



**BIBLIOTECONOMIA DE DADOS E CIÊNCIA DE DADOS
NO CONTEXTO DA E-SCIENCE**

DATA LIBRARIAN AND DATA SCIENCE IN THE CONTEXT OF E-SCIENCE

Makson de Jesus Reis

Doutorando pelo Programa de Pós-Graduação em Ciência da Informação (PPGCI/UFBA). Mestre em Gestão da Informação e do Conhecimento pela Universidade Federal de Sergipe (UFS). Bibliotecário na UniFTC. Orcid: <https://orcid.org/0000-0002-3707-6485>. E-mail: maksonreisprof@gmail.com

Normaci Correia dos Santos Sena

Doutoranda e Mestre pelo Programa de Pós-Graduação em Ciência da Informação (PPGCI/UFBA). Bibliotecária da Câmara Municipal de Salvador. Orcid: <https://orcid.org/0000-0003-2477-6201>. E-mail: normaci.correia@yahoo.com.br

RESUMO

Trata-se de uma pesquisa bibliográfica, para análise de publicações científicas da literatura nacional e internacional que envolvam o tema, com o objetivo de discutir a relação entre a Biblioteconomia de Dados e a Ciência de Dados, no contexto da *e-Science*, a fim de elucidar a competência do profissional bibliotecário de dados à luz das bibliotecas universitárias. A Ciência de Dados é um campo amplo, e a Biblioteconomia de Dados é a soma das práticas tradicionais da Biblioteconomia aplicadas a conjunto de dados. Com isso, o papel das bibliotecas e dos bibliotecários condiz com o desenvolvimento em infraestrutura cibernética e *e-Science*, incluindo um elo mais dinâmico e fluido ao longo ciclo de vida dos dados e, principalmente, a capacidade do bibliotecário de se posicionar como parceiro na pesquisa. Os resultados revelam que a atuação do bibliotecário acadêmico, como bibliotecário de dados, em repositório de dados de pesquisa, é inerente à sua formação, contudo é preciso uma educação continuada para lidar com os desafios que se apresentam a respeito da organização do conhecimento.

Palavras-chave: Bibliotecário de dados; Ciência de Dados; *e-Science*.

ABSTRACT

This is a bibliographic research for the analysis of scientific publications of national and international literature involving the subject with the objective of discussing the relationship between Data Librariany and Data Science in the context of *e-Science* in order to elucidate the competence of the data librarian professional in the light of university libraries. Data science is a broad field and Data Librariany is the sum of traditional Librariany practices applied to data set. As a result, the role of libraries and librarians is consistent with development in cyber and *e-science* infrastructure, including a more dynamic and fluid link throughout the life cycle of data and, especially, the librarian's ability to position itself as partners in research. The results reveal that the performance of the academic librarian as a data librarian in a research data repository is inherent to his training, however it takes a continuous education to deal with the challenges they have in relation to the organization of knowledge.

Keywords: Data librarian; Data Science; *e-Science*

1 INTRODUÇÃO

Percebe-se uma mudança mais incisiva nas bibliotecas, nas últimas décadas, comparando-se com os dois últimos séculos, sobretudo nas universitárias, referente à preservação e disseminação dos dados de pesquisa. Essa evolução dá-se, principalmente, pelo motor chamado “tecnologia”. Em decorrência da explosão informacional, no fim do século XX, e com grande ênfase na década de 2000, os estudos ligados a dados, à informação e ao conhecimento passaram por inúmeras transformações que são irreversíveis. Nesse período histórico, o grande volume de dados, conhecido como *Big Data*, influenciou a Biblioteconomia, a Ciência da Informação, a Arquivologia e as disciplinas relacionadas, o que demandou a produção de novos espaços, serviços e o desenvolvimento de competências dos profissionais atuantes nestes ambientes informacionais.

O empoderamento das tecnologias fortaleceu os métodos da *e-Science*, considerada o quarto paradigma das ciências por unificar as ciências empírica, teórica e computacional e por se preocupar com as interfaces entre o computador e as dimensões dos dados, ou seja, os processos de representação da informação e do conhecimento científico, por meio da utilização de tecnologias, que proporcionam o uso intensivo de dados. Conseqüentemente, o estudo sobre a *e-Science* evidenciou a Ciência de Dados, a Biblioteconomia de Dados e o bibliotecário de dados no contexto, principalmente, de bibliotecas universitárias.

Nesse cenário, o predomínio desses recursos digitais altera as atividades das bibliotecas e amplia o papel dos bibliotecários. Assim, este artigo tem por objetivo discutir a relação entre a Biblioteconomia de Dados e a Ciência de Dados no contexto da *e-Science*, a fim de elucidar a competência do profissional bibliotecário de dados à luz das bibliotecas universitárias. A presente pesquisa estabelece um diálogo entre as áreas da Ciência da Informação e Ciência de Dados frente a este profissional com habilidade em dados. Assim, na sequência, apresenta-se a metodologia utilizada para a construção da pesquisa e, logo em seguida, a abordagem teórica que permite esclarecer a interdisciplinaridade entre as áreas, a saber: Ciência de dados e Biblioteconomia de dados, um breve definição; Bibliotecário de dados frente ao *Big Data*; Bibliotecários de dados no contexto da *e-Science*. As considerações finais destaca a relação entre as áreas analisadas.

