

BIG DATA EM PERIÓDICOS DA ÁREA DE CIÊNCIA DA INFORMAÇÃO: uma abordagem voltada para a ciência aberta e a ciência de dados

Nivaldo Calixto Ribeiro

Doutorando em Gestão e Organização do Conhecimento
Escola de Ciência da Informação, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil

 <https://orcid.org/0000-0003-0650-0121> E-mail: nivaldo@biblioteca.ufla.br

Submetido em: 04-09-2020	Reapresentado em: 07-12-2020	Aceito em: 07-12-2020
--------------------------	------------------------------	-----------------------

RESUMO

Objetivou-se neste artigo analisar como têm sido abordadas as temáticas: *big data* e a ciência aberta em periódicos da área de Ciência da Informação. Trata-se de uma pesquisa bibliográfica baseada em artigos de periódicos científicos sobre os temas *big data* e a ciência aberta. Foram analisados artigos de 21 periódicos especializados na área de Ciência da Informação, sendo 3 com classificação Qualis A1, 3 como A2 e 15 títulos como B1, quadriênio 2013-2016. A coleta de dados foi realizada por meio de pesquisa na opção de busca das páginas web oficiais dos periódicos selecionados, para período de 2000 a 2019. Foi utilizado como estratégia de busca o termo *big data* e foram recuperados 88 artigos. Do total de artigos recuperados sobre o *big data*, 6 o tratam no contexto da ciência de dados e 13, com lentes voltadas para a ciência aberta. Os demais artigos foram desconsiderados. Concluiu-se que 2019 e 2016 foram os anos em que houve mais publicações sobre os temas. Com esse estudo, observou-se o espaço que as temáticas vem ganhando em periódicos da área de Ciência da Informação.

Palavras-chave: *Big data*. Ciência de dados. Ciência Aberta.

BIG DATA IN INFORMATION SCIENCE JOURNALS: an approach focused on open science and data science

ABSTRACT

The purpose of this article was to analyze how the themes: big data and open science in journals in the Information Science have been approached. This is a bibliographic search based on articles from scientific journals on the topics big data and open science. Articles from 21 journals specialized in the Information Science area were analyzed, 3 with Qualis A1 classification, 3 as A2 and 15 titles as B1, quadrennium 2013-2016. Data collection was carried out by searching the search option of the official pages of the selected journals, for the period from 2000 to 2019. The term big data was used as a search strategy and 88 articles were retrieved. Of the total articles retrieved on big data, 6 deal with it in the context of data science and 13, with lenses turned to open science. The other articles were disregarded. It was concluded that 2019 and 2016 were the years in which there were more publications on the topics. With this study, it was observed the space that the themes have been gaining in journals in the area of Information Science.

Keywords: Big data. Data science. Open Science.

1 INTRODUÇÃO

Com apoio das tecnologias a quantidade de dados emergentes vem transformando a comunicação científica, possibilitando novas formas de produção do conhecimento. O *big data*, a ciência de dados e a ciência aberta, por meio do acesso livre às publicações, bem como aos dados de pesquisa são realidades testemunhadas nos ambientes de pesquisa.

Ao abordar o *big data*, o entendimento é que está fazendo referência a um volume de dados que, certamente, excedem formas tradicionais de processamento, tratamento e análise. Este fenômeno tem despertado o interesse de diversos atores pela potencialidades de serem exploradas visando a algum tipo de benefício. Inúmeros são as áreas buscando solucionar problemas por meio de tantos dados. No contexto da Ciência Aberta, uma de suas prerrogativas é disponibilizar o acesso livre a dados de pesquisa.

Entre universos que envolvem tanto a ciência, como a economia surgem oportunidades, ramos especializados no trato e, principalmente, na análise e interpretação de grandes volumes de dados, como a ciência de dados. Porto e Ziviani (2014) expõem que essa área interdisciplinar aplica elementos de diversas disciplinas. Para os autores, emerge como um recurso cada vez mais importante, envolvendo ações de saúde, petrolíferas, energéticas, financeira, esporte, astronomia, bioinformática, Internet, mobilidade urbana, defesa cibernética, comunicação móvel e biodiversidades.

Devido ao tratamento de dados, de informações, da geração de novos conhecimentos e metodologias ocasionadas pelo o excesso de dados essa tripla abordagem do *big data*, ciência aberta e ciência de dados tem se mostrado interessantes para a área de Ciência da Informação. Assim, com esse estudo questiona-se como o *big data*, a ciência de dados e a ciência aberta têm sido abordados em artigos de periódicos da área de Ciência da Informação? Quais anos foram publicados artigos sobre os temas em questão? Quais periódicos publicam sobre os temas? O estudo justifica-se pela necessidade de entender o impacto dessas três abordagens para a área da Ciência da Informação.

