises, como, por exemplo, a ocorrência de terremotos. É possível afirmar que as discussões são recentes na área e com escassa produção ao longo dos anos, com um discreto acréscimo a partir de 2016.

Gráfico 1 – Distribuição temporal dos artigos no universo de pesquisa



Fonte: Elaborado pelos autores.

Com o auxílio do Microsoft Excel e Word, foi verificada a existência de sobreposição entre os artigos das diferentes bases. Após essa verificação, foi realizada a leitura dos títulos, resumos e palavras-chave dos artigos, objetivando-se confirmar se eles realmente possuíam correspondência com a busca realizada, reduzindo, assim, possíveis inconsistências. O total de artigos recuperados foi de 47 e, após a verificação da sobreposição, sobraram 30. Destaca-se que 4 artigos foram retirados do universo de pesquisa por não apresentarem discussões diretamente relacionadas ao tema *big data*. Portanto, o universo final de pesquisa foi composto por 26 artigos, conforme a tabela 1.

Quadro 1- Referências dos artigos selecionados para o universo de pesquisa

ARTIGOS DO UNIVERSO DE PESQUISA	
BENÍTEZ, A.S.; SÁNCHEZ-VIGIL; J.M. Tools and metrics in editorial marketing: from big data to artificial intelligence. Scire , Zaragoza, v. 26, n.1, p.35-46, 2020.	
ALHOORI, H.; SAMAKA, M.; FURUTA, R.; FOX, E.A. Anatomy of scholarly information behavior patterns in the wake of academic social media platforms, International Journal on Digital Libraries , Nova Iorque, v. 20, n. 4, p. 369-389, 2019.	
HJØRLAND, B. Data (with big data and database semantics). Knowledge Organization , Baden-Baden, v. 45, n. 8, p. 685-708, 2018.	
VICTORINO, M.; HOLANDA, M.T.; ISHIKAWA, E.; OLIVEIRA, E.C.; CHHETRI, S. Transforming open data to linked open data using ontologies for information organization in big data environments of the brazilian government: the brazilian database government open linked data DBgoldbr. Knowledge Organization , Baden-Baden, v. 45, n. 6, p. 443-466, 2018.	
CLEVERLEY, P.H.; MUIR, L.J. Using knowledge organization systems to automatically detect forward-looking sentiment in company reports to infer social phenomena. Knowledge Organization , Baden-Baden, v. 45, n. 2, p. 152-169, 2018.	
DU, W.; CHENG, X.; YANG, C.; SUN, J.; MA, J. Establishing interoperability among knowledge organization systems for research management: a social network approach. Scientometrics , Dordrecht, v. 112, n. 3, p. 1489-1506, 2017.	

MAKORI, E. O. Promoting innovation and application of internet of things in academic and research information organizations. Library Review , [s.l.], v. 66, n. 8-9, p. 655-678, 2017.
ZHANG, J.; SUN, Y.; YAO, C. Semantically linking events for massive scientific literature research. Electronic Library , Yorkshire, v. 35, n. 4, p. 724-744, 2017.
IBEKWE-SANJUAN, F.; BOWKER, G.C. Implications of big data for knowledge organization. Knowledge Organization , Yorkshire, v. 44, n. 3, p. 187-198, 2017.
MA, F.; CHEN, Y.; ZHAO, Y. Research on the organization of user needs information in the big data environment. Electronic Library , Yorkshire, v. 35, n. 1, p. 36-49, 2017.
ARAÚJO JÚNIOR, R. H.; SOUZA, R. T. B. Estudo do ecossistema de big data para conciliação das demandas de acesso, por meio da representação e organização da informação. Ciência da Informação , Brasília, v. 45, n. 3, p. 187-198, 2016.
XIA, C.; ZHANG, L. The application of linked data in Shanghai library's service of genealogy digital humanities. Library Journal , [s.l.], v. 35, n. 10, p. 26-34, 2016.
CERVONE, H. F. Organizational considerations initiating a big data and analytics implementation. Digital Library Perspectives , Yorkshire, v. 32, n. 3, p. 137-141, 2016.
KO, Y.M.; SONG, M.S.; LEE, S.J. Construction of the structural definition-based terminology ontology system and semantic search evaluation. Library Hi Tech , Yorkshire, v. 34, n. 4, p. 705-732, 2016.
SOERGEL, D. Unleashing the power of data through organization: structure and connections for meaning, learning and discovery. Knowledge Organization , Baden-Baden, v. 42, n. 6, p. 401-427, 2015.
SHIRI, Ali. Linked Data Meets Big Data: A knowledge organization systems perspective. Advances in Classification Research Online , Seattle, v. 24, p. 16-20, 2014.
LI, S.; HAO, Z.; DING, L.; XU, X.. Research on the application of information technology of Big Data in Chinese digital library. Library Management , Yorkshire, v. 40, n. 8-9, p. 518-553, 2019.
BARACHO, R. M. A.; CENDON, B. V.; MELO, M. O. T.; BARBOSA, C. R.; ALMEIDA, M. B. O caminhar da Ciência da Informação e o XV Encontro Nacional de Pesquisa em Ciência da Informação. Perspectivas em Gestão & Conhecimento , João Pessoa, v. 4, p. 198-211, 2014.
IWATA, S. Big Data era. Journal of Information Processing and Management , [s.l.], v. 55, n. 8, p. 543-551, 2012.
RIBEIRO, N. C.; OLIVEIRA, D. A.; ARAÚJO, R. F. Conjecturas da ciência aberta na contemporaneidade do <i>big data</i> . BIBLOS - Revista do Instituto de Ciências Humanas e da Informação , [s.l.], v. 33, n. 2, p. 163-179, 2019.
NAVARRO, F. P.; CONEGLIAN, C. S.; SANTAREM SEGUNDO, J. E. <i>Big data</i> no contexto de dados acadêmicos: o uso de machine learning na construção de sistema de organização do conhecimento. Informação & Tecnologia , João Pessoa, v. 5, n. 2, p. 181-200, 2018.
BARBOSA, E. C.; KOBASHI, N. Y. Extroversão e descoberta: visualização de dados no auxílio a buscas e recuperação de informações. Revista Brasileira de Biblioteconomia e Documentação , São Paulo, v. 13, p. 115-120, 2017.
SALES, L. F.; SAYÃO, L. F. Uma proposta de taxonomia para dados de pesquisa. Revista Conhecimento em Ação , [s.l.], v. 4, n. 1, p. 31-48, 2019.
ARAÚJO, C. A. V.; LIMA, G. A. B. O. Estruturação e representação semântica de <i>big data</i> no contexto de base de dados governamentais. Múltiplos Olhares em Ciência da Informação , Belo Horizonte, v. 9, n. 2, 2019.
FALSARELLA, O. M.; JANNUZZI, C. A. S. C. Inteligência organizacional e competitiva e <i>big data</i> : uma visão sistêmica para a gestão sustentável das organizações. Perspectivas em Ciência da Informação , Belo Horizonte, v. 25, n. 1, p. 179-204, 2020.
SANTANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. Informação & Informação , Londrina, v. 21, n. 2, p. 116-142, 2016.