2 METODOLOGIA

Para atingir o objetivo proposto, a metodologia, para o presente estudo, consistiu, primeiramente, em uma pesquisa bibliográfica em diferentes bases, incluindo a Base de Dados em Ciência da Informação (Brapci), Scielo e a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), em busca de literatura acadêmica e institucional sobre a Ciência de Dados e a Biblioteconomia de Dados, bem como sobre questões relacionadas aos papéis dos bibliotecários universitários frente ao movimento da *e-Science*.

Para a realização do levantamento bibliográfico foram definidas as seguintes palavras-chaves: Big Data, e-Science, data science, Ciência Aberta, Ciência de Dados, Biblioteconomia de Dados, bibliotecário de dados, biblioteca universitária e bibliotecários universitários. Como filtro,

foram definidos: idioma (inglês, espanhol e português); área do conhecimento (Comunicação e Informação e/ou Ciências Sociais Aplicadas e Ciências Exatas, esta por conta da proximidade da temática). Após este processo, os trabalhos recuperados foram analisados com base no grau de relevância para o estudo em questão.

3 CIÊNCIA DE DADOS E BIBLIOTECONOMIA DE DADOS: DEFINIÇÃO

Etimologicamente, a palavra Ciência deriva do latim *scientia*, que significa “conhecimento”, “saber”. Geralmente, pela sua natureza, a ciência comporta um conjunto de saberes pelos quais executam o método científico. Dessa maneira, a Ciência de Dados é um campo amplo que utiliza o método científico, a partir do *Big Data*, acerca de um determinado tema para produzir conhecimento. Além disso, essa ciência está voltada para ajudar indivíduos e organizações a tomarem melhores decisões a partir de dados armazenados, consumidos e gerenciados (KELLEHER; BRENDAN, 2018), ou seja, permite o uso de métodos teóricos, matemáticos, computacionais e outras práticas para estudar e avaliar dados. O seu objetivo principal é extrair informações necessárias que possam ser usadas para múltiplos fins, como tomada de decisão, desenvolvimento de produtos, análise de tendências e previsão, mas também obter *insights* e conhecimentos de qualquer tipo de dados — estruturados e não estruturados.

O fenômeno *Big Data* transita em nosso meio e ultrapassa a capacidade de compreensão humana, seja nos formatos, seja em fontes, participantes e outros interesses. A respeito disso, Taurion (2012b) afirma que há uma má compreensão do fenômeno.

O termo Big Data está cada vez mais popular, embora ainda esteja mal compreendido. [...] O Big Data vem chamando atenção pela acelerada escala em que volumes cada vez maiores de dados são criados pela sociedade. Já falamos comumente em petabytes de dados gerados cada dia e zetabytes começam a ser uma escala real e não mais imaginária e futurista. O que era futuro há uma década, terabytes, hoje nós já temos nas nossas próprias casas.

Para Taurion (2012a), o fenômeno possui cinco dimensões: **volume** – grandes quantidades de dados, que nem sempre é possível processá-los por meio das tecnologias tradicionais e banco de dados; **valor** – para Taurion (2012a), essa dimensão deve ser a principal a ser avaliada, pois, se não for possível transformar os dados em valor, eles não serão úteis; **velocidade** – diz respeito à forma como os dados são criados e propagados; **variedade** – está relacionada aos diversos tipos de dados estruturados e não estruturados que podem ser analisados pelas ferramentas pertinentes ao *Big Data*; e **veracidade** – por meio das ferramentas tecnológicas, é possível analisar os dados, cuja origem passa ser pouco confiável. Diante disso, notam-se habilidades inerentes à formação do bibliotecário frente a este fenômeno *Big Data* e revela-se a relevância dos objetos de estudo da Ciência da Informação.

A Ciência de Dados é uma área interdisciplinar que abrange qualquer tipo de análise de dados e que está relacionada não apenas à Ciência da Computação, mas também a outras áreas que podem ou não usar computadores, por exemplo, os campos básicos em Engenharia, ciência básica e a ciência natural (KELLEHER; BRENDAN, 2018).

O crescimento da produção de dados ocasionado pela massificação, a evolução da internet e, conseqüentemente, o aumento dos computadores proporcionaram um ambiente oportuno, para que os pesquisadores pudessem utilizar os grandes conjuntos de dados, para reutilizar em suas pesquisas, necessitando, portanto, de um rigor metodológico e científico da mesma forma que acontece como na ciência em geral. Ou seja, é imprescindível que os cientistas de dados (pesquisador da área de Ciência de Dados) tenham domínio no método científico e possuam habilidades que abarquem áreas como a Matemática, Estatística, Computação, entre outros.