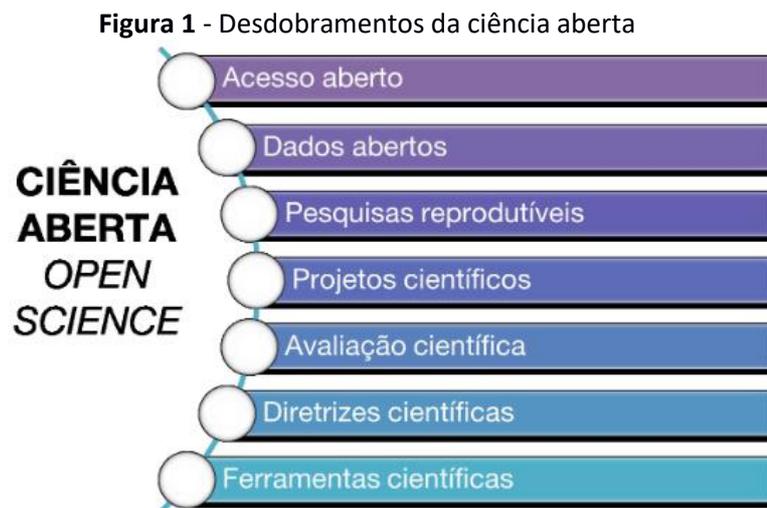
2 DISTINÇÕES CONCEITUAIS

Nos últimos tempos vivenciamos o quarto paradigma para a descoberta científica, que originou-se a partir da disponibilização de um volume exponencial crescente de: dados, evidenciados por recursos como sequenciadores de genoma, telescópios, fontes de luz, redes de sensores e; aparatos tecnológicos: dispositivos com capacidade de coletar e transmitir dados (GUTIERREZ, 2019). Considerando os objetivos do estudo, observou a necessidade de expor a diferença conceitual que circundam as expressões: ciência aberta, *big data* e Ciência de Dados.

Nas palavras de Silva (2019), *big data* pode ser entendido como uma coleção de conjuntos de dados, que apresentam dificuldades para serem processados por bancos de dados ou aplicações de processamento tradicionais. Para Daniel (2020) é um

fenômeno que envolve o crescimento complexo e dinâmico de dados gerados por pessoas, máquinas e recursos tecnológicos. “Tudo e nada ao mesmo tempo, talvez essa seja a melhor forma para representar o *big data*, diante da imensidão de dados e informações que surgem a cada instante” (RIBEIRO; OLIVEIRA; ARAÚJO, 2019, p. 174). Nessa perspectiva, os autores expõem que localizar o que se precisa na web, por exemplo, pode ser considerada uma tarefa tão árdua como encontrar uma “agulha no palheiro” e que o mesmo pode ocorrer com os dados científicos abertos, quando não estão encontráveis, acessível, interoperável e reutilizável, com previsto nas linhas que discorrem sobre Ciência Aberta .

Para Pontika *et al.* (2015) a ciência aberta é “[...]uma tendência universal que exige uma mudança no comportamento dos investigadores para conteúdos abertos e a adoção de um vasto leque de práticas e estratégias abertas relacionadas com todo o ciclo de vida da investigação”. Considerada como um movimento de movimentos por diversos autores, Ribeiro e Oliveira (2019) classifica as práticas de ciência aberta como desdobramentos que se apoiam em seis pilares fundamentais, proposto por Pontika *et al.* (2015): acesso aberto, pesquisa aberta reproduzível, avaliação aberta da ciência, políticas de ciência aberta, ferramentas de ciência aberta e dados abertos, Figura 1.



Fonte: Ribeiro e Oliveira (2019).

Com uma abordagem voltada para a filosofia que prevê a abertura da ciência, junto com os avanços das tecnologias, o acesso aos dados de pesquisa vem chamando

a atenção. Para Marín-Arraiza (2019), dados abertos são dados disponíveis online, sem custos e de forma acessível que podem ser utilizados, reutilizados e distribuídos, sempre seguindo as indicações da fonte dos dados. Para a autora, os dados abertos requerem a ampla documentação para garantir a adequada reutilização e podem ser publicados em repositórios de dados, em periódicos de dados (*data journals*) ou seguindo publicações como os artigos de dados (*data papers*). Em muitos casos, os dados devem passar por um processo de anonimização, pois podem conter conteúdos sensíveis como histórias clínicas de pacientes. Para Semeler e Pinto (2019, p. 119) os dados de pesquisa “[...] representam produto para pesquisa, bem como um insumo para a pesquisa”. De acordo com esses autores esse tipo de dados assume papel relevante referente aos novos processos de comunicação científica.