Fonte: Elaborado pelos autores.

Os passos metodológicos seguintes foram relacionados à classificação do universo de pesquisa (Figura 1) por meio da leitura de títulos, resumos,

palavras-chave e textos dos artigos. Esta classificação ocorreu nas dimensões propostas nas publicações da ISKO-Brasil (epistemológica, aplicada e social, cultural e política), nas categorias de Ekbia *et al.* (2015) para literatura sobre *big data* (orientada ao produto, aos processos e à cognição) e em contextos criados a partir das informações dos próprios artigos do universo de pesquisa (tecnológico, acadêmico, empresarial e social-epistêmico-metodológico).

Figura 1 - Quadro de classificação do universo da pesquisa sobre *big data* e Organização do Conhecimento

Dimensões da publicação "Estudos Avançados em Organização do Conhecimento", sob responsabilidade da ISKO-Brasil, aplicadas à produção científica sobre Big data e Organização do Conhecimento	Categorização de Ekbia et al. (2015) para literatura sobre Big data	Contextos criados a partir das informações dos artigos do universo de pesquisa
<u>Dimensão epistemológica da organização do conhecimento:</u> bases conceituais, históricas e metodológicas da organização do conhecimento assim como seus diálogos interdisciplinares e sua produção científica (GUIMARÃES; DODEBEI, 2012, p.17). <u>Dimensão aplicada da organização do conhecimento:</u> modelos, formatos, instrumentos, produtos e estruturas em organização do conhecimento (GUIMARÃES; DODEBEI, 2012, p.17). <u>Dimensão social e política da organização do conhecimento:</u> formação e atuação profissional, ética, contextos, cultura e identidade em organização do conhecimento assim como as relações da organização do conhecimento com o desenvolvimento sustentável (GUIMARÃES; DODEBEI, 2012, p.17).	<u>Orientada para o produto,</u> com foco quantitativo no tamanho dos dados. <u>Orientada a processos,</u> com foco nos processos envolvidos na coleta, curadoria e uso de dados. <u>Orientada a cognição,</u> com foco na maneira como os seres humanos, com suas capacidades cognitivas específicas, podem se relacionar com os dados.	<u>Contexto tecnológico:</u> pesquisas que almejam a consecução de recursos tecnológicos para lidar com o Big data, propondo estratégias que propiciem serviços digitais capazes de usufruir dos benefícios existentes neste contexto massivo de dados. <u>Contexto acadêmico:</u> pesquisas relacionadas ao âmbito acadêmico, e focadas, portanto, em compreender sob quais formas os dados científicos podem colaborar para o desenvolvimento da Ciência. <u>Contexto empresarial:</u> pesquisas alinhadas à realidade de dados comerciais, na qual se observa o volume mais considerável do Big data, bem como os principais problemas relacionados às questões éticas e de privacidade. <u>Contexto social-epistêmico-metodológico:</u> pesquisas voltadas ao viés social dos efeitos do Big data na vida dos cidadãos. Propiciando ainda reflexões epistemológicas e metodológicas da Organização do Conhecimento no que se refere a este contexto social massivo de dados.

Fonte: Elaborado pelos autores.

Após a realização das etapas metodológicas acima explicitadas, os resultados foram apresentados por meio de figuras elaboradas com o auxílio do Microsoft Excel e do Infogram, com o objetivo de sintetizar e destacar os resultados obtidos na pesquisa, os quais estão apresentados na seção seguinte.

4 Resultados

Os resultados da pesquisa aqui apresentados contribuíram para ampliar a compreensão sobre o fenômeno *big data* e seus impactos na Organização do Conhecimento. Primeiramente, foram analisadas as palavras-chave presentes no universo de pesquisa para saber sob quais perspectivas o *big data* foi abordado.

Na sequência, a produção científica foi classificada conforme os seguintes elementos: dimensões epistemológica, aplicada, social, cultural e política; produção científica orientada ao produto, aos processos e à cognição; contextos tecnológico, acadêmico, empresarial, social-epistêmico-metodológico. Conforme detalhado a seguir.

4.1 Principais palavras-chave do universo de pesquisa

Foram extraídas 140 palavras-chave dos resumos dos artigos do universo de pesquisa, no entanto, somente 16 foram mencionadas pelo menos 2 duas vezes. As palavras-chave mais frequentes estão apresentadas na figura 2.

Figura 2 - Palavras-chave mais frequentes no universo de pesquisa



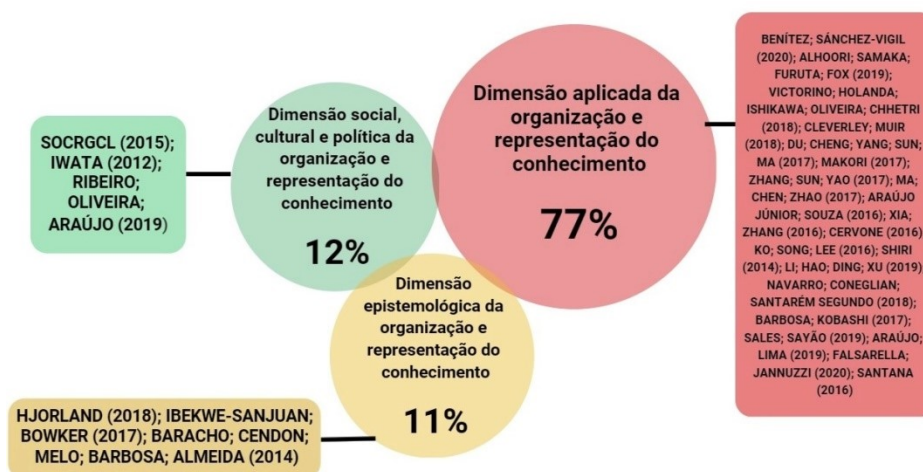
Fonte: Elaborado pelos autores.

