

# A curadoria de dados científicos na Ciência da Informação: levantamento do cenário nacional

## *Scientific data curation in information science: national scenario survey*

Liliane Chaves de Resende<sup>1</sup>, Marcello Peixoto Bax<sup>2</sup>

<sup>1</sup> Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil. ORCID: <https://orcid.org/0000-0002-4650-0540>

<sup>2</sup> Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil. ORCID: <https://orcid.org/0000-0003-0503-3031>

**Autor para correspondência/Mail to:** Liliane Chaves de Resende, [lilianederesende@gmail.com](mailto:lilianederesende@gmail.com)

**Recebido/Submitted:** 12 de setembro de 2019; **Aceito/Approved:** 16 de agosto de 2020



Copyright © 2020 Resende & Bax. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-Compartilhável 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em <http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice>.

### Resumo

**Introdução:** para a ciência contemporânea, o compartilhamento de dados constitui elemento primordial para seu progresso e colaboração entre comunidades científicas. No contexto da ciência aberta e *e-science*, a ciência da informação investiga soluções para os desafios do tratamento e uso da informação. O desconhecimento atual da importância das atividades de curadoria digital dos dados científicos e do grau de adesão com esse novo cenário, causam dificuldades para o profissional da informação se adequar nesse contexto e contribuir com soluções para os desafios de curadoria digital. **Objetivo:** investigar a importância que as atividades de curadoria digital de dados científicos têm para a área da Ciência da Informação no Brasil, frente à tendência internacional do gerenciamento do conhecimento científico. **Método:** pesquisa de abordagem quantitativa, de natureza aplicada e do tipo exploratória e descritiva. Utiliza procedimentos de uma pesquisa com Survey, para obter informações da curadoria digital por meio da opinião dos pesquisadores brasileiros da área da Ciência da Informação. **Resultados:** As informações analisadas revelam que o campo da Ciência da Informação brasileira está buscando se adaptar nesse cenário. **Conclusões:** os dados permitem afirmar, que ainda não há uma conscientização suficientemente assimilada pelos pesquisadores da Ciência da Informação em relação à necessidade de envolvimento mais consciente e comprometido nas atividades de Curadoria Digital. Novas pesquisas são indicadas para aprofundar esta discussão.

**Palavras-chave:** Curadoria Digital; Dados Científicos; Profissional da Informação; Ciência da Informação; Survey; e-Science.

### Abstract

**Introduction:** for contemporary science, data sharing is a key element for its progress and collaboration among scientific communities. In the context of open science and *e-science*, information science investigates solutions to the challenges of treating and using information. The current lack of knowledge of the importance of digital curation activities of scientific data and the degree of adherence to this new scenario, cause difficulties for the information professional to adapt in this context and contribute solutions to the challenges of digital curation. **Objective:** to investigate the importance that the digital curation activities of scientific data have for the area of Information Science in Brazil, in face of the international trend of scientific knowledge management. **Methods:** The research follows a quantitative approach, it is applied in nature and exploratory and descriptive. It uses procedures of a survey with Survey, to obtain information from the Digital Curator through the opinion of Brazilian researchers in the area of Information Science. **Results:** The information analyzed reveals that the field of Brazilian Information Science is seeking to adapt in this scenario. **Conclusions:** The data allow us to affirm that there is not a sufficiently assimilated awareness by the researchers of Information Science regarding the need for more conscious and committed involvement in Digital Curation activities. New research is indicated to deepen this discussion.

**Keywords:** Digital Curation; Scientific data; Information Professional; Information Science; Survey; e-Science.

## INTRODUÇÃO

A Curadoria Digital de Dados Científicos (CDDC) está ganhando importância junto às instituições de pesquisa. As comunidades acadêmicas estão se conscientizando da necessidade de manutenção da informação científica como recurso para continuidades de pesquisa ou novas oportunidades de descobertas do conhecimento.

Em várias comunidades científicas internacionais, o uso e a prática das atividades de CDDC já vêm sendo realizados em universidades dos EUA, Reino Unido, Europa, Austrália, Ásia e continente africano. Há uma notável participação de bibliotecários nesses projetos, executando atividades com propósito de fornecer suporte aos pesquisadores no gerenciamento dos dados científicos, auxiliando em atividades de armazenamento e disseminação apropriadas dos dados científicos. A execução de atividades necessárias às práticas da CDDC pode transformar as bibliotecas de pesquisas acadêmicas em um locus de gestão e curadoria de dados científicos.