Levando em consideração todos esses fenômenos: *big data*, Ciência Aberta, dados abertos, junto aos avanços tecnológicos, emerge a ciência de dados que vem impactar em diversas áreas atuação. Para Moraes e Kafure (2020) a explosão de dados na Internet e nas organizações têm provocado um crescente interesse entre cientistas, formuladores de políticas, profissionais de diversos setores no desenvolvimento de pesquisas que usam técnicas de *big data* e ciência de dados. Segundo Morettin e Singer (2020) essas duas expressões têm sido utilizadas em profusão, como se fossem conceitos novos, distintos daqueles com que os estatísticos lidam frequentemente há cerca de dois séculos, o que pode ser considerado equivocada.

Para Silva (2019) a ciência de dados é a disciplina que combina ideias da Estatística e da Ciência da Computação para resolver o problema da descoberta de conhecimento em bases dados, a explorar e analisar dados, fazer inferências e buscar padrões em meio de incertezas.

3 PERCURSO METODOLÓGICO

Esta investigação pode ser entendida como bibliográfica, qualitativa de caráter exploratória. Foram analisados artigos de 21 periódicos especializados na área de Ciência da Informação, sendo 3 com classificação Qualis A1 (Perspectiva Ciência da

Informação, Informação & Sociedade, Transinformação), 3 deles A2 (Informação & Informação, Em Questão, Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência Da Informação) e 15 títulos com classificação B1 (Ágora, Brazilian Journal of Information, Ciência da Informação, Incid: Revista de Ciência da Informação e Documentação, Intexto, Liinc em Revista, Perspectivas em Gestão & Conhecimento, Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, Ponto de Acesso, Revista Brasileira de Biblioteconomia e Documentação, Revista Digital de Biblioteconomia e Ciência da Informação, Revista Eletrônica de Comunicação, Informação & Inovação em Saúde, Revista Iberoamericana de Ciência da Informação, Tendências da Pesquisa Brasileira em Ciência da Informação) (CARVALHO, 2017). Justifica-se a escolha dos títulos por serem considerados pelo Qualis como publicações relevantes para a área de Ciência da Informação.

O primeiro passo foi o levantamento dos artigos que tratavam sobre os temas em estudo, realizado na primeira semana de 2020. Assim, as páginas web oficiais de todos os periódicos listados foram visitadas e na opção de busca, efetuou-se uma pesquisa pelo termo *big data*, no campo título e resumo, nos termos indexadores e até mesmo no texto completo. Essa coleta de dados compôs o período compreendido entre 2000 a 2019. Considerou essa margem adequada por entender que houve evolução das tecnologias de mineração e geração de dados, a partir do início do milênio. Com o universo de 88 artigos, esta pesquisa analisa a produção científica relacionada à ciência aberta e ciência de dados. Portanto, interessou à pesquisa os documentos que tratam-se dessas duas temáticas.

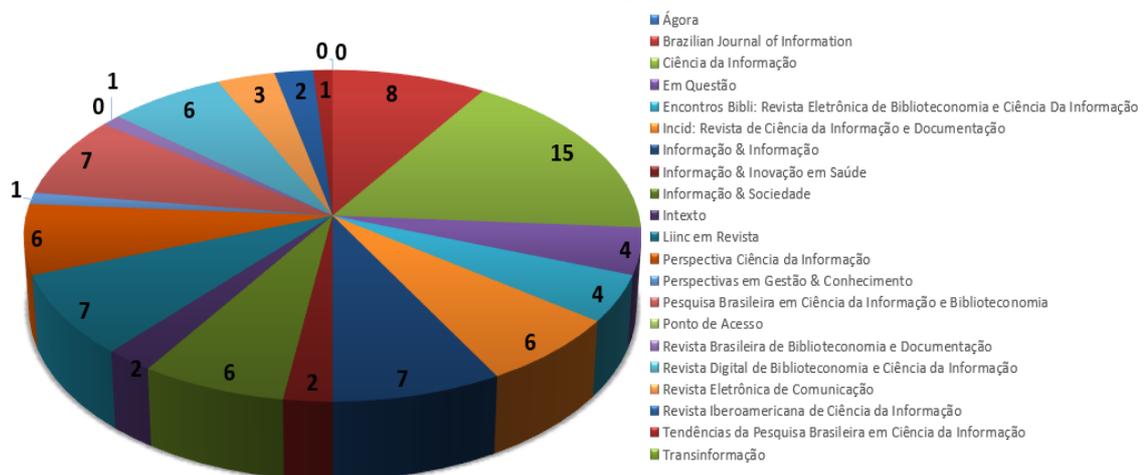
Em um segundo momento, fora realizada a análise dos artigos já pré-selecionados. Nessa etapa, a atenção voltou-se para aqueles que abordaram as duas temáticas ciência aberta e ciência de dados, independentemente, do objeto de estudo do artigo. No intuito de atender ao objetivo deste estudo, foram verificados os artigos que continham aspectos relacionados aos conteúdos investigados pelos pesquisadores. Entre outras características que pudessem traçar convergências entre as temáticas em análise, identificou-se as datas de divulgação dos artigos no intuito de se constatar quais os anos houveram mais artigos publicados.