Na concepção de Finzer (2013), a Ciência de Dados possui a seguinte composição: Matemática e Estatística, pois se refere a compreensão e interpretação dos princípios básicos e experiência adquiridos em Matemática e Estatística; conhecimentos substantivos – que dizem respeito à compreensão do contexto disciplinar para a escolha de uma metodologia de análise válida para dados; computação e habilidade em dados com o intuito de resolver os problemas, que possibilitam uma visualização da estrutura dos dados.

Isso evidencia o elo entre os cientistas de dados e a inteligência artificial, especialmente, seus subcampos de aprendizado de máquina e *deep learning*, para criar modelos e fazer previsões com algoritmos e outras técnicas, incluindo mineração de dados, análise de *Big Data*, extração de dados e recuperação de dados. Em vista disso, a Ciência de Dados cria relações interdisciplinares em várias áreas, o que permite que suas técnicas sejam uma produção com um olhar holístico, minucioso e refinado em dados brutos.

Smith (2006) descreve que tanto a Ciência de Dados quanto o cientista de dados têm abordado, em suas descrições e estudos, a captura de dados, metadados, recuperação rápida, troca e mineração de dados, tendo ainda como definição a extração da informação. A Biblioteconomia, ao longo da sua história, desenvolveu inúmeras técnicas e procedimentos para organizar, recuperar e compartilhar informações.

Em um ambiente tradicional de biblioteca universitária, por exemplo, os metadados se centralizavam nas descrições dos documentos físicos. Com o movimento da *e-Science*, novas ferramentas e outras abordagens sobre a geração e preservação de dados são inseridas, e, conseqüentemente, surge a Biblioteconomia de Dados (*e-Science Librarianhip*), fortalecendo as práticas anteriores e trazendo outras, na geração e preservação de dados (RICE; SOUTHALL, 2016). Com isso, os papéis para o bibliotecário são ampliados. É uma tendência de estudo desenvolvida principalmente em bibliotecas universitárias, pois tem como finalidade investigar problemas práticos relacionados ao armazenamento, ao descobrimento e ao acesso a dados de pesquisa científica.

4 BIBLIOTECÁRIO DE DADOS FRENTE AO *BIG DATA*

Em virtude do advento da Internet, os dados digitais trouxeram grandes melhorias no acesso e na reprodutibilidade. Segundo a publicação do *The Economist*, em 2017, atualmente, dados são considerados recursos muito mais valiosos do que o petróleo. E nesse avanço da Ciência de Dados, a *WEB* é uma grande aliada tanto para o acesso quanto para a visibilidade desses dados. A respeito disso, John Thompson (1998), ao relatar sobre a transformação da visibilidade, mostra como as mídias eletrônicas facilitam a criação de interação parassocial. Na

concepção do autor, esse tipo de interação fornece mais recursos e mais adaptações ao meio, ou seja, possibilita maior transparência no processo de comunicação.

Thompson (2008) evidencia que as transformações suscitadas pela mídia na publicidade mudaram as maneiras de exercer poder político, e a forma de gerenciamento e o fluxo de informações superaram a copresença, ou, a interação face a face. O autor mostra que a evolução da Internet exige um olhar diferenciado no gerenciamento dos dados e das informações.

No âmbito da Ciência da Informação, dados, informação e conhecimento são termos que estão interrelacionados, mas a natureza de cada uma das relações entre eles é discutível, assim como seus significados que possuem conceitos distintos. Para Zins (2007), dados, informação e conhecimento fazem parte de uma ordem sequencial em que os dados são a matéria-prima para a informação e a informação é a matéria-prima para o conhecimento. Sendo assim, o autor assevera que a Ciência da Informação, portanto, deve explorar os dados (construção da informação em estratos/blocos) e a informação, e jamais o conhecimento, pois “é uma entidade de ordem superior” (ZINS, 2007, p. 479).

Essa tríade é um lócus comum de muitas teorias, no entanto, Zins (2007, p. 480, tradução nossa), ao citar Elsa Barber Wellisch (1996), apresenta os seguintes conceitos:

Dados é a representação de conceitos ou outras entidades, fixas em ou em um meio em uma forma adequada para comunicação, interpretação ou processamento por seres humanos ou por sistemas automatizados (Wellisch, 1996). Informação é (1) uma mensagem usada por um remetente para representar um ou mais conceitos dentro de um processo de comunicação, com o objetivo de aumentar o conhecimento dos destinatários. (2) Uma mensagem gravada no texto de um documento. Conhecimento é saber, familiaridade adquirida com a experiência; alcance da pessoa de em formação; uma compreensão teórica ou prática de; a soma do que é conhecido.

Sendo dados os registros armazenados em um computador, é necessário discutir o seu uso e o seu compartilhamento. No âmbito da pesquisa acadêmica, as mudanças ocasionadas em decorrência do grande volume de dados têm evidenciado uma transição histórica. De um lado, a evolução da *e-Science* – movimento que prevê a concepção de uma ciência atuante no uso intensivo de dados e na colaboração por meio do uso de ferramentas – como um novo modo operante de comunicar resultados de pesquisas; na outra ponta, o *Big Data* manifesta-se exigindo habilidade do profissional bibliotecário na gestão desses dados.