É interessante observar a predominância de palavras-chave ligadas a sistemas de organização do conhecimento, dados abertos, recuperação da informação, inteligência artificial e classificação. Este cenário indica um direcionamento mais aplicado e voltado para o controle do atual contexto massivo de dados, fato natural e relevante, dado o intenso contexto tecnológico observado no fenômeno *big data*. A ausência de palavras-chave relacionadas às questões éticas e sociais entre as principais indica um caminho a ser explorado pelas pesquisas sobre *big data* e Organização do Conhecimento, ratificando sua importante função social no meio científico. A seguir, elencam-se as dimensões baseadas na ISKO-Brasil.

4.2 Dimensões: epistemológica, aplicada, social, cultural e política

As dimensões aqui apresentadas são oriundas da classificação das publicações presentes nos "Estudos Avançados em Organização do Conhecimento", sob responsabilidade da ISKO-Brasil. A classificação do universo de pesquisa com base nessas dimensões está representada na figura 3.

Figura 3 - Dimensões da ISKO-Brasil



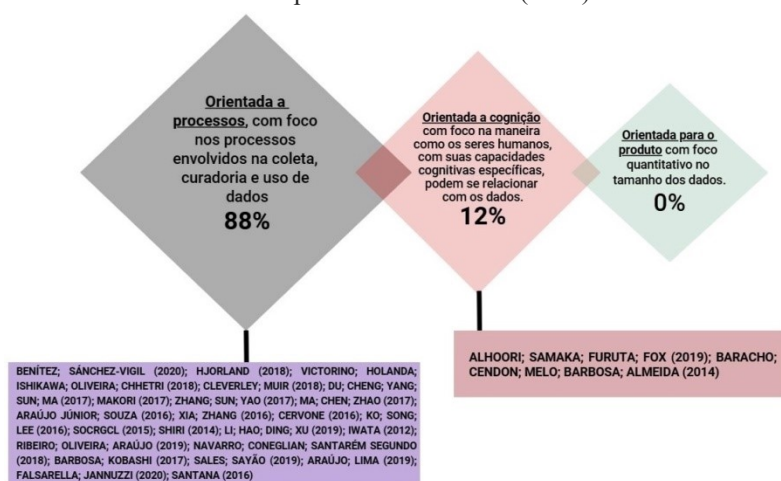
Fonte: Elaborado pelos autores.

A prevalência da dimensão aplicada (77%) ratifica a importância de aplicação em um contexto tecnológico, como o *big data*. Respostas práticas da Organização do Conhecimento perante o contexto massivo de dados são importantes para lidar com esta massa de possibilidades informacionais. É importante ressaltar a presença de pesquisas epistemológicas e sociais, culturais e políticas, para as questões vivenciadas em um contexto *big data*, já que a intensidade de dados afeta a vida dos cidadãos, interferindo em quesitos relacionados à privacidade e ao acesso igualitário dos dados para a melhoria de vida das pessoas. Além disso, as discussões epistemológicas são fundamentais para que a área desenvolva e garanta os melhores recursos para a sociedade usufruir positivamente esse contexto massivo de dados. A seguir, tem-se a categorização para literatura sobre *big data* proposta por Ekbja *et al.* (2015).

4.3 Produção científica orientada: produto, processos e cognição

A categorização a seguir foi elaborada por Ekbia *et al.* (2015) para literatura sobre *big data* e consiste na seguinte estrutura: orientada ao produto com foco quantitativo no tamanho dos dados; orientada aos processos, com foco nos processos envolvidos na coleta, curadoria e uso de dados; e, orientada à cognição com foco na maneira como os seres humanos, com suas capacidades cognitivas específicas, podem se relacionar com os dados. A seguir, a figura 4 representa uma síntese dos resultados.

Figura 4 - Categorização da produção científica sobre *big data* e Organização do Conhecimento a partir de Ekbia *et al.* (2015)



Fonte: Elaborado pelos autores.

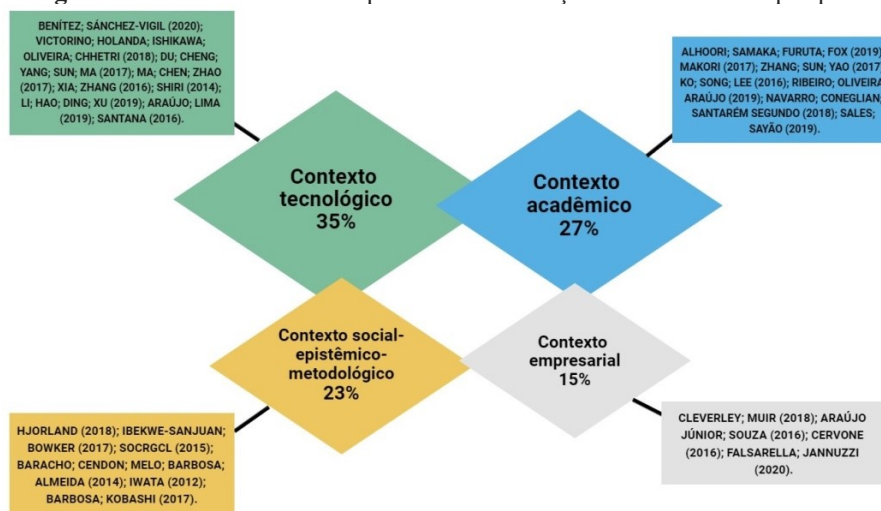
O primeiro fato em evidência nessa classificação é a ausência de artigos deste universo de pesquisa na literatura orientada para o produto com foco quantitativo no tamanho dos dados. Espera-se que este foco tenha mais abrangência em pesquisas com um viés mais voltado para a Ciência da Computação e as questões de armazenamento e transferência da intensidade de dados observada no contexto *big data*. O predomínio da categoria orientada a processos, com foco na coleta, curadoria e uso de dados (88%), confirma os princípios organizativos existentes nas origens da Ciência da Informação e Organização do Conhecimento. O foco nos elementos organizativos para ampliar as possibilidades informacionais em um contexto *big data* é relevante para que as áreas contribuam com o fenômeno de dados e com a proposição de uma diretriz social e voltada para o aspecto humano desse contexto. Há a

presença de pesquisas com orientação à cognição com foco na maneira como os seres humanos, com suas capacidades cognitivas específicas, podem se relacionar com os dados. Este direcionamento preconiza os cidadãos e suas relações com os dados, campo importante a ser explorado pela Organização do Conhecimento, na medida em que ratifica a função social da área e abre possibilidades de pesquisas que superem os quesitos tecnológicos, que também são importantes, mas não únicos no complexo contexto *big data*. A seguir, são apresentados os contextos criados a partir do próprio universo de pesquisa, onde as especificidades deste universo são mais amplamente discutidas.

