



XXII Encontro Nacional de Pesquisa em Ciência da Informação – XXII ENANCIB

ISSN 2177-3688

GT 6 – Informação, Educação e Trabalho

ATUAÇÃO BIBLIOTECÁRIA NA E-SCIENCE: PERFIL PROFISSIONAL E O PLANO DE GESTÃO DE DADOS

PERFORMANCE OF LIBRARIANSHIP IN E-SCIENCE: PROFESSIONAL PROFILE AND DATA MANAGEMENT PLAN

Fabício Rodrigues dos Santos Garrido. UFPB.

Gracy Kelli Martins. UFPB.

Modalidade: Trabalho Completo

Resumo: Este trabalho tem como objetivo evidenciar a relevância da atuação bibliotecária no suporte ao pesquisador, mais especificamente, no que se refere à estruturação do Plano de Gestão de Dados (PGD), exigido para a elaboração e o desenvolvimento de projetos de pesquisa, dentro do contexto de *e-Science*. A Ciência Orientada a Dados, também conhecida com *e-Science*, vem se consolidando, e os dados de pesquisa, nesse contexto, precisam cada vez mais de organização e de gestão. O(a) pesquisador(a) necessita de apoio nessa gestão, a qual envolve um ciclo com diversas fases, além de métodos e padrões de organização desses dados, como os princípios FAIR. Uma das ferramentas gerenciais usadas para essa gestão é o PGD. Por meio de uma revisão de literatura, apresentam-se os principais conceitos que fundamentam as técnicas e as metodologias existentes entre a Ciência da Informação, a Biblioteconomia e a Ciência Orientada de Dados e as correlações entre elas. O estudo revelou que esse profissional da informação, que outrora atuava em pós-publicações de livros, periódicos, trabalhos acadêmicos etc., agora deverá se preocupar com a pré-publicação, ou seja, com a gestão dos dados advindos da *e-Science*.

Palavras-chave: Bibliotecário de Dados. Curadoria de Dados. Plano de Gestão de Dados.

Abstract: This work aims to highlight the relevance of the librarian's performance in supporting the researcher, more specifically in what it refers to the structuring of the Data Management Plan (PGD), required for the elaboration and development of research projects, within the context of *e-Science*. Data Drive Science, also known as *e-Science*, has been consolidating, and research data, in this context, increasingly need organization and management. The researcher needs support in this management, which involves a cycle with several phases, in addition to methods and standards for organizing these data, such as the FAIR principles. One of the management tools used for this management is the PGD. Through a literature review, the main concepts that underlie the existing techniques and methodologies between Information Science, Librarianship and Data Drive Science and the correlations between them are presented. The study revealed that this information professional, who once worked in post-publications of books, journals, academic works, etc. now you will have to worry about pre-publication, that is, with the management of data coming from *e-Science*.

Keywords: Data Librarian. Data Curation. Data Management Plan.



1 INTRODUÇÃO

A atuação dos profissionais da informação, que vêm acompanhando o desenvolvimento informacional da comunidade acadêmica, tem evidenciado a imersão do(a) bibliotecário(a) no contexto da Ciência Orientada a Dados - ou *e-Science* – trabalhando com os dados de pesquisa. Nessa configuração, a produção massiva de dados provenientes do *Big Data* proporciona novas tendências e exigências no que se refere aos dados. Dentre as possibilidades de atuar nesse universo, destacam-se os processos de planejamento e a gestão dos dados de pesquisa, atividades tradicionalmente desenvolvidas por um (a) profissional bibliotecário(a).

Os dados gerados no *Big Data*, em que a *e-Science* está inserida, têm um ciclo de vida estruturado em etapas pelas quais esses dados passam, desde sua criação até o armazenado e/ou o descarte. Nesse percurso, que exige a Gestão de Dados, um dos processos imprescindíveis para que o ciclo se processe de maneira eficiente é o Plano de Gestão de Dados (PGD), de responsabilidade do(a) pesquisador(a). Durante todo o Ciclo de Vida dos Dados (CVD), o PGD orienta as diversas fases da produção, da manutenção e da disponibilização dos dados que, para que sejam usados e reusados, existem padrões normatizadores, como os princípios FAIR¹, que demonstram como os dados devem ser tratados para que possam ser encontráveis, acessíveis, interoperáveis e reusáveis.

A tendência das agências de fomento ou instituições de pesquisa, públicas ou privadas, é de que, para além de apresentar os projetos de pesquisa, os(as) pesquisadores(as) elaborem o PGD, porquanto esse documento é imprescindível para a pesquisa financiada e provê a gestão, o armazenamento e a disponibilização dos dados da pesquisa, garantindo o reuso dos dados coletados.

Tradicionalmente, o profissional da Biblioteconomia é capacitado para gerir a informação e os dados, já que apresenta um perfil apto para esse contexto de pesquisa científica, quanto ao uso e ao reuso de dados criados e recriados de forma colaborativa, além de auxiliar o(a) pesquisador(a) a gerir seus dados de pesquisa e elaborar o PGD.