No Brasil, esse cenário ainda está em processo embrionário. São poucas iniciativas das atividades de curadoria digital sendo, efetivamente, implementadas nas comunidades científicas brasileiras. Percebe-se que há grande interesse das instituições em desenvolver a CDDC, mas ainda não se tem conhecimento de uma política sólida, com definições claras sobre como isso poderá se tornar atividade obrigatória nas instituições.

Segundo Kouper (2016), a iniciativa de obter informações sobre Curadoria Digital (CD), diretamente daqueles que pesquisam assuntos da área, amplia o conhecimento existente das principais competências da CD e aprimora

a compreensão dos conhecimentos, valores e experiência cotidianas dos pesquisadores da área de Ciência da Informação (CI). Portanto, pode-se perguntar, no sentido de declarar o problema questionado por esta pesquisa: qual é o cenário da CI frente ao movimento de curadoria digital de dados científicos, em âmbito nacional?

O objetivo desta pesquisa é descrever qual o cenário atual da área de CD no Brasil por meio da opinião e atitudes dos pesquisadores da área de CI em relação às atividades básicas de armazenamento, uso e reuso de dados brutos produzidos em suas pesquisas.

## BREVE PANORAMA DA CURADORIA DIGITAL

A curadoria digital (CD) é uma disciplina emergente, aplicada e interdisciplinar, com dimensões acadêmicas e profissionais, progredindo a partir da dimensão técnica da informação digital como objeto de pesquisa.

A CD é voltada para o ambiente digital cuja função específica é “cuidar” de recursos digitais. Como objetivo principal, busca desenvolver estratégias para resolver problemas do fluxo da informação digital. Essa é sua atividade central. Contudo, existem implicações que são causadas pela multiplicidade e diversidade de elementos envolvidos com a informação digital. Isso faz com que diferentes níveis de curadoria sejam apropriados para diferentes produtores e consumidores da informação digital (Higgins, 2018; National Research Council, 2015).

A evolução da área da CD cresceu de práticas originadas da curadoria tradicional, surgindo juntamente com a definição de *e-Science*, ou ciência eletrônica, definida por Gray, Szalay, Thakar, e Stoughton (2002), revolucionando o formato da investigação científica. Para Gray, os pesquisadores especialistas de um determinado domínio que queiram executar atividades de curadoria deveriam aprender, também, conceitos e técnicas relevantes utilizadas por bibliotecários como, por exemplo, as técnicas de anotação, descrição e preservação dos dados (Gray et al., 2002; National Research Council, 2015).

Kouper (2016) afirma que a área de CD cresce na interseção entre biblioteconomia, arquivística e tecnologia da informação e suas práticas requerem uma combinação de habilidades oriundas dessas áreas e de novas a serem definidas. A competência desejada inclui familiaridades técnicas para uso das tecnologias digitais e ferramentas de CD, a capacidade de iniciar e gerenciar projetos, programas e serviços, com fortes habilidades interpessoais e organizacionais.

Uma contribuição para a evolução da CD se deu com a publicação do Modelo de Referência do Sistema de Informação em Arquivamento Aberto (OAIS), em 2003, pela Organização Internacional para Padronização (ISO). Esse modelo, embora não prescreva diretamente práticas de CD, propõe uma estruturação de responsabilidades, processos e funções de arquivamento, preservação e acesso a longo prazo da informação digital.

Portanto, a Curadoria de Digital de Dados Científicos (CDDC) surge nesse cenário como oportunidade para a área de CI e profissionais da informação. Atividades de CDDC constituem um recurso que transforma o ciclo da comunicação científica, corroborando com novas descobertas de conhecimentos por meio da colaboração entre pesquisadores (Sales & Sayão, 2012) apesar de apresentar-se como um desafio complexo frente as demandas do gerenciamento dos dados científicos brutos produzidos pelas pesquisas.

A proposta da CDDC é, especificamente, atribuir valor aos dados produzidos pelas pesquisas científicas. Esse valor agregado se refere, principalmente, à manutenção de dados que são impossíveis de recriar, objetivando os mais diversos fins: uso no ensino, validação dos resultados de pesquisa publicados, analisando os que podem ou não ser descartados, mantendo sua integridade e qualidade para reuso futuro em outras pesquisas, podendo gerar novos conhecimentos (Poole, 2016).

O National Research Council dos EUA, em seu estudo intensivo da CD, concluiu que a demanda por informações digitais prontamente acessíveis, precisas, úteis e utilizáveis, fez surgir limitações relacionadas ao gerenciamento do dado digital. E para resolver isso é necessário o desenvolvimento de políticas, serviços, tecnologias e conhecimento em CD, como forma de articular melhores oportunidades para reduzir custos e aumentar benefícios para a toda sociedade do conhecimento (National Research Council, 2015).