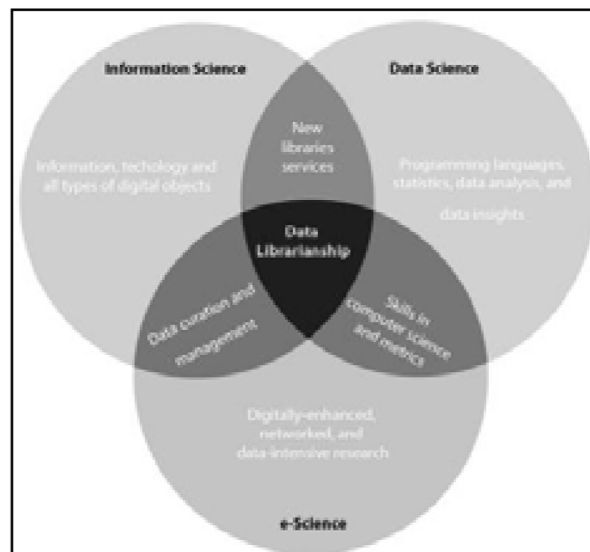
Nesse contexto, surge a gestão de dados como um novo elemento no processo de comunicação da ciência, conseqüentemente, isso leva os bibliotecários a buscar uma educação continuada na execução das suas funções e perfis profissionais no tratamento, armazenamento, para a preservação, disseminação e uso dos dados, o que faz surgir, assim, o bibliotecário de dados.

No entanto, Souza, Almeida e Baracho (2013) argumentam que os cursos de graduação, para formação de bibliotecários e arquivistas, não enfatizam a pesquisa; e os de pós-graduação recebem egressos de áreas distintas e que, na prática, não contribuem para as mudanças necessárias na Ciência da Informação, o que dificulta o processo de organização entre as ciências dedicadas a dado, informação e conhecimento.

A discussão sobre Biblioteconomia de Dados tem sido realizada desde 1960 na América do Norte e na Europa. Ela enfatiza a criação de serviços de biblioteca relacionados ao uso de dados digitais, repositórios de dados e gerenciamento de dados. Ou seja, conforme expresso anteriormente, o impacto das tecnologias de dados e a nova pesquisa intensiva em dados influenciam a Biblioteconomia de Dados (SEMELER; PINTO, 2020). Para os autores, os bibliotecários de dados necessitam combinar habilidades e conhecimentos imprescindíveis para a gestão dos dados.

Dessa maneira, discute-se a Biblioteconomia de Dados, como um fundo teórico, para permitir que bibliotecários de dados participem das atividades inerentes a esses profissionais, por exemplo, a curadoria e a gestão de dados. O elo entre a Ciência da Informação, Ciência de Dados e a Biblioteconomia evidencia a Biblioteconomia de Dados, como um campo interdisciplinar em virtude das conexões práticas e teóricas existentes entre as disciplinas, conforme apontam Semeler e Pinto (2020) na Figura 1.

Figura 1 - União e interseção entre a Ciência da Informação, Ciência de Dados, e-Science.



Fonte: Semeler e Pinto (2020, p. 3).

A Figura 1, portanto, apresenta uma ideia de como o bibliotecário de dados se envolve com três conjuntos teóricos:

- a) a **CI** – com foco em tecnologia, objetos digitais, na natureza e na interação com as pessoas;
- b) **e-Science** – envolve ferramentas técnicas e dispositivos construídos com tecnologia contemporânea na ciência intensiva em dados e refere-se ao uso de tecnologia computacional em processos de investigação científica;
- c) **data science** (Ciência de Dados) – nesse contexto, referem-se a habilidades, métodos, técnicas e tecnologias de Estatística e Ciência da Computação usados para extrair conhecimento e criar novos produtos e serviços a partir de dados;
- d) por fim, o bibliotecário de dados – que se encontra como a intersecção

central entre os três conjuntos (Ciência da Informação, Ciência de Dados e *e-Science*) e representa o impacto das tecnologias de dados e novas pesquisas intensivas em dados em bibliotecas.

Dessa maneira, o ponto de interseção (bibliotecário de dados) mostra que os bibliotecários têm interesse em entender o uso e a reutilização de tecnologias de dados, que podem ser usadas para criar serviços de bibliotecas universitárias (SEMELER; PINTO, 2020). E, por consequência, revela que cada uma das áreas e seus respectivos conceitos são essenciais para a compreensão das atividades do bibliotecário de dados.

Em consonância com Semeler e Pinto (2020), Souza (2006) destaca que as questões relacionadas à formação do profissional bibliotecário e cientistas da informação levam a responder a problemas práticos e teóricos e que afetam suas próprias atitudes e qualidades pessoais.

Em uma síntese dessas posições, pode-se dizer que as questões atuais relacionadas à formação do profissional bibliotecário e cientista da informação perpassam por sua capacitação para responder a problemas práticos e teóricos que se produzem em: Gestão e administração da informação, Tratamento da informação, Tecnologia da Informação, Atendimento e interação com o Usuário e as que afetam suas próprias Atitudes e qualidades pessoais (SOUZA, 2006, p. 26).