4.4 Contextos: tecnológico, acadêmico, empresarial e social-epistêmico-metodológico

Nesta parte da pesquisa, são apresentados os contextos criados a partir das informações dos próprios artigos do universo de pesquisa, tendo por objetivo ampliar a compreensão das práticas científicas a respeito do *big data* e da Organização do Conhecimento. Abaixo, na figura 5, apresenta-se a distribuição do universo de pesquisa na proposta classificatória.

Figura 5 - Contextos criados a partir das informações do universo de pesquisa



Fonte: Elaborado pelos autores.

O contexto tecnológico concentra o maior número de trabalhos (35%), cenário este motivado pelas intensas tecnologias envolvidas no fenômeno *big*

data. As repostas da Organização do Conhecimento para lidar com esta massa de dados estão interligadas com o desenvolvimento de tecnologias. Da mesma forma, o estabelecimento de proposições práticas para organizar estes dados não estruturados e expandir semanticamente as possibilidades informacionais deste ambiente massivo de dados. Em seguida, com 27%, apresenta-se o contexto acadêmico. O resultado está relacionado às diversas pesquisas sobre *big data* que foram desenvolvidas na área, onde são destacados os “dados científicos” produzidos nos ambientes acadêmicos. O contexto social-epistêmico-metodológico aparece com 23%. Este contexto é mais amplo e diverso, concentrando artigos voltados ao viés social, quesito relevante para os impactos do *big data* no cotidiano da sociedade. Agrega também os estudos que refletem traços epistemológicos e metodológicos da Organização do Conhecimento e suas contribuições para lidar com essa intensidade de dados. O contexto social-epistêmico-metodológico é importante para que as contribuições da área sejam consistentemente fundamentadas e impactem positivamente a sociedade. Por último, com 15%, apresenta-se o contexto empresarial. Este contexto representa uma aproximação da Organização do Conhecimento com uma realidade de dados comerciais, na qual o maior volume de dados do *big data* circula e os impactos nas questões de privacidade e ética são mais evidentes. Ressalta-se que esta aproximação pode ratificar a função social da Organização do Conhecimento, já que uma estruturação destes dados, pautada na ética, pode representar uma contribuição necessária para este ambiente massivo de dados, nem sempre controlado. A seguir, as características do universo de pesquisa de cada um destes contextos são especificamente explanadas.

4.4.1 Contexto tecnológico

Este contexto concentra interessantes pesquisas que almejam a consecução de recursos tecnológicos para lidar com o *big data*, propondo estratégias que propiciem serviços digitais capazes de usufruir dos benefícios existentes neste contexto massivo de dados. A figura 6 sintetiza as principais características do contexto tecnológico.

Figura 6 - Temas presentes no contexto tecnológico



Fonte: Elaborado pelos autores.

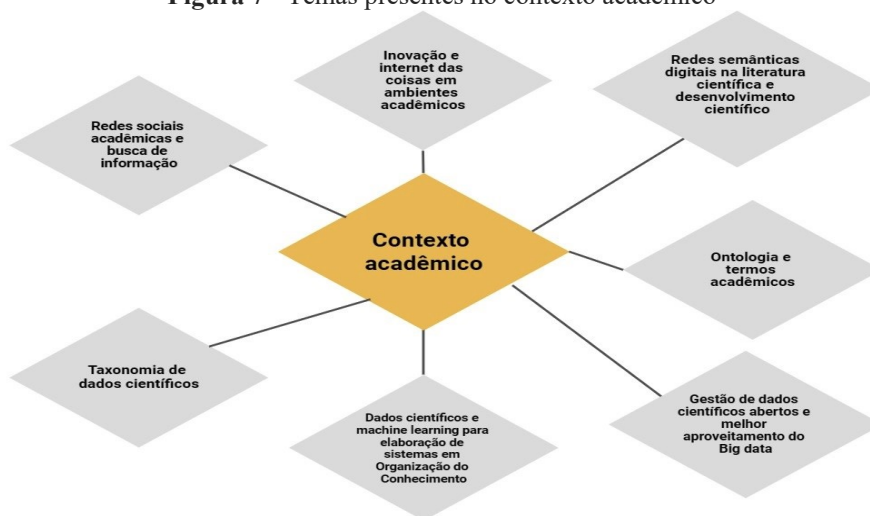
As contribuições das pesquisas presentes neste contexto são explicitadas a seguir: Benítez e Sánchez-Vigil (2020) (Proposição de ferramentas e métricas relacionadas a marketing, *big data* e Inteligência artificial para conhecer hábitos de leitura e realizar recomendações de livros em uma estrutura consistente de representação e organização do conhecimento); Victorino, Holanda, Ishikawa, Oliveira e Chhetri (2018) (Propostas de ferramenta e modelo conceitual para a construção de ontologia baseada em dados abertos governamentais como forma de apoio para formulação e gerenciamento de políticas públicas); Du, Cheng, Yang, Sun e Ma (2017) (Abordagem de redes sociais para mapeamentos de atividades científicas e sociais com base em Sistemas de Organização do Conhecimento para gerenciamento de pesquisas não focado em análise semântica e metadados); Ma, Chen e Zhao (2017) (Estabelecimento de um modelo conceitual, focado na necessidade dos usuários e sua privacidade, para aprimoramento da organização da informação em ambiente *big data*); Xia e Zhang (2016) (Linked Data como recurso para práticas em Humanidades digitais e Organização do Conhecimento, com base em reorganização do conhecimento e descoberta de novos conhecimentos); Shiri (2014) (Aplicativos de dados baseados em Sistemas de Organização do Conhecimento para dados

massivos, como novas ferramentas para a análise, organização, representação, visualização e acesso a estes dados massivos); Li, Hao, Ding e Xu (2019) (Aplicação de tecnologias relacionadas à *big data*, Inteligência artificial e Block chain em bibliotecas digitais chinesas, como formas de aprimoramento dos serviços digitais prestados); Araújo e Lima (2017) (Identificação de quais tecnologias da Web Semântica podem contribuir com uma padronização de estruturação e representação de dados abertos governamentais no contexto de *big data*); e Santana (2016) (Otimização do ciclo de vida de dados, com foco em propiciar pontes entre os usuários e suas necessidades de dados). As pesquisas classificadas neste contexto estão voltadas para a consecução de recursos capazes de usufruir dos benefícios que o *big data* pode propiciar e suas complexas possibilidades informacionais, considerando, assim, este contexto massivo de dados como uma espécie de provedor para a realização de novos estudos informacionais. A seguir, as especificidades do contexto acadêmico são apresentadas.