Por fim, foram realizadas as análises qualitativas dos textos recuperados, expondo os objetivos dos artigos, conceitos, curiosidades, bem como suas principais conclusões e contribuições relevantes para o contexto abordado nesta pesquisa envolvendo o *big data*, ciência aberta e ciência de dados, essa última identificada como relevante para o contexto do estudo.

4 RESULTADOS

Diante da pesquisa inicial foram recuperados 88 artigos que tratam da temática *big data*. Conforme o Gráfico 1, pode-se observar que o título Ciência da Informação é o periódico que mais publicou sobre o tema, com 15 artigos recuperados. No periódico Brazilian Journal of Information foram recuperados 8 artigos e nos periódicos Informação & Informação, Liinc em Revista e Pesquisa Brasileira em Ciência da Informação e Biblioteconomia foram recuperados 7. Nota-se que 3 títulos não publicaram artigos com o tema *big data*.

Gráfico 1 - Quantidade de artigos por periódicos

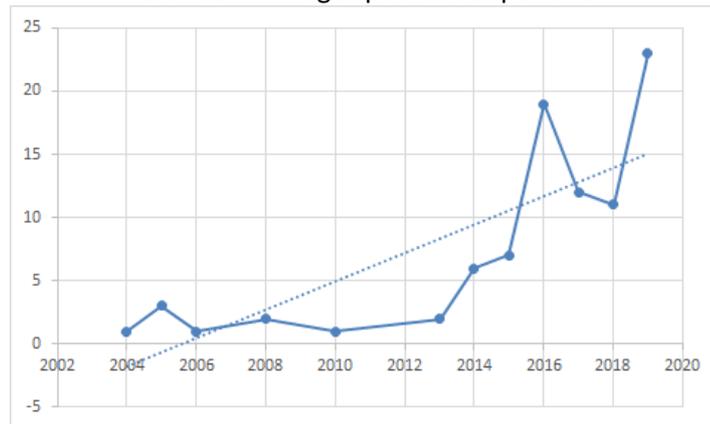


Fonte: Dados da pesquisa.

Ao observar o período em que mais se publicou sobre o *big data* nos periódicos de Ciência da Informação, em análise, identificou-se uma tendência para os próximos anos de ampliar o aceite de artigos com a temática. Esta afirmação pode ser

confirmada por meio do Gráfico 2, devidamente, especificada pela linha tracejada cruzando o gráfico. Identificou-se que o primeiro artigo foi publicado em 2004. O gráfico nos mostra que o *big data* vem chamando a atenção da área, do total de artigos recuperados, 81,81% deles foram publicados nos últimos 5 anos, 7 em 2015, 19 em 2016, 12 em 2017, 11 em 2018 e 23 em 2019.

Gráfico 2 - Artigos publicados por ano



Fonte: Dados da pesquisa

Em seguida, foram analisados os artigos, observando a convergência entre *big data* e a ciência aberta. Do total de artigos recuperados sobre o *big data*, 6 o tratam no contexto da ciência de dados e 13, com lentes mais voltadas para a ciência aberta. Os demais artigos foram desconsiderados. Assim, passa-se a expor as principais abordagens, ideias e discussões expostas em 19 artigos.