1 Anacrônico de *Findable* (encontráveis), *Accessible* (acessíveis), *Interoperable* (interoperáveis) e *Reusable* (reutilizáveis).



Nesse âmbito, o objetivo deste trabalho é de mostrar, com base na literatura das áreas de Biblioteconomia, Ciência da Informação e Ciência de Dados, a importância da atuação bibliotecária, a partir da estruturação do PGD, como eixo orientador para o(a) pesquisador(a), em relação ao processo exigido na gestão dos dados de pesquisa, no contexto de *e-Science*. Para alcançá-lo, a metodologia adotada envolveu uma pesquisa bibliográfica e uma revisão de literatura, em que se abordam os principais conceitos que fundamentam técnicas e metodologias existentes entre a Ciência da Informação, a Biblioteconomia e a Ciência Orientada de Dados e as correlações entre elas, para, por meio de um arcabouço teórico, demonstrar as possibilidades que o bibliotecário de dados tem de auxiliar o(a) pesquisador(a) a gerir seus dados de pesquisa por meio do PGD. Para isso, o trabalho revisitou os conceitos de dados e dados de pesquisa propostos por Setzer (2015), Borgman (2015), Semeler e Pinto (2019) e de gestão de dados de pesquisa sob o ponto de vista de Silva Segundo e Araújo (2019) e Sant'Ana (2013; 2019).

2 O QUARTO PARADIGMA CIENTÍFICO

A *e-Science*, além de ser uma nova forma de se fazer pesquisa, é um campo de estudo cujo objetivo é de explorar dados e definições distintas em cada área do conhecimento, tanto no campo científico quanto no corporativo, de natureza interdisciplinar. Seu estudo, como um campo científico, está fortemente atrelado à Ciência da Informação. Ferreira (2018) analisa um estudo realizado por Costa e Cunha (2015) que [...] "evidencia a pertinência de pesquisas relacionadas à *e-Science* entrelaçada com a área da Ciência da Informação, devido à relevância e à atualidade do tema e à produção na área ainda pequena" (FERREIRA, 2018, p. 15-16).

Para Gray (1994), o quarto paradigma da Ciência derivou-se da Ciência de Dados e reconhece que a *e-Science* é o encontro da Computação com a Ciência, que passa a possibilitar pesquisas com o uso e o reuso intensivo dos inúmeros dados que são originados nas investigações científicas. Com base nessa noção, buscam-se melhores definições do termo dados, diferenciando-o do conceito de informação, e a necessidade de estudos por meio de uma melhor gerência e curadoria desses dados, para que se libertem dos discos rígidos e dos *pen-drives* pessoais dos pesquisadores e alcancem repositórios digitais de dados, consolidando a pesquisa colaborativa proporcionada pela *e-Science*.

Compreende-se que a Ciência de Dados, que emergiu do *Big Data*, é proveniente, principalmente, da massiva produção e do acúmulo de dados, originados pelo forte



desenvolvimento das Tecnologias Digitais de Comunicação e Informação (TDIC) que, devido à sua popularização e ao surgimento do Movimento Acesso Aberto (*Open Access*), marcado pela Declaração de Budapeste de 2002, bem como a consolidação da Ciência Aberta (*Open Science*), facilitou o acesso, a criação e o compartilhamento de dados.

Para entender a Ciência Orientada a Dados, este trabalho visa mostrar a diferença entre dados e dados de pesquisa científica. Setzer (2015) e Borgman (2015) referem que os dados são fatos, números, letras e símbolos que podem ser quantificados ou quantificáveis, e os dados de pesquisa científica são oriundos de pesquisas científicas, que demandam um "tratamento especial", e cujo uso é essencial para a Ciência de Dados. No que tange ao tratamento desses dados, é preciso gerir todo o seu ciclo de vida, com especial atenção para a coleta, o armazenamento, a recuperação e o descarte, porquanto serão usados, compartilhados e reusados tanto por seu criador quanto por outros pesquisadores.

Dessa necessidade, surgiram os princípios FAIR - acrônimo de Encontráveis (*Findable*), Acessíveis (*Accessible*), Interoperáveis (*Interoperable*) e Reutilizáveis (*Reusable*). Para que seja possível usá-los e reusá-los, é necessário utilizar metadados, o que vai além da simples definição de "dados sobre dados", pois os metadados devem descrever os dados, tendo em vista que se trata de "[...] um objeto potencialmente informativo que descreve outro objeto potencialmente informativo" (POMERANTZ, 2015, p. 26), o que, na linguagem da Biblioteconomia, significa catalogar e indexar dados de pesquisas.