4.4.2 Contexto acadêmico

Neste contexto são apresentadas pesquisas relacionadas ao âmbito acadêmico e focadas, portanto, em compreender sob quais formas os dados científicos podem colaborar para o desenvolvimento da ciência. A figura 7 apresenta os principais temas do contexto acadêmico.

Figura 7 - Temas presentes no contexto acadêmico



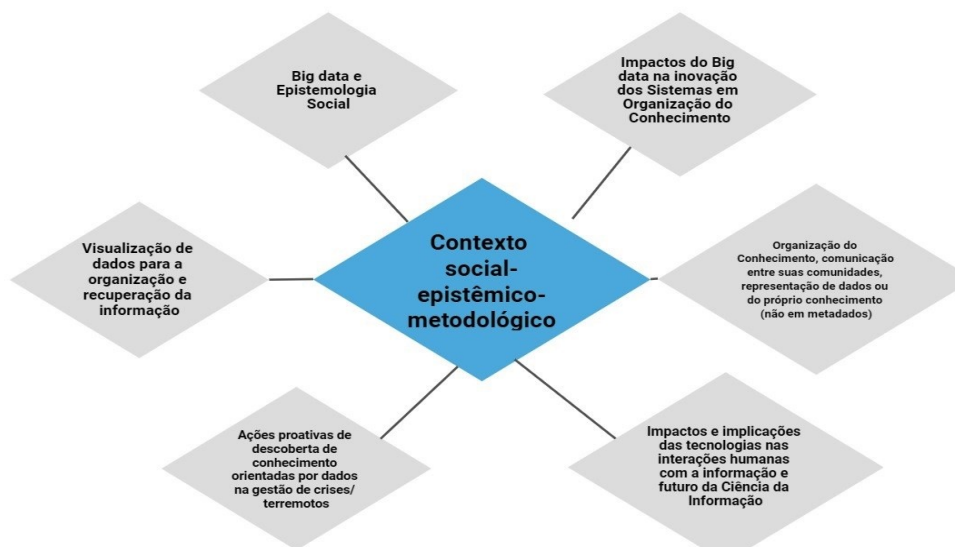
Fonte: Elaborado pelos autores.

As características das pesquisas presentes no contexto acadêmico são apresentadas a seguir: Alhoori, Samaka, Furuta e Fox (2019) (Redes sociais acadêmicas e busca de informação no meio digital, compreendendo o *big data* como possibilidade de descobertas científicas pelos pesquisadores); Makori (2017) (Inovação e internet das coisas em ambientes acadêmicos como promotoras do acesso a conteúdos científicos); Zhang, Sun e Yao (2017) (Redes semânticas digitais na literatura científica para promover o desenvolvimento científico por meio de um modelo de rede de organização da informação); Ko, Song e Lee (2016) (Construção de um sistema de ontologia terminológica voltado a termos acadêmicos, considerando recuperação e mineiração de *big data* por meio de pesquisas semânticas); Ribeiro, Oliveira e Araújo (2019) (Gestão de dados científicos abertos e melhor aproveitamento do *big data* para acesso ao conhecimento científico); Navarro, Coneglian e Santarém Segundo (2018) (Dados científicos e machine learning para elaboração de sistemas em Organização do Conhecimento); e Sales e Sayão (2019) (Proposta de uma taxonomia de dados científicos para a descoberta de novos conhecimentos). As pesquisas alocadas neste contexto abordam formas de ampliar as descobertas científicas por meio da organização dos dados científicos. Estas pesquisas, portanto, tendem a considerar que a imensidão de dados presente no *big data* oculta conhecimentos científicos, sendo necessária a obtenção de recursos para viabilizar o acesso a estes conhecimentos. A seguir, características do contexto social-epistêmico-metodológico são elencadas.

4.4.3 Contexto social-epistêmico-metodológico

Este contexto é o mais diversificado, com pesquisas voltadas ao viés social dos efeitos do *big data* na vida dos cidadãos. Propicia ainda reflexões epistemológicas e metodológicas da Organização do Conhecimento no que se refere ao contexto massivo de dados e explora as possibilidades de contribuições da área para o desenvolvimento informacional da sociedade. A figura 8 demonstra as principais temáticas do contexto social-epistêmico-metodológico.

Figura 8 - Temas presentes no contexto social-epistêmico-metodológico



Fonte: Elaborado pelos autores com o auxílio do Infogram.

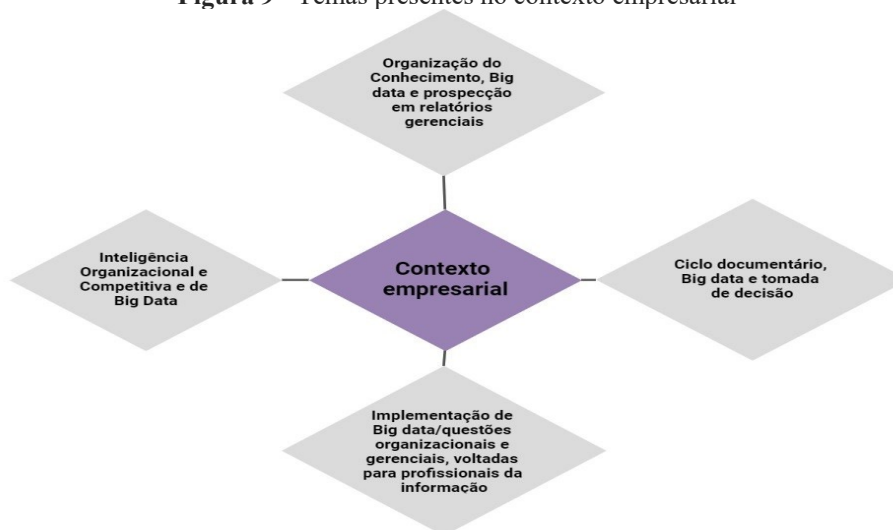
As contribuições das pesquisas classificadas neste contexto são explicitadas a seguir: Hjørland (2018) (Compreensão teórica do fenômeno *big data* sob a perspectiva da Epistemologia Social, considerando a natureza e intencionalidade existentes na produção massiva dos dados); Ibekwe-Sanjuan e Bowker (2017) (Impactos e implicações do *big data* na inovação dos Sistemas em Organização do Conhecimento); Soergel (2015) (Discute a Organização do Conhecimento e a comunicação entre suas comunidades, considerando os impactos da tecnologia na representação de dados ou do próprio conhecimento); Baracho, Cendon, Melo, Barbosa e Almeida (2014) (Impactos e implicações das tecnologias nas interações humanas com a informação e futuro da Ciência da Informação); Iwata (2012) (Ações proativas de descoberta de conhecimento orientadas por dados na gestão de crises, como, por exemplo, a ocorrência de terremotos); e Barbosa e Kobashi (2017) (Visualização de dados em repositórios digitais como suporte para a organização e recuperação da informação). As pesquisas presentes neste contexto consideram a importância do fator social na produção e acesso do contexto massivo de dados, bem como refletem sobre os notáveis impactos informacionais deste fenômeno de dados na vida em sociedade. As pesquisas também procuram discutir o papel relevante da Ciência da Informação e Organização do Conhecimento na descoberta de conhecimento

no universo do *big data*. A seguir, são apresentadas especificidades do contexto empresarial.