Nesse ínterim, as atividades específicas de uma biblioteca universitária englobam procedimentos que demandam que os recursos informacionais impressos e digitais dos dados sejam disponibilizados para a comunidade acadêmica, para tanto, é preciso considerar alguns atributos para se fazer uma gestão de qualidade: tipologia, volume, fluxo de dados, métodos de análise e objetivo primário (Quadro 1).

Quadro 1 - Big Data e análise de dados tradicionais.

Atributo	Big Data	Análise de dados tradicional
Tipo de dados	Não estruturados	Formatados em linhas e colunas
Volume de dados	100 terabytes até petabytes	Volume de dados 100 terabytes até petabytes
Fluxo de dados	Fluxo de dados constante	Conjunto estático de dados
Métodos de análise	<i>Machine Learning</i>	Baseado em hipóteses
Objetivo primário	Produtos baseados em dados	Suporte à tomada de decisão interna

Fonte: Davenport (2014, p. 4).

Nesse ambiente acadêmico, é imprescindível que os bibliotecários de dados se tornem especialistas em informação científica. Contudo a literatura mostra a necessidade de treinamentos adicionais e estratégias práticas, para construir relacionamentos entre os bibliotecários de dados e os pesquisadores científicos, a fim de aprimorar o gerenciamento dos dados e facilitar a abertura, a reprodutibilidade e a interoperabilidade intensificada pelo movimento de acesso aberto (COATES, 2014; GOLD, 2007).

Dada a importância das bibliotecas nas instituições, o bibliotecário de dados é o profissional que deve aplicar as habilidades a respeito da biblioteconomia, da *e-Science* e da Ciência de Dados (GOLD, 2007, tradução nossa). Além disso, resolver os problemas referentes ao ciclo de vida dos dados, processo que envolve alguns estágios: coleta (obtenção dos dados), tratamento, análise, preservação e visualização de dados em repositórios digitais. Cada estágio requer distintos níveis de organização e exige do bibliotecário uma transformação nos moldes tradicionais de atendimento às necessidades dos pesquisadores, ou seja, apoio à gestão desses dados científicos, sendo não apenas consumidor, mas também produtor de dados (SILVA, 2019).

Gold (2007) afirma que o discurso, que envolve o papel das bibliotecas, o que inclui a universitária, e dos bibliotecários, condiz com o desenvolvimento em infraestrutura cibernética e *e-Science*, incluindo um elo mais dinâmico e fluido ao longo do fluxo ou ciclo de vida dos dados e, principalmente, a capacidade do bibliotecário de se posicionar como parceiro na pesquisa. A partir disso, a Biblioteconomia começa a ser entendida como uma área que apoia e se envolve com a conduta de pesquisa e investigação científica, e não somente com o uso e consumo de dados (RICE; SOUTHALL, 2016).

Nesse contexto, frente aos protagonistas do bibliotecário de dados, a literatura aponta algumas ações possíveis:

- a) participar de conferências profissionais de curadoria de dados;
- b) participar de treinamentos relacionados à curadoria de dados e à ciência de dados;
- c) compor rede com gestores de dados que apoiam programas de pesquisa e trabalham para entender suas perspectivas, práticas e cultura;
- d) compreender e apoiar tecnologias e sistemas de publicação de dados que permitem a reutilização de dados;
- e) aconselhar e defender sistemas e padrões de descrição de dados que estimulem a interoperabilidade;
- f) articular dados e questões de informática, entre outros (GOLD, 2007, online, tradução direta).

Além dessas, os bibliotecários podem participar de projetos relacionados à *e-Science*, por isso Garritano e Carlson (2009) identificaram cinco categorias de habilidades que auxiliam esses profissionais, a saber:

- a) conhecimento nas áreas de Biblioteconomia e CI;
- b) conhecimento no tema da pesquisa;

- c) parcerias e divulgação (internas e externas);
- d) participação em pesquisas; e,
- e) trabalho em equipe.

Já para trabalhar com *Big Data*, Stryker (2014) aponta as seguintes habilidades:

- a) habilidade em estruturar ideias;
- b) habilidade em encontrar *insights* nos grandes volumes de dados;
- c) habilidade de transformar esses *insights* em questões utilizáveis e úteis à sociedade;
- d) habilidade de raciocínio e ações rápidas.

Em suma, diante do fenômeno *Big Data*, o bibliotecário de dados exerce papel fundamental na verificação das possíveis mudanças nas práticas de dados e no sistema sociotécnico de pesquisa acadêmica, e a relação com a comunidade colaboradora é que vai fortalecer a gestão e a preservação desses dados a longo prazo.

5 BIBLIOTECÁRIO DE DADOS NO CONTEXTO DA E-SCIENCE

A era digital trouxe mudanças incríveis na forma como as informações e os dados são produzidos, consumidos, adaptados e compartilhados, e, por conseguinte, exige uma transformação de recursos e serviços. A biblioteca tradicional era fácil de definir, era a estrutura de tijolos e argamassa com um ponto de entrada claro e controlado que continha e protegia os recursos físicos selecionados sobre os quais a biblioteca afirmou o controle e a responsabilidade curatorial.