Tradicionalmente, a Biblioteconomia dá suporte para a recuperação e o acesso às coleções de materiais informacionais, por meio das técnicas de catalogação, indexação, classificação, serviços de referência, treinamentos, entre outros. No contexto da *e-Science*, esses atributos se mantêm em relação ao tratamento do uso e do reuso intensivo de dados e dão suporte à elaboração de metadados e à promoção de treinamentos para os pesquisadores, no que diz respeito à organização dessa forma de se fazer pesquisa, com práticas recomendadas para tratar os dados de pesquisa em todo o seu ciclo de vida, incluindo a estruturação do Plano de Gestão de Dados (PGD).

3 A IMPORTÂNCIA DO PLANO DE GESTÃO DE DADOS

O termo 'dados de pesquisa', segundo Borgman (2015), é uma atividade de pesquisa acadêmica. Nessa perspectiva, é possível afirmar que os dados de pesquisa são os obtidos por meio de uma pesquisa acadêmica? De acordo com a Organização para Cooperação e



Desenvolvimento Econômico — OCDE (2007, p. 13, tradução nossa), os dados de pesquisa são “registros de fatos usados como fontes primárias de pesquisa científica e que são comumente aceitos na comunidade científica como necessários para validar resultados da pesquisa”.

Em seu livro, *‘Big data, little data, no data: scholarship in the networked world’*, Christine Borgman (2015), além de dissertar sobre dados e dados de pesquisa, faz uma associação com as bolsas de estudo na Academia, afirmando que “[...] dados são representações de observações, objetos ou outras entidades usadas como evidência de fenômenos para fins de pesquisa ou bolsa de estudos” (BORGMAN, 2015, p. 28, tradução nossa). No contexto brasileiro, existem as agências de fomento que financiam pesquisas científicas que, conseqüentemente, são geradoras de muitos dados.

Ao gerar os dados de pesquisa, o(a) cientista deve ponderar que ele não é mais aquele que trabalha sozinho em sua garagem sem apoio, mas que, agora, a Ciência é global, e as pesquisas científicas são realizadas em rede. “Grosso modo, o paradigma que precisa ser destruído é a ideia de que nós, cientistas, existimos como indivíduos fora da rede” (WILBANKS, 2011, p. 219).

Não obstante o surgimento do *Big Data*, os recursos disponíveis de tecnologias digitais na comunicação científica e os pressupostos da ciência aberta revelam que “estimativas recentes indicam que mais de 50% dos achados científicos não aparecem na literatura publicada e residem em gavetas de arquivos e discos rígidos pessoais” (FERGUSON et al, 2014, p. 1443). Esse fenômeno prejudica todo o sistema de pesquisa científica, principalmente no contexto da *e-Science*, e gera um registro incompleto de dados e, muitas vezes, uma duplicação desnecessária de esforços científicos que seriam evitados ao propiciar acesso, o que contribui para que haja falhas na replicação dos experimentos realizados, e a Ciência seja inviabilizada. Essa ciência invisível se encontra na cauda longa dos dados² e “[...] constitui um território onde os dados são altamente heterogêneos e dispersos entre instituições, projetos, laboratórios, pequenos grupos de pesquisa e pesquisadores individuais” (SALES, SAYÃO, 2018, p. 4185).

2 “Os dados de cauda longa [...] podem ser definidos como conjuntos de dados pequenos e granulares [publicáveis], coletados por laboratórios individuais no decorrer da pesquisa do dia-a-dia. [...] A cauda longa dos dados também é composta de “dados escuros”, dados não publicados que incluem resultados de experiências com falha e registros que são vistos como auxiliares de estudos publicados (por exemplo, registros de cuidados veterinários). Embora esses dados possam não ser considerados úteis no sentido tradicional, os esforços de compartilhamento de dados podem iluminar informações e descobertas importantes ocultas nessa longa cauda (FERGUSON et al, 2014, p. 1442).



Partindo dessa perspectiva, o *Digital Curation Centre* (DCC) - Centro de Curadoria Digital – assevera que a curadoria digital é necessária para a pesquisa científica e mostra sua importância, principalmente, para “desinvisibilizar” a ciência da cauda longa:

O gerenciamento ativo dos dados de pesquisa reduz as ameaças ao seu valor de pesquisa a longo prazo e reduz o risco de obsolescência digital. Enquanto isso, dados selecionados em repositórios digitais confiáveis podem ser compartilhados entre a comunidade de pesquisa mais ampla do Reino Unido. (*DIGITAL CURATION CENTRE*, 2020, *online*, tradução nossa).

Assim, devido a essa necessidade, surgiram na *e-Science* novos modelos científicos, o ciclo de vida dos dados, conhecidos, na literatura internacional, como *Data Lifecycle*. Sobre isso, Renata Anjos (2019, p. 39) informa que são “[...] modelos aplicados ao processo de curadoria de dados científicos, que são compostos por etapas de gestão e compartilhamento de dados que, em sua maioria, objetivam a preservação desses dados para gerações atuais e futuras de usuários”.