4.4.4 Contexto empresarial

Este contexto está alinhado à realidade de dados comerciais, na qual se observa o volume mais considerável do *big data*, bem como os principais problemas relacionados às questões éticas e de privacidade. A figura 9 sintetiza as principais características do contexto empresarial.

Figura 9 - Temas presentes no contexto empresarial



Fonte: Elaborado pelos autores com o auxílio do Infogram.

As características das pesquisas presentes no contexto empresarial são apresentadas a seguir: Cleverley e Muir (2018) (Organização do Conhecimento, *big data* e prospecção em relatórios gerenciais para detecção de fenômenos sociais); Araújo Júnior e Souza (2016) (Análise do ecossistema do *big data*, abrangendo as etapas do ciclo documentário, para a tomada de decisão na administração pública brasileira); Cervone (2016) (Implementação de *big data* para questões organizacionais e gerenciais de unidades de informação, objetivando usufruir dos benefícios gerados por este contexto massivo de dados); e Falsarella e Jannuzzi (2020) (Inteligência organizacional e competitiva e de *big data* para a gestão sustentável das organizações por meio do monitoramento de dados). As pesquisas deste contexto estão focadas em

propiciar estratégias para explorar as vantagens do *big data* para a consecução de um desenvolvimento organizacional. Mas, também, existe potencial para o desenvolvimento da função social da Organização do Conhecimento na medida em que os principais problemas éticos residem nos dados comerciais. A estruturação desses dados, pautada em elementos de ética, pode representar contribuições importantes da área para este ambiente intenso de dados e com necessidades organizativas urgentes. Os resultados da pesquisa são discutidos a seguir.

5 Discussão

Os resultados obtidos nesta pesquisa permitiram ampliar a compreensão do fenômeno *big data* e suas relações com a Organização do Conhecimento por meio de uma proposta classificatória da produção científica, bem como da análise das principais palavras-chave dessa produção. A relação das palavras-chave com os sistemas de organização do conhecimento, dados abertos, recuperação da informação, inteligência artificial e classificação indica esforços da área voltados para um controle organizacional do contexto massivo de dados, almejando, portanto, a ampliação das possibilidades informacionais destes dados. Tendo em vista as complexidades na gestão de dados e suas implicações na privacidade dos cidadãos, outro destaque é a ausência de palavras-chave relacionadas às questões éticas e sociais. Explorar pesquisas nesses temas em Organização do Conhecimento pode contribuir para a função social da área e trazer benefícios para a sociedade na estruturação de dados com base em valores éticos.

Quando se analisa as dimensões - aplicada, epistemológica e social, cultural e política -, a elevada concentração na primeira dimensão indica que a Organização do Conhecimento propicia respostas práticas para lidar com esta massa de dados no sentido de ampliar as condições informacionais no contexto de dados. Embora tal praticidade seja relevante ao tratar de um fenômeno de dados intensamente ligado a quesitos tecnológicos, conforme já observado por Shiri (2014) e Borgman (2015a), onde os autores enaltecem a consecução de tesouros, linguagens documentárias e ontologias como recursos para lidar com o

big data, é importante destacar que a realização de pesquisas abordando questões epistemológicas, sociais, culturais e políticas também deve ser estimulada, cenário já evidenciado por Ibekwe-Sanjuan e Bowker (2017). Essa seria uma maneira de contribuir com recursos bem fundamentados para que a sociedade possa usufruir positivamente do contexto *big data*.

Outra maneira seria através da aplicação de parâmetros pautados no acesso ético e igualitário a esses dados, objetivando a melhoria de vida dos cidadãos no que se refere à ampliação do acesso informacional. Esse foco no acesso é reafirmado por Almeida e Crippa (2009) e Guimarães (2008) na medida em que defendem a urgência do diálogo entre os conhecimentos científicos e a sociedade como elementos fundamentais para que os dados possibilitem a consecução de informações capazes de produzir efeitos positivos na vida das pessoas, pautando-se em análises do conteúdo informacional que podem ser oriundas do Tratamento Temático da Informação (TTI).

A classificação da literatura sobre *big data* baseada em Ekbja *et al.* (2015), utilizada nesta pesquisa, revelou o predomínio da categoria orientada a processos, com foco nos processos envolvidos em coleta, curadoria e uso de dados. Tal resultado demonstra a importância dos princípios organizativos e informacionais presentes na Organização do Conhecimento. Ressalta-se que esses princípios devem possuir uma diretriz social direcionada para o fator humano no contexto massivo de dados. Por isso, destaca-se também a existência de pesquisas voltadas à cognição humana específica nas relações com os dados, o que coloca em evidência o protagonismo social que a Organização do Conhecimento deve priorizar, isto é, para além das questões tecnológicas.

Quanto aos contextos criados a partir dos artigos do universo desta pesquisa (tecnológico, acadêmico, social-epistêmico-metodológico e empresarial), o contexto tecnológico prevalece pelas interligações entre *big data* e desenvolvimentos tecnológicos para a estruturação informacional destes dados pela Organização do Conhecimento, pautando-se, portanto, em bases semânticas para a expansão das capacidades informacionais do contexto massivo de dados. Um ponto que deve ser ressaltado é que as pesquisas classificadas neste contexto poderiam explorar a obtenção de recursos para o alcance de benefícios

neste fenômeno de dados, compreendendo este cenário como um estímulo para a realização de novos estudos informacionais.