De acordo com Nielsen e Hjørland (2014), os dados digitais de pesquisa são diferentes de documentos impressos. A manutenção de tais conjuntos de dados requer conhecimento científico do domínio e conhecimento avançado em tecnologia para organizá-los e armazená-los, com vistas à preservação e à reutilização. Para os autores, “dados são frequentemente compreendidos como matéria-prima do processamento de informação e aquisição de conhecimento” (Nielsen & Hjørland, 2014, p. 3), sendo necessário um conhecimento específico sobre como gerenciar dados de pesquisas de um determinado domínio. As várias prioridades, considerações e estimativas a partir das perspectivas de domínio que levam a uma exigência de competências mais profundas de gestores e curadores.

Nesse sentido, a CDDC poderá contribuir com a demanda do gerenciamento dos dados produzidos pelas pesquisas, propondo alternativas para sua preservação a longo prazo, garantindo acesso e reuso. Inclusive, repositórios institucionais já são considerados como principais instrumentos para disseminação do conhecimento, apoiando a criação de novas descobertas científicas (Sales & Sayão, 2012).

As bibliotecas de pesquisa também estão se movimentando para desenvolver novos papéis e implementar novos modelos de serviço. Bibliotecários com funções específicas de tecnologia da informação podem assumir funções de liderança do gerenciamento de dados científicos, atuando em parceria com os pesquisadores (Cox, Searle, Wolski, Simons, & Richardson, 2015).

O avanço científico proporciona, cada vez mais, a necessidade perante o acesso a dados de pesquisa, causando impactos à eficiência e à eficácia das atividades científicas. Nesse sentido, bibliotecários precisarão se tornar especialistas de domínio, propondo meios e recursos para as necessidades de pesquisa em diferentes domínios, formando uma parceria com pesquisadores. Poderão auxiliá-los em questões da articulação de infraestrutura de pesquisa dentro das instituições, e em atividades de curadoria de dados (Lee & Stvilia, 2017) (Nielsen & Hjørland, 2014, p. 3).

Os repositórios institucionais já são considerados como principais instrumentos para disseminação do conhecimento, apoiando novas descobertas científicas. Sobretudo com o movimento da Ciência Aberta, que proporcionou a democratização da informação científica. Mas sua criação ainda é recente e seu uso pouco disseminado (Shintaku, Duque, & Suaiden, 2015).

A necessidade da integração das atividades de CD nas atividades dos pesquisadores pode transformar a CDDC como parte da missão dos profissionais da informação da área da CI. Lee e Stvilia (2017) entrevistaram profissionais que atuam em repositórios institucionais e descobriram que existem diferentes níveis nas atividades de CDDC executadas em repositórios institucionais e que um encontro inicial com o pesquisador é muito importante para definir a extensão das necessidades de CDDC. A exemplo dessas atividades, estão as questões de conceituação, planejamento, criação, carregamento ou *upload*, armazenamento e publicação dos dados científicos. Durante todo esse processo, o pesquisador é necessário e irá interagir com vários profissionais, principalmente bibliotecários, arquivistas, profissionais de TI, atuando de acordo com cada atividade das etapas do ciclo de vida do gerenciamento de dados científicos. É por meio dessa interação que as necessidades de CDDC serão identificadas (Lee & Stvilia, 2017).

O papel dos repositórios institucionais, nesse sentido, se alinhou à visão tradicional das bibliotecas em salvaguardar e ofertar acesso à informação. Contudo, apresenta uma potencialidade ainda maior que a biblioteca tradicional por proporcionar apoio, geração, preservação e disseminação da informação científica. Por isso, o uso de repositórios digitais e das bibliotecas de pesquisa para organizar dados científicos é uma decorrência necessária e altamente indicada para as instituições acadêmicas (Lee & Stvilia, 2017; Shintaku et al., 2015; Tibbo & Hank, 2015).

No cenário brasileiro, a CDDC ainda se encontra em um processo de implantação recente e embrionário. Algumas iniciativas das atividades de curadoria digital estão sendo efetivamente implementadas nas comunidades científicas brasileiras e governamentais. Percebe-se que há grande interesse das instituições em desenvolver a CDDC, mas ainda não se tem conhecimento de uma política sólida, com definições claras sobre como isso poderá se tornar atividade obrigatória nessas instituições.