No universo dos resultados da pesquisa, os trabalhos que buscaram conceituar o *big data*, concordam que este fenômeno possui cinco pressupostos que possibilitam uma modelagem de dados e sua posterior recuperação, pautada pelos cinco V - Volume, Velocidade, Variedade, Veracidade e Valor. Nas palavras de Araújo Júnior e Sousa (2016), o *big data* tem como uma de suas características possibilitar a gestão dos dados em grande velocidade de processamento, mas não se refere especificamente à sua organização, nem a rastreabilidade dos dados processados. Esse autor analisou o ecossistema de *big data* como subsídio no processo de tomada de

decisão, apresentando a aplicação do estudo no ambiente da Secretaria de Inspeção do Trabalho (SIT) do Ministério do Trabalho (MTB).

Já a pesquisa de Souza, Almeida e Baracho (2013), discutiu o panorama da Ciência da Informação e algumas das suas perspectivas, em face dos fenômenos informacionais contemporâneos e expôs que o termo *big data* se tornou banal, na busca por tentar descrever quaisquer amontoados de dados que não tenham como ser processados sem estruturas de computação específicas à disposição. Os autores acentuam ainda o grande impacto deste fenômeno em áreas como a Ciência da Informação e a Ciência da Computação.

Situado na abordagem da Ciência Aberta, o artigo de Victorino *et al.* (2017) aborda o governo aberto, apresentando uma proposta de sistema de *big data* para dar suporte à análise de dados abertos governamentais conectados, o que pode incrementar a transparência da administração pública. Outro documento recuperado, Caldas e Caldas (2019), trata o *big data*, voltado para a administração, analisou os impactos que as novas tecnologias e alguns de seus potenciais fenômenos correspondentes, entre eles os processos eleitorais e, conseqüentemente, o modelo democrático existente na atualidade. Já a investigação de Rodrigues, Sant'ana e Ferneda (2015), primou por identificar, na fase de recuperação, atributos disponíveis nos momentos em que se realiza pesquisas por conjuntos de dados em repositórios governamentais, a partir do modelo de Ciclo de Vida de Dados para a Ciência da Informação (CVD-CI), proposto por Sant'Ana (2013). O trabalho de Bertin *et al.* (2019), que destaca o *big data*, comenta e analisa a Política de Governança de Dados, Informação e Conhecimento da Embrapa, com foco nas questões relacionadas à gestão de dados de pesquisa.

O trabalho de Bufrem *et al.* (2016) teve como objetivo apresentar a produção científica mundial relativa à ciência orientada a dados a partir dos termos *e-Science* e *Data Science* na *Scopus* e na *Web of Science*, entre 2006 e 2016. Os autores adotaram em sua pesquisa o conceito de *e-Science* proposta por Hey e Trefethen (2005) quando afirmam que a concepção de *e-Science* deve estar ligada à infraestrutura capaz de permitir aos cientistas que seus trabalhos sejam desenvolvidos de modo mais preciso,

mais rápido e/ou diferente, tendo forte relação com o incremento da produtividade e da inteligência científica por meio da partilha, replicação e gestão de dados científicos. Entre os resultados da sua investigação, identificou-se que o termo *big data* esteve entre os destaques da produção científica analisada. Já o artigo Costa e Cunha (2015) contextualiza o surgimento da *e-Science* e sua relação com a Ciência da Informação, buscando conhecer a comunidade científica que produz literatura sobre esses em um universo formado por 202 registros de documentos das bases de dados LISA e LISTA.

Relacionado ao acesso aberto, um dos braços da ciência aberta, identificou-se o artigo Van Reenen (2006), que revisou um trabalho desenvolvido com o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e o progresso mundial do Movimento de Acesso Aberto. Segundo o autor, em 9 de dezembro de 2003 foi proposto ao Governo Brasileiro e/ou o Ibict fazerem uma declaração de apoio ao movimento global de acesso aberto, acatando a Declaração de Berlim sobre o Acesso Aberto ao Conhecimento nas Ciências e Humanidades, durante o 50º Aniversário do Ibict. O autor defende a teoria de que o acesso aberto é um sistema adaptativo complexo criado a partir da publicação acadêmica que utiliza a internet, para influenciar a conexão nas infinitas variedades, na criatividade, no trabalho, na mudança, no conhecimento e na economia da informação.

Entre os documentos recuperados o artigo de Shintaku, Duque e Suaiden (2015) levantou e analisou a adesão dos repositórios brasileiros em relação a: Dados Abertos; Arquivos Abertos; Acesso Aberto; Ciência Aberta; Curadoria Digital e *e-science*. Entretanto, à época do estudo, revelou-se, assim, um cenário ainda incipiente, mas promissor quanto ao acesso aberto.