Luís Sayão e Luana Sales (2015, p. 11) referem que “os dados e as coleções de dados de pesquisa possuem um tempo de vida maior que os projetos de pesquisa que os criaram.” Portanto, eles podem ser utilizados mesmo depois de findados os projetos e os financiamentos dos quais são provenientes e reutilizados em novos projetos de pesquisa, reiniciando um novo ciclo de vida para os futuros estudos. Assim, entende-se que esse modelo é “[...] constituído de estágios que conduzem à gestão dos dados científicos e estimulam a colaboração entre pesquisadores e instituições através do compartilhamento de parcerias institucionais” (OLIVEIRA, 2016, p. 19).

Os dados de pesquisa podem diferir em formato e em conteúdo (podem estar em planilha, imagem, vídeos etc., conter números, variações semânticas e fenômenos naturais etc.), no entanto vários elementos básicos são essenciais para gerenciar eficazmente esses dados, os quais, a princípio, devem ser gerenciados para que qualquer pesquisador, coletor ou criador dos dados seja capaz de descobri-los, usá-los e interpretá-los depois de determinado período de tempo. Sendo assim, o PGD “[...] envolve a gestão de dados de pesquisa desde a sua gênese, assegurando a preservação por longo prazo, a descoberta, a interpretação e o reuso” (SAYÃO; SALES, 2012, p. 179).

Nesse contexto, “um componente chave do gerenciamento de dados é a descrição abrangente dos dados e das informações contextuais que futuros pesquisadores precisam entender para usar os dados” (STRASSER; COOK; MICHENER; BUDDEN, 2012, p. 2), por meio



dos metadados, o que é equivalente à catalogação de livros, mas, nesse caso, dos dados de pesquisa.

No Brasil, a Comissão Nacional de Energia Nuclear ligada ao Ministério da Ciência, Tecnologia, Inovações e Comunicações (BRASIL, 2021, *online*) sugere dez razões para se criar um PGD de pesquisa (como uma ferramenta da curadoria), a saber:

1. Ajustar o seu projeto de pesquisa às políticas mandatórias da sua instituição e/ou dos órgãos de fomento à pesquisa;
2. Assegurar a integridade da pesquisa e o seu potencial de replicação;
3. Assegurar que os dados e demais registros de pesquisa sejam acurados, completos, autênticos e confiáveis;
4. Aumentar a sua eficiência como pesquisador – um plano que organize os dados e seu armazenamento permita que você foque na sua pesquisa. Você estará mais capacitado a localizar e usar seus dados e compartilhá-los com seus colaboradores;
5. Permitir que os seus dados sejam compreensíveis agora e no futuro – se os dados são bem documentados antes e durante a formação da coleção de dados, eles serão mais facilmente entendidos e reutilizados;
6. Economizar tempo e recursos em longo prazo;
7. Aumentar a segurança dos dados e minimizar os riscos de perda;
8. Evitar a duplicação de esforços na coleta ou regeneração dos dados, possibilitando que outros pesquisadores se beneficiem dos seus dados e os interprete em outros contextos e com novas visões;
9. Aumentar a visibilidade da pesquisa – se seus dados foram planejados para estar organizados e corretamente arquivados, eles poderão ser identificados, recuperados e citados, aumentando a visibilidade de sua pesquisa e seu prestígio como pesquisador;
10. Tornar mais fácil a preservação e o arquivamento – ajustando antecipadamente a geração de dados, tomando como referência as práticas, os formatos e os demais padrões mais adequados ao arquivamento e à preservação de longo prazo, torna a gestão de dados mais fácil e menos custosa; além do mais, tornam os dados mais aderentes aos requisitos dos repositórios e centros de dados (BRASIL, 2021, *online*).

Existem algumas ferramentas gratuitas e *online* para elaborar o PGD, como o DMPTool³, ligado, entre outros, à instituição DataONE e disponível em inglês e em português, e o DMPonline⁴, respectivamente associados aos Planos de Gestão de Dados nos EUA e nos países do continente europeu (DCC), disponíveis em alemão, inglês (britânico e estadunidense), espanhol e francês.

3 <https://dmptool.org/>

4 <https://dmponline.dcc.ac.uk/>



Como está disponível em português, tomaremos o DMPTool como exemplo, porquanto é um aplicativo *online*, gratuito e de código aberto, que dispõe de um modelo de PGD comum que qualquer pesquisador poderá utilizar. No website do DMPTool, também há planos personalizados por instituições associadas, algumas das quais são brasileiras, como: USP, UNESP, Unicamp, UNIFESP, UFSCAR, UFABC, FAPESP, PUC-CAMPINAS, UNINOVE, Instituto Butantan, ITA, UFRGS, UNITAU, UNISA, UNIMAR, Mackenzie, UNOESTE, Instituto de Botânica, UNISO, FAMEMA, UFRJ, FMJ E IBICT.

Nesse sistema, o(a) pesquisador(a) pode fazer um PGD-teste, editá-lo em qualquer momento e disponibilizá-lo na própria plataforma da DMPTool para qualquer outro usuário do sistema e acessá-lo com código identificador único, *Digital Object Identifier (DOI)*, data de criação e data da última modificação. É possível localizar o PGD de outros pesquisadores por órgão financiador, instituição de pesquisa, idioma e assunto.