As pesquisas do contexto acadêmico reconhecem a intensidade de dados existentes no *big data* como forma de ampliação das descobertas científicas por meio da organização dos dados científicos. Essas descobertas tornam-se necessárias para viabilizar o acesso ao conhecimento que possivelmente está perdido nessa massa de dados. Já o contexto social-epistêmico-metodológico - em sua diversidade de pesquisas - considera o aspecto social da Organização do Conhecimento como elemento fundamental na produção e no acesso aos dados.

Quanto ao contexto empresarial, nota-se que as pesquisas têm por foco compreender o *big data* como um recurso para aprimorar o desenvolvimento das organizações, baseado em informações a partir de dados estruturados. É importante ressaltar que os dados comerciais do *big data* são os mais polêmicos no que se refere aos quesitos de privacidade e ética. Logo, representam mais uma oportunidade para que o elemento social da Organização do Conhecimento seja reafirmado e valorizado por meio de pesquisas que abarquem tanto dados científicos quanto comerciais, a difusão de fronteiras entre estes dados e sua necessidade organizacional. A seguir, são elencadas as considerações finais.

6 Considerações finais

O fenômeno *big data* e seus impactos nos campos científico e social alteram significativamente as práticas vivenciadas por seus membros, impondo novos desafios organizacionais, éticos e informacionais. A Organização do Conhecimento e seus princípios organizativos e informacionais têm potencial para contribuir com o desenvolvimento da capacidade informacional advinda do contexto massivo de dados. Ao realizar uma proposta classificatória da produção científica sobre *big data*, bem como uma análise das principais palavras-chave, considera-se que os objetivos da pesquisa foram alcançados e que foi possível ampliar a compreensão das relações entre *big data* e Organização do Conhecimento.

O perfil predominantemente aplicado, tecnológico - orientado a processos relacionados à coleta, curadoria e utilização de dados observado no

universo desta pesquisa - indica que a Organização do Conhecimento lida objetivamente com o fenômeno do *big data*, propondo estratégias e recursos para a ampliação das possibilidades informacionais neste contexto. Embora também existam, neste universo, pesquisas relacionadas às questões sociais, cognitivas, epistemológicas e metodológicas, destaca-se a importância da realização de mais investigações nesses âmbitos, bem como a elaboração de estudos voltados para a ética neste cenário. A estruturação destes dados, pautada em elementos sociais, éticos e igualitários, sob a perspectiva da Organização do Conhecimento, pode resultar em notáveis contribuições desta área para o potencial desenvolvimento informacional advindo do *big data*. Portanto, as oportunidades de pesquisas no contexto *big data* são tão amplas quanto o crescimento exponencial desse segmento.

Referências

- ALHOORI, H.; SAMAKA, M.; FURUTA, R.; FOX, E. A. Anatomy of scholarly information behavior patterns in the wake of academic social media platforms, **International Journal on Digital Libraries**, Nova Iorque, v. 20, n. 4, p. 369-389, 2019.
- ALMEIDA, M. A.; CRIPPA, G. De Bacon à internet: considerações sobre a organização do conhecimento e a constituição da Ciência da Informação. **Ponto de Acesso**, Salvador, v. 3, n. 2, p. 109-131, 2009.
- ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 8, n. 15, p. 18-40, 2003.
- ARAÚJO, C. A. V.; LIMA, G. A. B. O. Estruturação e representação semântica de *big data* no contexto de base de dados governamentais. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, 2019.
- ARAÚJO JÚNIOR, R. H.; SOUZA, R. T. B. Estudo do ecossistema de *big data* para conciliação das demandas de acesso, por meio da representação e organização da informação. **Ciência da Informação**, Brasília, v. 45, n. 3, p. 187-198, 2016.
- ARBOIT, A. E.; BUFREM, L. S.; FREITAS, J. L. Configuração epistemológica da Ciência da Informação na literatura periódica brasileira por meio de análise de citações (1972-2008). **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 1, p. 18-43, 2010.

- BAGCHI, M. A large scale, knowledge intensive domain development methodology. **Knowledge Organization**, Baden-Baden, v. 48, n. 1, p. 8-23, 2021a.
- BAGCHI, M. Towards knowledge organization ecosystems. **arXiv:2105.10923**, 2021b. Disponível em: <https://arxiv.org/pdf/2105.10923.pdf>. Acesso em: 21 jul. 2021.
- BARBOSA, E. C.; KOBASHI, N. Y. Extroversão e descoberta: visualização de dados no auxílio a buscas e recuperação de informações. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 13, p. 115-120, 2017.
- BARACHO, R. M. A.; CENDON, B. V.; MELO, M. O. T.; BARBOSA, C. R.; ALMEIDA, M. B. O caminhar da Ciência da Informação e o XV Encontro Nacional de Pesquisa em Ciência da Informação. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 4, p. 198-211, 2014.
- BENÍTEZ, A. S., SÁNCHEZ-VIGIL; J. M. Tools and metrics in editorial marketing: from big data to artificial intelligence. **Scire**, Zaragoza, v. 26, n. 1, p. 35-46, 2020.
- BORGMAN, C. Data scholarship. In: BORGMAN, C. **Big data, little data, no data: scholarship in the networked world**. Cambridge: MIT Press, 2015a. p. 31-53.
- BORGMAN, C. Data Scholarship in the Social Sciences. In: BORGMAN, C. **Big data, little data, no data: scholarship in the networked world**. Cambridge: MIT Press, 2015b. p. 125-160.
- BRÄSCHER, M.; GUIMARÃES, J. A. C. Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. **Liinc em revista**, Rio de Janeiro, v. 14, n. 2, p. 241-258, 2018.
- CERVONE, H. F. Organizational considerations initiating a big data and analytics implementation. **Digital Library Perspectives**, Yorkshire, v. 32, n. 3, p. 137-141, 2016.
- CLAVIER, V. Knowledge organization, data and algorithms: the new era of visual representations. **Knowledge Organization**, Baden-Baden, v. 46, n. 8, p. 615-621, 2019.
- CLEVERLEY, P. H.; MUIR, L. J. Using knowledge organization systems to automatically detect forward-looking sentiment in company reports to infer social phenomena. **Knowledge Organization**, Baden-Baden, v. 45, n. 2, p. 152-169, 2018.

DE MAURO, A.; GRECO, M.; GRIMALDI, M. A formal definition of big data based on its essential features. **Library Review**, [s.l.], v. 65, n. 3, p. 122-135, 2016.

DU, W.; CHENG, X.; YANG, C.; SUN, J.; MA, J. Establishing interoperability among knowledge organization systems for research management: a social network approach. **Scientometrics**, Dordrecht, v. 112, n. 3, p. 1489-1506, 2017.