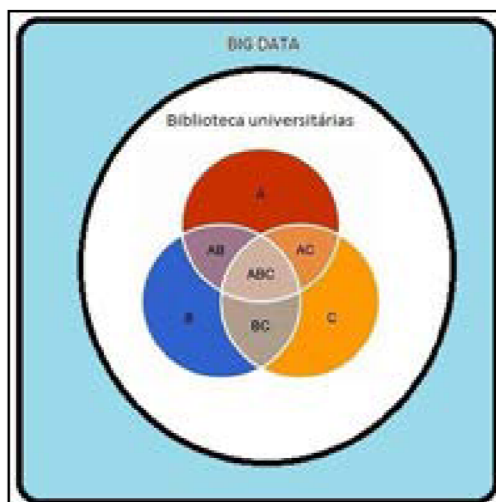
Uma característica da biblioteca tradicional é que tanto a coleção como o seu catálogo utilizam o papel como suporte de registro da informação. Todavia, no final do século XIX, houve uma grande revolução na biblioteca com a introdução do catálogo em fichas e o abandono do catálogo sob a forma de livro (CUNHA, 1999, p. 258).

As bibliotecas atuais implicam em um novo conceito para a armazenagem da informação e para a sua disseminação; e com isso, exigem transformações na maneira de organizar os dados oriundos das instituições a que pertencem.

Bibliotecário de dados não é um conceito novo, mas o estado atual do dilúvio de dados e o aumento de dados digitais exigem mais bibliotecários profissionais para gerenciar dados em instituições acadêmicas, o que resulta na criação de mais cargos para esses bibliotecários.

Os conceitos relacionados entre o *Big Data* e a Ciência de Dados englobam um conjunto de soluções tecnológicas de uma ciência feita a partir de meta base de dados que organiza, avalia e dá sentido a esses dados.

Figura 2 - Relação entre a Ciência de Dados, Big-Data, e-Science, a Biblioteconomia e a Universidade.



Legenda:

A - e-Science

B biblioteconomia C - Ciência de Dados

AB - Biblioteconomia de Dados AC - Cientista de Dados

BC - Ainda em definição ABC - Bibliotecário de dados

Fonte: Elaboração dos autores.

A Figura 2 mostra uma tentativa de relacionar a Ciência de Dados, *Big-Data*, *e-Science*, a Biblioteconomia e a Universidade, com o intuito de evidenciar os pontos de convergências entre os termos e com foco na interseção deles que é o Bibliotecário de dados. Este profissional deve estar presente em todas as vertentes que tem a informação e os dados, como insumo para o avanço científico, ou seja, com toda sua habilidade técnica e competência, é o elo atuante entre todas essas transformações e ampliações tecnológicas e informacionais.

A tecnologia do *Big Data* deve ser vista como um macrossistema que analisa e converte os dados em *insights* de inovações e produtos de dados. Inserido nesse macrossistema tecnológico, estão os subconjuntos, muitas vezes complexos, diversificados, nos quais não seria possível aplicar apenas uma única técnica para o armazenamento e compartilhamento de dados; esses subconjuntos estão relacionados entre si quanto ao uso, acesso e consumo de dados.

Como o *Big Data* se tornou um desafio a ser enfrentado por vários setores da sociedade, a informatização desses ambientes é também um grande problema. No âmbito universitário, as bibliotecas desempenham um papel fundamental no gerenciamento dos dados científicos e acadêmicos. Contudo, Cunha (2010), ao investigar os obstáculos e a missão das bibliotecas e sua relação com os grandes volumes de dados, apontou a *e-Science* como elemento importante nos ambientes de bibliotecas. O autor demonstra que é necessário se pensar na *Big Data* dentro das bibliotecas universitárias como possibilidades do futuro.

Para Cunha (2010), o surgimento de acervo ligado aos dados de pesquisa (científicos) exige bibliotecários com novas habilidades e competências diante do cenário do *Big Data* e da *e-Science*.

Para o referido autor, “[...] o desafio mais crítico será remover os obstáculos que a impedem de responder às necessidades de uma clientela em mudança, transformar os processos e estruturas administrativas que caducaram e questionar as premissas existentes.” (CUNHA, 2010, p. 88).

Sendo assim, as Figuras 1 e 2 estabelecem elos entre a *e-Science*, a Ciência de Dados e a Biblioteconomia. Ao mesmo tempo que essas áreas são independentes pela filosofia que propõe cada uma delas, elas se conectam criando outras áreas: *e-Science* mais Ciência de Dados dão origem aos cientistas de dados; *e-Science* e Biblioteconomia dão origem à Biblioteconomia de Dados; Biblioteconomia e Ciência de Dados originam algo que ainda não foi possível identificar; e na interseção de todos, encontra-se o bibliotecário de dados – profissional que tem de agregar as técnicas tradicionais com as novas ferramentas e estabelecer relação com o pesquisador (GOLD, 2007; RICE; SOUTHALL, 2016).