No plano, o(a) pesquisador(a) deve mencionar os detalhes essenciais de seu projeto, como título, resumo etc., inserir colaboradores, mostrar como os dados serão usados, os resultados pretendidos, fazer o download do plano e finalizá-lo, publicá-lo ou deixá-lo privado. Todas as etapas são preenchidas no formulário do próprio sistema.

No que diz respeito à importância do PGD e de suas características, o profissional que atua nos sistemas de gerenciamento de coleções para bibliotecas é capacitado para indexar e catalogar diversos tipos de informações em vários suportes diferentes e habilitado para trabalhar com esse tipo de ferramenta. Ele atua tradicionalmente “para uma gestão eficiente da informação” e pode auxiliar, sem muitas dificuldades, o(a) pesquisador(a), que, em muitas circunstâncias, dispõe de pouco tempo para atividades como a curadoria de dados de pesquisa (COSTA; CUNHA, 2014, *online*).

4 BIBLIOTECONOMIA E DADOS DE PESQUISA

A formação bibliotecária tem tido poucas atualizações oficializadas nacionalmente. A última foi feita pelo Conselho Nacional de Educação/Câmara de Educação Superior (CNE/CES), que emitiu, em julho de 2001, o Parecer CNE/CES 492/2001⁵, que define as Diretrizes Curriculares Nacionais (DCN) para alguns cursos superiores, entre eles, o de Biblioteconomia, e estabelece o perfil dos formandos, as competências, as habilidades, os conteúdos

5 <http://portal.mec.gov.br/cne/arquivos/pdf/CES0492.pdf>



curriculares, os estágios, as atividades complementares, a estrutura do curso e a avaliação institucional.

Conforme as Diretrizes Curriculares para os Cursos de Biblioteconomia, os formandos em Biblioteconomia devem ser formados e preparados

[...] para enfrentar com proficiência e criatividade os problemas de sua prática profissional, produzir e difundir conhecimentos, refletir criticamente sobre a realidade que os envolve, buscar aprimoramento contínuo e observar padrões éticos de conduta [...]. (BRASIL, 2001, p. 32).

No Brasil, ao longo do processo do ensino da Biblioteconomia e dos processos de formação do profissional da área, promoveram-se discussões sobre a ênfase do ensino, em um caráter mais humanístico ou mais tecnicista, porém, em toda formação curricular, disciplinas como Catalogação, Classificação e Indexação sempre foram mantidas por compor um espaço nuclear para a área (GUIMARÃES, 2017), que trabalha com a informação e, atualmente, em uma organização mais contemporânea, com dados.

Convém enfatizar que os sistemas de organização do conhecimento se desenvolveram por meio de diversas tecnologias, especialmente as computacionais, e das novas formas de representar, como os metadados e as linguagens de máquinas, dando um enfoque semântico e estabelecendo a necessidade de organizar informações e dados de forma compreensiva por humanos e por máquinas.

O advento da Internet e a popularização da criação e do compartilhamento de dados e informações de forma exponencial tornaram a capacidade operacional de organizar esse gigantesco volume cada vez mais desafiadora. “Nesse contexto, a Ciência da Informação estabelece-se como uma ciência de organização e recuperação do relevante em meio ao excessivo, evidenciada por Jesse Hauk Shera”, ainda entre as décadas de 1950 e 1960 (MARTINS, 2014, p. 92).

Ribeiro (2017) assevera que é possível propor novas abordagens para representar no cenário atual, por meio das mudanças impostas pelas tecnologias, associadas ao dilúvio de dados e de informações que, presentes na sociedade atual, requerem novas demandas de organização e representação. Costa e Cunha corroboram essa assertiva afirmando que,

na Ciência da Informação, a *e-science* traz implicações relevantes sobre a comunicação científica, afinal os dados oriundos da *e-science* são de fato dados científicos primários. Também gera efeitos nos serviços e produtos de informação, bem como afeta diretamente as bibliotecas digitais exigindo



reflexões sobre preservação digital e o planejamento das bases de dados (COSTA; CUNHA, 2014, *online*).

A partir dessa concepção, a discussão sobre informação e dados requer um(uma) profissional da Biblioteconomia especializado(a) como bibliotecário(a) de dados na sociedade atual, com base no uso intensivo de dados pela *e-Science*.

5 O PROFISSIONAL DA BIBLIOTECONOMIA E O PGD

O Serviço de Referência, que se institucionalizou a partir do século XX, aproximou bem mais o profissional da Biblioteconomia de seu usuário e se tornou um verdadeiro orientador de suas necessidades informacionais. “É junto ao bibliotecário de referência que o usuário requisita ajuda para suas necessidades informacionais e busca auxílio para alcançar a informação que deseja. Nessa perspectiva, ocorre o processo de busca da informação” (LIMA; ARAÚJO, 2020, p. 3).