Uma das premissas da ciência aberta é a colaboração aberta para solução de problemas sociais ou científicos. Nesse contexto Mansell (2014) examinou o potencial para colaboração entre profissionais da ciência formal e grupos frouxamente conectados online que empregam *crowdsourcing* para gerar recursos de informação digital. O autor sugeriu que a ciência aberta precisa tornar-se mais flexível ao tratar da construção de colaborações com grupos pouco conectados em termos equitativos. *Crowdsourcing* pode ser entendido como um processo compartilhado em que as

tarefas de produção e resolução de problemas são distribuídos para uma multidão anônima de pessoas, envolve uma comunidade que concorda em ajudar com o processo (HARRIS, 2013).

Neste ponto passamos a analisar os artigos que abordam o *big data* no contexto da ciência de dados, começando pelo artigo de Rautenberg e Carmo (2019) que procurou evidenciar a diferença e a complementaridade dos conceitos de *big data* e ciência de dados. Os autores sinalizam que o *big data* é um termo intrinsecamente ligado à infraestrutura do hardware e de serviços de computação na nuvem, necessários para o armazenamento, o processamento e a distribuição de recursos. Os autores ressaltam ainda que a Ciência da Informação tem papel fundamental na consolidação dos ecossistemas de *big data*. Quanto à ciência de dados, entende-se que este conceito é conexo à camada dos softwares, a qual metodologicamente transforma os dados em informação para o apoio à tomada de decisão.

A pesquisa de Curty e Serafim (2016) teve como objetivo caracterizar e compreender os aspectos formativos do cientista de dados em 93 cursos em ciência de dados ofertados por instituições estadunidenses. Os resultados do estudo mostraram que existem algumas iniciativas de formação de cientistas de dados no Brasil, na modalidade *lato sensu* de cursos voltados para a ciência de dados e também cursos de Ciência da Informação com essa ênfase. Os autores concluíram que são atribuídas à formação em ciência de dados habilidades voltadas para computação, estatística, matemática, incluindo programação e modelagem avançada, sendo que muitas destas são pré-requisitos para ingresso nestes cursos.

5 CONSIDERAÇÕES FINAIS

Quanto às questões iniciais, observou-se que a revista Ciência da Informação tem dado mais atenção à temática *big data* e que se identificou uma tendência de ampliar o aceite de artigos sobre esse assunto para os próximos anos. Os anos com maior registro de publicações sobre os temas foi 2019, seguido pelo ano de 2016.

No estudo, constatou-se que os autores têm apresentado diversas abordagens para os temas em estudos, reforçando a interdisciplinaridade da área. Os textos tratam de experimentos que, da área da saúde à agrária, demonstram o potencial do *big data* e da ciência aberta.

Tamanho foi a revolução dos dados que houve o surgimento de uma nova disciplina, a ciência de dados, e a necessidade de formação de um novo profissional, o cientista de dados. A ciência de dados bem como o cientista de dados continuam em evolução e vêm ganhando cada vez mais espaço no mercado e na pesquisa da área.

REFERÊNCIAS

ARAÚJO JÚNIOR, R. H. de; SOUSA, 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, DF, v. 45, n. 3, set./dez. 2016. Disponível em: <http://revista.ibict.br/ciinf/article/view/4057>. Acesso em: 17 ago. 2020.

BERTIN, P. R. B. *et al.* A Política de Governança de Dados, Informação e Conhecimento da Embrapa como mecanismo para a gestão de dados de pesquisa agropecuários. **Liinc em Revista**, Rio de Janeiro, v. 15, n. 2, p. 194-204, nov. 2019. DOI: <https://doi.org/10.18617/liinc.v15i2.4798>. Disponível em: <http://revista.ibict.br/liinc/article/view/4798>. Acesso em: 17 ago. 2020.

BUFREM, L. S. *et al.* Produção Internacional Sobre Ciência Orientada a Dados: análise dos termos Data Science e E-Science na Scopus e na Web of Science. **Informação & Informação**, Londrina, v. 21, n. 2, p. 40-67, maio/ago. 2016. DOI: <http://dx.doi.org/10.5433/1981-8920.2016v21n2p40>. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/26543/20114>. Acesso em: 17 ago. 2020.

CALDAS, C. O. L.; CALDAS, P. N. L. Estado, democracia e tecnologia: conflitos políticos e vulnerabilidade no contexto do big-data, das fake news e das shitstorms. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 24, n. 2, p. 196-220, abr./jun. 2019. DOI: <https://doi.org/10.1590/1981-5344/3604>. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362019000200196&lng=en&nrm=iso. Acesso em: 17 ago. 2020.