As habilidades do bibliotecário estão voltadas ao gerenciamento de dados de pesquisa, pela utilização de dados de pesquisa, como recurso ou apoio aos pesquisadores nessas atividades. Os bibliotecários podem oferecer recursos sobre maneiras de lidar com *Big Data* para prepará-los para análise, assim como oferecem recursos em muitos outros tópicos de pesquisa. Os bibliotecários já oferecem tipos semelhantes de recursos como novas fontes de dados, informações de fundo sobre os assuntos e conselhos sobre como lidar com metadados.

Conforme a Figura 3, o bibliotecário de dados tem em seu papel a importância de auxiliar os serviços das bibliotecas e os pesquisadores, em relação a gerenciamento de dados, curadoria e outros serviços, a partir das competências construídas, frente a Ciência de Dados. O *Dataone Usability*¹ adaptou o ciclo de vida dos dados para expor perguntas que bibliotecas universitárias e bibliotecários de dados podem ter em relação aos serviços de dados de pesquisa em contexto.

Figura 3 - Ciclo de vida dos dados e as ações do bibliotecário de dados.



Fonte: Tenopir, Birch e Allard (2012, p. 12).

1 Para mais informações, ver: <http://datacurationprofiles.org>

O ciclo de vida dos dados é um processo que envolve os dados criados por uma empresa ou instituição que vai desde a coleta até a análise (parte final). Cada etapa do ciclo se complementa e exige capacitação para desenvolvê-la, o que viabiliza habilidades diversas para o bibliotecário de dados. Para os autores, a capacidade do bibliotecário de dados em se comunicar com o pesquisador sobre conceitos relacionados aos dados é um dos serviços mais importantes nesse processo. Isso porque irá ajudar os bibliotecários a desenvolver serviços de gestão de dados de pesquisa para suas instituições, bem como promover a compreensão das necessidades dos dados dos pesquisadores e os tipos de dados que eles querem compartilhar, fazer a curadoria e preservar.

Sendo assim, as rápidas mudanças no cenário da pesquisa tornam necessária a continuação de estudos sobre os serviços de dados oferecidos aos pesquisadores. Os autores mostram o papel abrangente e estratégico para as bibliotecas: além de seu papel administrativo mais óbvio no fornecimento de bibliometria, as bibliotecas precisam se interessar por todos os aspectos da atividade acadêmica e participar ativamente da curadoria, assessoria e preservação das produções de pesquisa (TENOPIR; BIRCH; ALLARD, 2012, p. 13).

Os bibliotecários são únicos por terem uma base de conhecimento especializada em relação a descoberta, armazenamento e preservação da informação. Essa habilidade pode ser revertida para uso particular de cientistas de dados, que consultam bibliotecários para ajudá-los a chegar a planos de gerenciamento de dados, por exemplo. Os bibliotecários têm um conjunto geral de expertizes e habilidades de dados que os tornam exclusivamente capazes de aconselhar os clientes da biblioteca sobre as melhores práticas para coleta, gerenciamento e organização de dados.

CONSIDERAÇÕES FINAIS

Percebe-se que a tecnologia possibilita a inovação em diferentes formas e de forma abrangente nos espaços de trabalho. Essas ações são potencializadas pelo processo de descobertas e relações no desenvolvimento ao longo do século. Com o *Big Data*, novas potencialidades de ocupações e estudos evoluíram, representadas nos aspectos de informação, tecnologia e dados. Com isso, a atuação do bibliotecário de dados nas bibliotecas universitárias está vinculada à sua formação e educação continuada, de modo a aprofundar e lidar com os diferentes aspectos que envolvem criação e gerenciamento de metadados referentes ao contexto acadêmico.

Pode-se perceber que, neste artigo, os estudos sobre as relações entre a Biblioteconomia de Dados e a Ciência de Dados, no contexto da *e-Science*, no Brasil, estão incitando, a partir de suas premissas, o surgimento de publicações científicas que começam a estabelecer diálogos entre as áreas. Vale destacar que este artigo tende a promover essas novas discussões devido à interdisciplinaridade entre a Ciência de Dados e a Biblioteconomia, apresentando um campo de atuação profissional: o bibliotecário de dados, que pode atuar frente aos dados, dentre eles, os científicos e os de pesquisa.

REFERÊNCIAS

COATES, Heather L. Building data services from the ground up: strategies and resources. **Journal of eScience Librarianship**, v. 3, n. 1, e1063, p. 52-59, 2014. DOI: 10.7191/jeslib.2014.1063. Disponível em: escholarship.umassmed.edu/jeslib/vol3/iss1/5/. Acesso em: 20 ago. 2021.

CUNHA, Murilo Bastos da. A biblioteca universitária na encruzilhada. **DataGramaZero: Revista de Ciência da Informação**, v. 11, n. 6, p. 1-21, dez. 2010. Disponível em: <https://www.brapci.inf.br/index.php/article/download/17036>. Acesso em: 14 set. 2021.