Com a evolução das TDIC e, principalmente, a solidificação da *Wikis* e do Google Acadêmico e a popularização da base de dados científicas, como a *Web of Science*, por exemplo, diversos usuários se tornaram autossuficientes na busca por seu referencial teórico. Nesse contexto, o(a) bibliotecário(a) de referência deve acompanhar esse cenário e, segundo Gracirlei Lima e Carlos Araújo (2020), auxiliá-lo a se tornar autossuficiente em suas pesquisas bibliográficas com o “Processo de Referência Educativo (PRE), [que] se caracteriza pela orientação e pela capacitação de um ou mais usuários para o acesso e para o uso de fontes de informação sobre determinada área ou tema de interesse” (LIMA; ARAÚJO, 2020, p. 7).

Considerando que esse processo autossuficiente poderia causar um esvaziamento de pesquisadores nas Bibliotecas Universitárias, já que muitos usuários sabem pesquisar e podem encontrar grande parte de suas fontes de pesquisa *online*, qual seria o papel do(a) bibliotecária(o) nesses espaços?

O bibliotecário, em geral, tem formação voltada para o armazenamento, a organização, o acesso e o uso da informação, no escopo da organização da informação bibliográfica (física), eletrônica e digital, visando à recuperação e à apropriação efetiva dos usuários. Esse profissional está intimamente ligado a essa nova era de produção e compartilhamento de dados, em que grande parte das informações são geradas de forma privada ou colaborativa e exigem que esses serviços sejam prestados por diferentes tipos de bibliotecas, sejam elas tradicionais ou digitais/virtuais.



Então, o bibliotecário é capacitado para gerir dados e, nesse mundo competitivo, ele tem antelação diante dos demais profissionais para assumir esse posto, tendo em vista o panorama crescente de produção, organização e armazenamento de dados que atuam na gestão de dados de pesquisa. Por isso, é considerado um cientista de dados.

[...] no panorama dos grandes volumes de dados, o cientista de dados desponta como uma das profissões promissoras e desafiantes da atualidade diante do fenômeno do Big Data e da gestão desses mega dados no contexto da Ciência da Informação (RODRIGUES; DUARTE; DIAS, 2017, *online*).

Pesquisas recentes demonstraram algumas habilidades bibliotecárias relacionadas à gestão de dados de pesquisa, que devem se iniciar antes da pesquisa em si, pois, como visto na seção 3, atualmente o PGD é essencial para o desenvolvimento e a gestão de projetos de pesquisas. Nesse sentido, o profissional da Biblioteconomia apresenta competências e habilidades para atuar em diversas fases da gestão de dados de pesquisa. Uma delas, essencialmente, é o plano de gestão de dados, como asseveram Sales, Sayão, Maranhão, Drumond e Silva (2019, p. 307):

A participação dos bibliotecários deve ocorrer já nas fases que precedem a pesquisa, no planejamento, na elaboração de plano de gestão de dados, no apoio na geração/coleta, na identificação de dados de pesquisa de outros pesquisadores, que podem ser reutilizados pelos pesquisadores locais, e na seleção de repositórios adequados à publicação dos dados gerados, incluindo o controle de versionamento, armazenamento e backups, e, na pesquisa finalizada, auxiliando na publicação dos dados e na criação de ambientes confiáveis para armazenamento por longo prazo, caso seja necessário

Portanto, o profissional da Biblioteconomia está apto a auxiliar o(a) pesquisador(a) em sua “pré-pesquisa”. No Brasil, ele(a) pode ser encontrado(a) na Biblioteca Universitária, onde, historicamente, o(a) pesquisador(a) sempre buscou auxílio em diversas bases de dados, em periódicos, livros, entre outros. Atualmente, esse usuário os encontra disponíveis *online* com facilidade. A Biblioteca Universitária, que sempre auxiliou o(a) pesquisador(a) durante sua pesquisa e, em seu pós-pesquisa, assistiu-o com suas publicações acadêmicas, agora, desempenha um papel fundamental no pré-pesquisa, principalmente o(a) amparando-o(a) na preparação de seu PGD e no planejamento da organização de seus dados e metadados, para que ele possa fazer sua pesquisa e controlar facilmente o acesso aos dados dela, para que sejam encontráveis, acessíveis, interoperáveis e reutilizáveis.



Vê-se, nesse cenário, que esse será o futuro das bibliotecas universitárias que vêm desenvolvendo seus repositórios científicos e, brevemente, também deverão atender à demanda dos repositórios de dados de pesquisa, porquanto, no Brasil, as Universidades são importantes fomentadoras de pesquisa científica. Então, no curso natural do desenvolvimento da e-Science e da adesão direta das universidades a esse contexto, a curadoria de dados, que exige o PGD, terá o(a) bibliotecário(a) como o(a) profissional com perfil para auxiliar a organizar, a armazenar e a recuperar os dados das pesquisas.