Durante a execução dessa pesquisa, pôde-se levantar um total de nove repositórios especificamente voltados para armazenamento, uso e reuso de dados científicos no Brasil, disponibilizados pela Re3data ([www.r3data.org](http://www.r3data.org)), um registro global de repositórios de dados científicos que inclui conjuntos de dados para pesquisadores, órgãos financiadores, editores e instituições acadêmicas. O objetivo da Re3data é promover a cultura de compartilhamento e acesso a dados de pesquisa. A exemplo de alguns projetos em andamento no Brasil visando o mesmo objetivo, podemos citar:

- A plataforma CarpeDIEN de Informações em Energia Nuclear do Instituto de Engenharia Nuclear (IEN/CNEN), vinculada ao Conselho Superior de Pesquisas Científicas da Espanha (CSIC) desde 2014, propõe integrar instituições de pesquisas na área nuclear, com mais de 45 instituições brasileiras já participantes, dentre elas a USP, Fiocruz e a EMBRAPA. A finalidade dessa plataforma é promover a preservação, a curadoria e a disseminação da memória digital técnico-científica produzida pelas instituições parceiras. Tem como participação a pesquisadora Dra. Luana Sales, bibliotecária e uma das idealizadoras da plataforma;
- O Centro de Documentação e Acervo Digital da Pesquisa (CEDAP), órgão auxiliar da Faculdade de Biblioteconomia e Comunicação (FABICO) da Universidade Federal do Rio Grande do Sul (UFRGS), cuja finalidade é fornecer suporte às pesquisas científicas institucionais. Oferece vários serviços para armazenamento, acesso e compartilhamento de dados científicos;
- A iniciativa da Rede de Dados científicos Brasileira (RDP Brasil) é uma parceria da Rede Nacional de Ensino e Pesquisa (RNP) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em colaboração com a Universidade Federal do Rio Grande do Sul (UFRGS) e a Universidade Federal do Rio Grande (FURG). Tem como finalidade desenvolver um projeto conjunto sobre o Acesso Aberto a Dados de Pesquisa (AADP), cuja finalidade é desenvolver atividades que contribuam com a identificação das práticas do acesso aberto a dados de pesquisa em instituições brasileiras, mapeando requisitos e desenvolvendo uma prototipação de sistemas de AADP para levantar uma comparativo de serviços e soluções tecnológicas de

AADP que visa facilitar a divulgação de informações científicas;

- O Repositório da Produção Científica CRUESP, do Conselho de Reitores das Universidades Estaduais de São Paulo – USP, UNICAMP e UNESP, cuja finalidade é reunir, preservar e proporcionar acesso aberto, público e integrado à produção científica de docentes, pesquisadores, alunos e servidores das universidades participantes. Integradas ao CRUESP, estão os repositórios institucionais de produção intelectual (científica, artística, acadêmica e técnica) da USP, a Biblioteca Digital da Produção Intelectual da Universidade de São Paulo (BDPI), a Biblioteca Virtual da FAPESP, a Biblioteca Digital da Produção Intelectual e Científica da UNICAMP e o Repositório Institucional da UNESP.

Além desses projetos, já existem no Brasil propostas de conscientização da participação dos profissionais da informação em projetos de CDDC. Mas é importante descobrir qual o engajamento dos pesquisadores da área da Ciência da Informação perante esse processo e como estão se organizando para contribuir com o desenvolvimento da área de CD frente às demandas do campo científico, como já vem acontecendo em outros países. Partindo desse pressuposto, obter opiniões acerca das atividades de CDDC executadas por pesquisadores brasileiros da área da CI é imprescindível para que seja possível descrever como está evoluindo a área de CD no Brasil e qual seu estado atual, especificamente para o campo da CI.

## MÉTODO

A pesquisa segue abordagem quantitativa, é de natureza aplicada, do tipo exploratória e descritiva. Utiliza-se de uma pesquisa com *Survey* para obter informações da Curadoria Digital por meio da opinião dos pesquisadores brasileiros da área da Ciência da Informação. O quadro 1 abaixo apresenta a caracterização da pesquisa com *Survey*:

<b>Técnica de pesquisa</b>	Pesquisa com <i>Survey</i> exploratório.
<b>Métodos</b>	Pesquisa de levantamento de opiniões.
<b>Objetivo</b>	Descrever a população investigada num determinado momento para analisar o contexto da CDDC em âmbito nacional.
<b>Natureza</b>	Quantitativa, exploratória e descritiva.
<b>População</b>	Pesquisadores da área da ciência da informação participantes de programas de pós-graduação stricto sensu nas universidades brasileiras recomendados pela CAPES ao Conselho Nacional de Educação – CNE/MEC, registrados na plataforma Sucupira.
<b>Amostra</b>	Pesquisadores na área de CI que responderam ao questionário on-line.
<b>Seleção da amostra</b>	Amostragem de probabilidade aleatória (igual probabilidade de seleção da população), amostragem auto selecionada.
<b>Instrumento de coleta</b>	Questionário <i>on-line</i> .
<b>Métodos de análise</b>	Análises estatísticas por meio de frequências absolutas das respostas obtidas, descrevendo atitudes, preferências e tendências da população.
<b>Validade do instrumento</b>	Aplicação do questionário como pré-teste intencional com um desenho de amostragem aleatória aplicado a 3 pesquisadores.