EKBIA, H. *et al.* Big data, bigger dilemmas: a critical review. **Journal of the Association for Information Science and Technology**, Hoboken, v. 66, p. 1523-1545, 2015.

FALSARELLA, O. M.; JANNUZZI, C. A. S. C. Inteligência organizacional e competitiva e big data: uma visão sistêmica para a gestão sustentável das organizações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 25, n. 1, p. 179-204, 2020.

GLÄNZEL, W. **Bibliometrics as a research field: a course on theory and application of bibliometric indicators**. Bélgica: Magyar Tudományos Akadémia, 2003.

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 1, n. 1, p. 77-99, 2008.

HJØRLAND, B. Data (with big data and database semantics). **Knowledge Organization**, Baden-Baden, v. 45, n. 8, p. 685-708, 2018.

IBEKWE-SANJUAN, F.; BOWKER, G. C. Implications of big data for Knowledge Organization. **Knowledge Organization**, Baden-Baden, v. 44, n. 3, p. 187-198, 2017.

IWATA, S. Big data era. **Journal of Information Processing and Management**, [s.l.], v. 55, n. 8, p. 543-551, 2012.

KO, Y.M.; SONG, M.S.; LEE, S.J. Construction of the structural definition-based terminology ontology system and semantic search evaluation. **Library Hi Tech**, Yorkshire, v. 34, n. 4, p. 705-732, 2016.

LI, S.; HAO, Z.; DING, L.; XU, X. Research on the application of information technology of big data in Chinese digital library. **Library Management**, Yorkshire, v. 40, n. 8-9, p. 518-553, 2019.

MA, F.; CHEN, Y.; ZHAO, Y. Research on the organization of user needs information in the big data environment. **Electronic Library**, Yorkshire, v. 35, n. 1, p. 36-49, 2017.

MACIAS-CHAPULA, C. A. O papel da informetria e da cientometria e sua perspectiva nacional e internacional. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 134-140, 1998.

MAKORI, E. O. Promoting innovation and application of internet of things in academic and research information organizations. **Library Review**, [s.l.], v. 66, n. 8-9, p. 655-678, 2017.

MUELLER, S. P. M. Literatura científica, comunicação científica e ciência da informação. In: TOUTAIN, Lídia Maria Batista Brandão (org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007. p. 125-144.

NAVARRO, F. P.; CONEGLIAN, C. S.; SANTARÉM SEGUNDO, J. E. *Big data* no contexto de dados acadêmicos: o uso de *machine learning* na construção de sistema de organização do conhecimento. **Informação & Tecnologia**, João Pessoa, v. 5, n. 2, p. 181-200, 2018.

RENAULT, L. V. Paradigmas e modelos: proposta de análise epistemológica para a Ciência da Informação. **Informação & Sociedade**, João Pessoa, v. 17, n. 2, p. 53-60, 2007.

RIBEIRO, N. C.; OLIVEIRA, D. A.; ARAÚJO, R. F. de. Conjecturas da ciência aberta na contemporaneidade do *big data*. **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, [s.l.], v. 33, n. 2, p. 163-179, 2019.

SALES, L. F.; SAYÃO, L. F. Uma proposta de taxonomia para dados de pesquisa. **Revista Conhecimento em Ação**, [s.l.], v. 4, n. 1, p. 31-48, 2019.

SANTANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116-142, 2016.

SHIRI, A. Linked data meets big data: a knowledge organization systems perspective. **Advances in Classification Research Online**, Seattle, v. 24, p. 16-20, 2014.

SILVA JÚNIOR, E. M.; KARPINSKI, C.; DUTRA, M. L. Conhecimento científico no contexto *big data*: reflexões a partir da epistemologia de Popper. **Brazilian Journal of Information Science: Research trends**, Marília, v. 14, n. 4, 2020.

SOERGEL, D. Unleashing the power of data through organization: structure and connections for meaning, learning and discovery. **Knowledge Organization**, Baden-Baden, v. 42, n. 6, p. 401-427, 2015.

VICTORINO, M.; HOLANDA, M.T.; ISHIKAWA, E.; OLIVEIRA, E.C.; CHHETRI, S. Transforming open data to linked open data using ontologies for information organization in big data environments of the brazilian government:

the brazilian database government open linked data DBgoldbr. **Knowledge Organization**, Baden-Baden, v. 45, n. 6, p. 443-466, 2018.

XIA, C.; ZHANG, L. The application of linked data in Shanghai library's service of genealogy digital humanities. **Library Journal**, [s.l.], v. 35, n. 10, p. 26-34, 2016.

ZHANG, J.; SUN, Y.; YAO, C. Semantically linking events for massive scientific literature research. **Electronic Library**, Yorkshire, v. 35, n. 4, p. 724-744, 2017.

Big data and Knowledge Organization: initial reflections from a classification proposal for scientific production

Abstract: Considering the impacts and the complexities that exist in the big data context and observed in the Knowledge Organization, reflections are necessary that highlight the possible contributions of this area to the informational development of society. With this in mind, the present research seeks to understand the big data phenomenon and its impacts on the Knowledge Organization through a classificatory proposal of scientific production. For this purpose, articles related to big data and Knowledge Organization were collected in the Scopus, Web of Science, Library and Information Science Abstracts and Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação databases. The articles were classified by reading the titles, abstracts, keywords and texts of the articles. The results demonstrate the predominance of a more applied, technological profile and oriented to processes related to the collection, curation and use of data. It concludes that it is necessary to expand research related to social, cognitive, epistemological and methodological issues, as well as the elaboration of more studies focused on ethics in this intense data scenario. It contributes to understand and expand the research scenarios about big data in the Knowledge Organization.

Keywords: Big data; Knowledge Organization; Scientific production; Scientific production; Information Science.

Recebido: 03/05/2021

Aceito: 27/07/2021

Declaração de autoria

Concepção e elaboração do estudo: Fabio Orsi Meschini, Marivalde Moacir Francelin.

Coleta de dados: Fabio Orsi Meschini.

Análise e interpretação de dados: Fabio Orsi Meschini, Marivalde Moacir Francelin.

Redação: Fabio Orsi Meschini, Marivalde Moacir Francelin.

Revisão crítica do manuscrito: Marivalde Moacir Francelin.

Como citar:

MESCHINI, Fabio Orsi; FRANCELIN, Marivalde Moacir. Big data e organização do conhecimento: reflexões iniciais a partir de uma proposta classificatória da produção científica. **Em Questão**, Porto Alegre, v. 28; n. 1, p. 35-63, 2022. <https://doi.org/10.19132/1808-5245281.35-63>