6 CONSIDERAÇÕES FINAIS

A importância da Biblioteconomia, no contexto da pesquisa orientada para os dados, ainda é incipiente. Apesar disso, já se percebe que ela vem se evidenciando cada vez mais, principalmente na Curadoria de Dados, e atentando para o CVD e a elaboração do PGD, respeitando os princípios FAIR.

O estudo mostrou que o PDG é sobremaneira importante como um dos processos do CVD, que é exigido por agências de fomento, como a FAPESP e algumas instituições de ensino e pesquisa, que existem ferramentas para auxiliar os pesquisadores a elaborarem seus PGD e que o profissional da Biblioteconomia tem competências para orientar tais procedimentos, para se capacitar e auxiliar o(a) pesquisador(a) nessa fase da pesquisa, uma vez que sua formação é direcionada ao armazenamento, à organização, à apropriação e ao uso da informação, no escopo da organização da informação bibliográfica (física), eletrônica e digital, visando à recuperação e à apropriação efetiva dos usuários. Nessa perspectiva, o cerne do PGD consiste em possibilitar que os dados da pesquisa sejam recuperáveis e que possam se apropriar deles.

Conclui-se, então, que a biblioteca universitária sempre foi voltada para pós-publicações, como livros, periódicos, trabalhos acadêmicos etc. só que, agora, ela deverá se preocupar com a pré-publicação, ou seja, com a gestão desses dados advindos da *e-Science*, a fim de que, para além dos resultados das pesquisas, seus dados possam ser consultados e reutilizados, proporcionando economia de tempo e recursos para outros pesquisadores. Logo, considerando o exposto, é possível e justificável reconhecer o(a) bibliotecário(a) como um profissional com as competências e as habilidades necessárias para facilitar a vida do usuário, atuando nesse cenário, por meio da Curadoria de Dados de Pesquisa e do auxílio a pesquisadores(as) na elaboração do planejamento dos dados de sua pesquisa com o PGD.



Para além dessas colocações, e como provocação, convém ressaltar a necessidade de se discutir a respeito da atualização dos currículos dos Cursos de Biblioteconomia no país, a fim de incluir componentes curriculares que garantam competências e habilidades para a atuação do(a) profissional bibliotecário(a) nessas atividades que se revelem promissoras para a formação desses profissionais.

REFERÊNCIAS

ANJOS, Renata Lemos dos. **Análise dos componentes do ciclo de vida dos dados na matriz curricular dos Cursos de Biblioteconomia**: um estudo comparado. 2019. 282 f. Dissertação (Mestrado) - Curso de Ciência da Informação, CCSA, Universidade Federal da Paraíba, João Pessoa, 2019.

BORGMAN, C.L. **Big data, little data, no data**: scholarship in the networked world. Cambridge, Massachusetts; London, England: MIT Press Books, 2015. 383 p.

BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. **10 razões para você criar um plano de gestão de dados de pesquisa**. Elaborada pela Comissão Nacional de Energia Nuclear. Disponível em: <http://antigo.cnen.gov.br/centro-de-informacoes-nucleares/sobre-o-cin/78-cin/dicas-academicas/232-10-razoes-para-voce-criar-um-plano-de-gestao-de-dados-de-pesquisa>. Acesso em: 02 jul. 2021.

BRASIL. Ministério da Educação. CNE/CES 492/2001. Diretrizes Curriculares Nacionais dos Cursos de Filosofia, História, Geografia, Serviço Social, Comunicação Social, Ciências Sociais, Letras, Biblioteconomia, Arquivologia e Museologia. **Diário Oficial da República Federativa do Brasil**, Brasília, 9 jul. 2001, Seção 1e, p. 50. Disponível em: <http://portal.mec.gov.br/cne/arquivos/pdf/CES0492.pdf> Acesso em: 26 mar. 2020.

COSTA, Maira Murrieta; CUNHA, Murilo Bastos da. O bibliotecário no tratamento de dados oriundos da e-Science: considerações iniciais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 3, p. 189-206, set. 2014. Disponível em: <https://www.scielo.br/j/pci/a/VNQTBgB8ZRCJNbWhqbVkpZH/?lang=pt>. Acesso em: 03 jun. 2022

DIGITAL CURATION CENTRE. Edimburgo: DCC. (ed.). **What is digital curation?** Disponível em: <http://www.dcc.ac.uk/digital-curation/what-digital-curation>. Acesso em: 05 maio 2020.

FERGUSON, Adam R; NIELSON, Jessica L; CRAGIN, Melissa H; BANDROWSKI, Anita e; MARTONE, Maryann e. Big data from small data: data-sharing in the 'long tail' of neuroscience. data-sharing in the 'long tail' of neuroscience. **Nature Neuroscience**, [S.l.], v. 17, n. 11, p. 1442-1447, 28 out. 2014. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/nn.3838>. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4728080/citedby/>. Acesso em: 01 maio 2020.