CARVALHO, J. Saiba o que é Qualis Capes e quais os periódicos da CI com essa classificação. **Biblio Cultural**, [S.l.], 3 mar. 2017. Disponível em:

<https://biblioo.cartacapital.com.br/saiba-o-que-e-qualis-capes/>. Acesso em: 17 ago. 2020.

COSTA, M. M.; CUNHA, M. B. A literatura internacional sobre e-science nas bases de dados LISA e LISTA. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 20, n. 44, p. 127-144, set./dez. 2015. ISSN 1518-2924. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2015v20n44p127>. Acesso em: 17 ago. 2020.

CURTY, R. G.; SERAFIM, J. S. A formação em ciência de dados: uma análise preliminar do panorama estadunidense. **Informação & Informação**, Londrina, v. 21, n. 2, p. 307–331, maio/ago. 2016. DOI: <http://dx.doi.org/10.5433/1981-8920.2016v21n2p307>. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27945>. Acesso em: 17 ago. 2020.

DANIEL, B. K. Big Data e ciência de dados: uma revisão crítica de questões para a pesquisa educacional. Tradução Mirtes Dâmares Santos de Almeida Maia e Danilo Garcia da Silva. **PerCursos**, Florianópolis, v. 21, n. 45, p. 80 - 103, jan./abr. 2020. DOI: 10.5965/1984724621452020080. Disponível em: <https://www.periodicos.udesc.br/index.php/percursos/article/view/1984724621452020080>. Acesso em: 9 dez. 2020.

GUTIERREZ, M. A. Ciência de dados e inteligência artificial em medicina. **Journal of Health Informatics**, v. 11, n. 4, p. 1-2, out./dez. 2019. Disponível em: <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/viewFile/746/375>. Acesso em: 17 ago. 2020.

HARRIS, C. G. **Applying human computation methods to information science**. 2013. 209 f. Thesis (Doctor of Philosophy) - University of Iowa, 2013. Disponível em: <https://doi.org/10.17077/etd.6kebet6l>. Acesso em: 17 ago. 2020.

HEY, T.; TREFETHEN, A. E. Cyberinfrastructure for e-Science. **Science**, London, v. 308, n. 5723, p. 817-821, maio 2005. Disponível em: <https://science.sciencemag.org/content/308/5723/817/tab-article-info>. Acesso em: 12 fev. 2020.

MANSELL, R. Open collaboration for social problem solving: converging or diverging norms of governance authority?. **Liinc em revista**, Rio de Janeiro, v. 10, n. 2, p. 451-459, nov. 2014. Disponível em: <http://revista.ibict.br/liinc/article/view/3594/3073>. Acesso em: 17 ago. 2020.

MARÍN-ARRAIZA, P. **Multimodalidade na publicação científica ampliada: considerações semióticas e modelo de representação**. Orientadora: Silvana Aparecida

Borsetti Gregorio Vidotti. 2019. 189 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências - Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2019. Disponível em:

<https://repositorio.unesp.br/handle/11449/191491>. Acesso em: 2 set. 2020.

MORAES, L. L.; KAFURE, I. Bibliometria e ciência de dados: um exemplo de busca e análise de dados da Web of Science (WoS). **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 18, p. e020016, jun. 2020. DOI:

<https://doi.org/10.20396/rdbci.v19i0.8658521>. ISSN: 1678-765X.

MORETTIN, P. A; SINGER, J. M. **Introdução à Ciência de Dados**: fundamentos e Aplicações. São Paulo: USP, 2020. Disponível em:

<https://www.ime.usp.br/~jmsinger/MAE0217/cdados2020mai14.pdf>. Acesso em: 17 ago. 2020.

PONTIKA, N. *et al.* Fostering Open Science to Research using a Taxonomy and an eLearning Portal. *In: INTERNATIONAL CONFERENCE ON KNOWLEDGE TECHNOLOGIES AND DATA DRIVEN BUSINESS*, 15., 2015, Graz. **Anais [...]**. Graz, Áustria: The Autors, out. 2015. DOI: <http://dx.doi.org/10.1145/2809563.2809571>. Disponível em:

<http://oro.open.ac.uk/44719/>. Acesso em: 17 ago. 2020.

PORTO, F.; ZIVIANI, A. **Ciência de Dados**. 2014. Disponível em:

<https://www.Incc.br/~ziviani/papers/III-Desafios-SBC2014-CiD.pdf>. Acesso em: 17 ago. 2020.