CUNHA, Murilo Bastos da. Desafios na construção de uma biblioteca digital. **Ciência da Informação**, v. 28, n. 3, p. 257-268, set./dez. 1999. DOI: <https://doi.org/10.1590/S0100-19651999000300003>. Disponível em: <https://www.scielo.br/j/ci/a/Wb33LWZdjFTqxTrRhpDbwcp/#:~:text=A%20biblioteca%20tra> aquela%20onde%20a%20maioria,tablete%20de%20argila%2C%20o%20papiro%20e%20o%20pergaminho%29. dicional%20%C3%A9%20

DAVENPORT, Thomas H. **Big Data at Work: Dispelling the Myths, Uncovering the Opportunities**. Boston: Harvard Business Review, 2014.

GARRITANO, Jeremy.; CARLSON, Jake R. **A subject librarian's guide to collaborating on e-science projects**. Issues in Science and Technology Librarianship, Purdue University, n. 57, spring 2009. Disponível em: https://docs.lib.purdue.edu/lib_research/140/. Acesso em: 21 out. 2021.

GOLD, Anna. Cyberinfrastructure, data, and libraries: part 1: a cyberinfrastructure primer for librarians. **D-Lib Magazine**, v. 13, n. 9/10, 2007. Disponível em: www.dlib.org/dlib/september07/gold/09gold-pt1.html. Acesso em: 19 ago. 2021.

FINZER, William. The data science education dilemma. **Technology Innovations in Statistics Education**, California, v. 7, issue 2, 2013. DOI: <https://doi.org/10.5070/T572013891>. Disponível em: <https://escholarship.org/uc/item/7gv0q9dc>. Acesso em: 22 ago. 2021.

KELLEHER, John; BRENDAN; Tierney. **Data Science**. London: MIT, 2018.

RICE, Robin; SOUTHALL, John. **The data librarian's handbook**. London: Facet Publishing, 2016. DOI: dx.doi.org/10.5195/jmla.2019.601. Disponível em: https://www.researchgate.net/publication/330154840_The_Data_Librarian's_Handbook Acesso em: 15 ago. 2021.

SEMELER, Alexandre Ribas; PINTO, Adilson Luiz. Data librarianship as a field study. **TransInformação**, Campinas, v. 32, e200034, p. 1-12, 2020. DOI: <https://doi.org/10.1590/2318-0889202032e200034>. Disponível em: <https://www.scielo.br/j/tinf/a/LNskQZxLgCVPRvbbFqN9vZz/?format=pdf>. Acesso em: 15 nov. 2021.

SILVA, Fabiano Couto Corrêa da. **Gestão de dados científicos**. Rio de Janeiro: Interciência, 2019.

SMITH, Jack. Data Science as an academic discipline. **Data Science Journal**, v. 5, p.163–164, 2006. Disponível em: <https://datascience.codata.org/articles/abstract/10.2481/dsj.5.163/> > Acesso em: 25 ago. 2021.

SOUZA, Francisco das Chagas de. A Formação acadêmica de bibliotecários e cientistas da informação e sua visibilidade, identidade e reconhecimento social no Brasil. **Informação & Sociedade: Estudos**, João Pessoa, v. 16, n. 1, p. 23-34, jan./jun. 2006. Disponível em: <http://www.biblionline.ufpb.br/ojs/index.php/ies/article/viewFile/439/1491>. Acesso em: 22 out. 2021.

SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos; BARACHO, Renata Maria Abrantes. 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: 18 out. 2021.

STRYKER, Charlie. **Big data will impact every part of your life**. 22 set. 2014. (9:10). Video. Disponível em: <https://www.youtube.com/watch?v=0Q3sRSUYmys>. Acesso em: 22 out. 2021.

TAURION, César. O caos conceitual e os 5V's do big data. **Revista CIO**, 11 maio 2012a. Disponível em: <https://cio.com.br/tendencias/o-caos-conceitual-e-os-5-vs-do-big-data/>. Acesso em: 10 nov. 2021.

TAURION, César. **Você realmente sabe o que é big data?** 30 abr. 2012b. Disponível em: <https://www.tiespecialistas.com.br/voce-realmente-sabe-o-que-e-big-data/>. Acesso em: 10 nov. 2021.

TENOPIR, Carol; BIRCH, Ben; ALLARD, Suzie. Academic libraries and research data services: current practices and plans for the future. **Association of College and Research Libraries (ACRL)**, White Paper. June 2012. Disponível em: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf. Acesso em: 12 ago. 2021.

THOMPSON, John B. **Mídia e modernidade**: uma teoria social da mídia. Tradução de Wagner de Oliveira Brandão. 5 ed. Petrópolis, RJ: Vozes, 1998.

ZINS, Chaims. Conceptual Approaches for defining Data, Information and Knowledge. **Journal of The American Society for Information Science and Technology**, v. 58, n. 4, p.479–493, 15 feb. 2007. DOI: 10.1002/asi.20508. Disponível em: <https://www.researchgate.net/publication/220432993>. Acesso em: 14 out. 2021.

Recebido/ Received: 13/09/2021

Aceito/ Accepted: 10/12/2021

Publicado/ Published: 31/12/2021