FERREIRA, Valdinéia. Barreto. **E-Science e políticas públicas para ciência, tecnologia e inovação no Brasil [online]**. Salvador: EDUFBA, 2018, 256 p. ISBN: 978-85-232-1865-2.

GRAY, Jim. **E-Science**: um método científico transformado. Transcrição de palestra ministrada por Jim Gray no Conselho Nacional de Pesquisa (EUA), 11 Jan. 2007. In: HEY, Tony;

GUIMARÃES, José Augusto Chaves. Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. **Informação & Informação**, [s.l.], v. 22, n. 2, p. 84-98, 29 out. 2017. Disponível em:

<http://www.uel.br/revistas/uel/index.php/informacao/article/view/31443>. Acesso em: 21 jul. 2020.

LIMA, Gracirlei Maria de Carvalho; ARAÚJO, Carlos Alberto Ávila. Serviço de referência: práticas informacionais do bibliotecário. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 16, p. 1-23, 2020. Disponível em:

<https://rbbd.febab.org.br/rbbd/article/view/1336>. Acesso em: 27 maio 2022.

MARTINS, Gracy Kelli. **Institucionalização cognitiva e social da Organização e Representação do Conhecimento na Ciência da Informação no Brasil**. 2014. 184 f. TESE (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. OECD **Principles and guidelines for access to research data from public funding**. Paris: Oecd Publishing, 2007. 22 p.

OLIVEIRA, Adriana Silva de Oliveira. **Desvendando a autoridade colaborativa na e-Science sob a ótica dos direitos de propriedade intelectual**. 2016. 300 f. Tese (Doutorado em Ciência da Informação) - Curso de Ciência da Informação, CCSA, Universidade Federal da Paraíba, João Pessoa, 2016.

POMERANTZ, Jeffrey. **Metadata**. London: The Mit, 2015. 239 p.

RIBEIRO, Cláudio José Silva. Desafios para a representação documental no âmbito da Web Semântica. In: ZAFALON, Zaira Regina; DAL'EVEDOVE, Paula Regina. **Perspectivas da representação documental**: discussão e experiências. São Carlos: Cpoi/ufscar, 2017. p. 287-302.

RODRIGUES, Adriana Alves; DUARTE, Emeide Nóbrega; DIAS, Guilherme Ataíde. Desafios da gestão de dados na era do Big data: perspectivas profissionais. **Encontro Nacional de Pesquisa em Ciência da Informação**. 2017. Disponível em:

<https://periodicos.ufpb.br/index.php/itec/article/view/40538>. Acesso em 23 mai. 2021.

SALES, Luana Farias; SAYÃO, Luís Fernando. A ciência invisível: revelando os dados da cauda longa da pesquisa. **XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB)**. 2018. GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação. Disponível em: <https://www.brapci.inf.br/index.php/res/v/103678>. Acesso em: 01 mai. 2020.



SALES, Luana Farias; SAYÃO, Luís Fernando; MARANHÃO, Ana Maria Neves; DRUMOND, Geisa Meirelles; SILVA, Maria Helena Ferreira Xavier da. Competências dos bibliotecários na gestão dos dados de pesquisa. **Ciência da Informação**, Brasília, v. 48, n. 3, p. 303-313, set. 2019.

SÃO PAULO. Secretaria de Desenvolvimento Econômico. **Diretrizes para Planos de Gestão de Dados (PGD) para propostas de centros**. Elaborada pela Fundação de Amparo à Pesquisa do Estado de São Paulo. Disponível em: <https://fapesp.br/14974/diretrizes-para-planos-de-gestao-de-dados-pgd-para-propostas-de-centros>. Acesso em: 27 maio 2022.

SAYÃO, LUÍS FERNANDO; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, v. 22, n. 3, p. 179-191, set./dez. 2012. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/12224/8586>. Acesso em: 16 abr. 2020.

SAYÃO, Luís Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. 90 p.

SEMLER, A. R.; PINTO, A. L. Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. **Ciência da Informação**, v. 48, n. 1, 2019. Disponível em: <http://revista.ibict.br/ciinf/article/view/4461/4102>. Acesso em: 29 ago. 2019.

SETZER, V. W. **Dado, informação, conhecimento e competência**. 2015. Departamento de Ciência da Computação, Universidade de São Paulo. Disponível em: <https://www.ime.usp.br/~vwsetzer/dado-info.html>. Acesso em: 30 ago. 2019.

STRASSER, Carly; COOK, Robert; MICHENER, William; BUDDEN, Amber. **Primer on Data Management: what you always wanted to know**. 2012. UC Office of the President: California Digital Library. Disponível em: <https://escholarship.org/uc/item/7tf5q7n3>. Acesso em: 22 jul. 2020.

WILBANKS, John. Ver a mudança de paradigma: e somos nós. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma: descobertas científicas na era da e-Science**. São Paulo: Oficina de Textos, 2011. Cap. 25. p. 217-222. Tradução de Leda Maria Marques Dias Beck.