RAUTENBERG, S.; CARMO, P. R. Big data e Ciência de dados. **Brazilian Journal of Information Science: research trends**, Marília, v. 13, n. 1, p. 56-67, mar. 2019. DOI: <https://doi.org/10.36311/1981-1640.2019.v13n1.06.p56>. ISSN: 1981-1640. Disponível em: <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/8315>. Acesso em: 14 jul. 2020.

RIBEIRO, N. C.; OLIVEIRA, D. A. Universidades públicas federais brasileiras. **Múltiplos Olhares em Ciência da Informação**, v. 9, n. 2, 2019. Disponível em:

<https://periodicos.ufmg.br/index.php/moci/article/view/19184>. Acesso em: 2 set. 2020.

RIBEIRO, N. C.; OLIVEIRA, D. A.; ARAÚJO, R. F. Conjecturas da Ciência Aberta na contemporaneidade do Big data. **BIBLOS**, [S.l.], v. 33, n. 2, p. 163-179, jul./dez. 2019. ISSN 2236-7594. Disponível em:

<https://periodicos.furg.br/biblos/article/view/9648/7280>. Acesso em: 2 set. 2020. Doi: <https://doi.org/10.14295/biblos.v33i2.9648>.

RODRIGUES, F.; SANT'ANA, R. C.; FERNEDA, E. Análise do processo de recuperação de conjuntos de dados em repositórios governamentais. **InCID: Revista de Ciência da**

Informação e Documentação, Ribeirão Preto, v. 6, n. 1, p. 38-56, mar./ago. 2015. DOI: 10.11606/issn.2178-2075.v6i1p38-56. ISSN Eletrônico: 2178-2075. Disponível em: <http://www.revistas.usp.br/incid/article/view/73496>. Acesso em: 14 jul. 2020.

SANT'ANA, R. C. G. Ciclo de vida dos dados e o papel da ciência da informação. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 14., 2013, Florianópolis. **Anais [...]**. Florianópolis: UFSC, 2013. ISBN 978-85-65044-06-6. Disponível em: <http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib/paper/viewFile/284/319>. Acesso em: 17 ago. 2020.

SEMELER, R.; PINTO, L. Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. **Ciência da Informação**, Brasília, DF, v. 48, n. 1, p. 113-129, jan./abr. 2019. Disponível em: <http://revista.ibict.br/ciinf/article/view/4461>. Acesso em: 2 set. 2020.

SHINTAKU, M.; DUQUE, C.; SUAIDEN, E. Análise da adesão às tendências da Ciência pelos repositórios institucionais brasileiros. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 6, n. 2, p. 148-169, 2 out. 2015. DOI: <https://doi.org/10.11606/issn.2178-2075.v6i2p148-169>. Disponível em: <http://www.revistas.usp.br/incid/article/view/89191>. Acesso em: 14 jul. 2020.

SILVA, R. D. R. Big data e ciência de dados: como transformar dados em conhecimento. *In: PESQUISAR*, 8., 2019, Aparecida de Goiânia. **Anais [...]**. Aparecida de Goiânia: UNIFAN, 2019. ISSN 2447-2239. Disponível em: http://www.unifan.edu.br/unifan/aparecida/?page_id=2776. Acesso em: 2 set. 2020.

SOUZA, R. R.; ALMEIDA, M. B.; BARACHO, R. M. A. Ciência da Informação em transformação: big data, nuvens, redes sociais e web semântica. **Ciência da Informação**, Brasília, DF, v. 42 n. 2, p.159-173, maio/ago. 2013. Disponível em: <http://revista.ibict.br/ciinf/article/view/1379>. Acesso em: 17 jan. 2020.

VAN REENEN, Johann. Open access and connectedness: stimulating unexpected innovation through the use of institutional open archives. **Ciência da Informação**, Brasília, DF, v. 35, n. 2, p. 17-26, maio/ago. 2006. DOI: <https://doi.org/10.1590/S0100-19652006000200003>. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000200003&lng=en&nrm=iso. Acesso em: 3 set. 2020.

VICTORINO, M. de C. *et al.* Uma proposta de ecossistema de big data para a análise de dados abertos governamentais conectados. **Informação & Sociedade: Estudos**, João Pessoa, v. 27, n. 1, p. 225-252, jan./abr. 2017. DOI: <https://doi.org/10.22478/ufpb.1809-4783.2017v27n1.29299>. Disponível em: <https://periodicos.ufpb.br/index.php/ies/article/view/29299>. Acesso em: 14 jul. 2020.