**Quadro 1.** Caracterização da Pesquisa com *Survey*.

Fonte: Elaborado pelos autores (2019).

A população alvo foi composta por todos os pesquisadores brasileiros na área de CI, vinculados a programas de pós-graduação stricto sensu (Mestrado e Doutorado) na área de Ciência da Informação, selecionados na Plataforma Sucupira da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), com chances iguais de participação.

Neste estudo, o método de amostragem utilizado foi amostragem aleatória simples. A amostragem foi auto selecionada, ou seja, composta por pesquisadores que aceitaram o convite em participar do levantamento. De acordo com Oliveira e Grácio (2005), também é necessário garantir que a amostra apresente as mesmas características gerais da população. Sendo assim, o cálculo do tamanho da amostra é determinado com base na quantidade de elementos necessários para compor a amostra, a fim de obter resultados válidos.

Os pesquisadores foram contatados por meio de mensagens de correio eletrônico (*e-mail*) enviadas para as suas respectivas caixas postais, individualmente. Os *e-mails* foram obtidos por meio de informações públicas, disponibilizadas nas páginas oficiais dos programas de pós-graduação na área da CI ou por meio de artigos publicados em periódicos de acesso aberto. Os questionários foram enviados, por *e-mail*, aos 390 membros da população-alvo. Cerca 78 elementos da população-alvo ficaram impossibilitados de participar por não ser possível localizar referência de *e-mail* nos sites institucionais, de acesso público.

De acordo a taxa de respondentes obtida, por meio de 97 respondentes, pôde-se obter uma abrangência representativa de pesquisadores na pesquisa. Marconi e Lakatos (2003) afirmam que questionários enviados para os entrevistados alcançam, em média, 25% de devolução. Oliveira e Grácio (2005) sugere que, para que uma amostragem seja representativa, ela deve abranger uma porcentagem fixa da população, de, aproximadamente, 10% a 20% e que essa porcentagem deve representar, pelo menos, de 30 a 40 elementos da população, caso contrário é considerada muito pequena. Portanto, o quantitativo obtido dos respondentes sugere uma amostragem representativa e válida, uma vez que se encontra dentro da estimativa sugerida pelos autores, ou seja, 25% da população-alvo, e suficiente para generalizar os achados.

Os dados foram coletados durante os meses de maio, junho, julho e agosto de 2019, por meio de um link para acesso ao formulário *web* do Google Forms. Optou-se pelo uso dessa ferramenta devido a sua simplicidade, facilidade de uso e agilidade para a obtenção dos dados, uma vez que é perfeitamente possível usar essa ferramenta para obter um número razoável de dados, em um período de tempo curto e em diferentes áreas geográficas, sem custos. Até mesmo para possibilitar maior flexibilidade ao respondente (Gil, 2008).

Utilizou-se uma versão gratuita de teste do software SPSS *Statistics Subscription (Statistical Package for the Social Sciences* – para tabular as respostas obtidas, que foram analisadas por meio de suas frequências absolutas. As questões elaboradas no instrumento de pesquisa utilizaram 26 questões de múltipla escolha e duas questões abertas (Apêndice 1).

Esta pesquisa demanda uma compreensão inicial básica dos conceitos de CDDC por parte dos respondentes. Dada a natureza emergente da disciplina, esse pode ter sido um fator que limitou ou não tenha motivado a participação de mais respondentes na pesquisa.

## RESULTADOS

Para descrever o estado atual da CD no cenário brasileiro, foram obtidos dados para: 1) caracterização do perfil dos pesquisadores brasileiros na área de CI que atuam em programas de pós-graduação, mestrado e doutorado; 2) o nível de conhecimento desses pesquisadores sobre CD; 3) o nível de envolvimento nas atividades de CDDC quando desenvolvem suas pesquisas; 4) tendências e opiniões desses pesquisadores a respeito da CDDC.

### Perfil do pesquisador brasileiro na Ciência da Informação

Para apresentar o perfil do pesquisador na área de CI foram levantados os dados pessoais, acadêmicos básicos e atividades de CD. Os dados obtidos são apresentados na Tabela 1 abaixo.

Resumidamente, o participante tem idade acima de 40 anos, é em maioria do gênero feminino e quase todos possuem título de doutorado, com mais de 10 anos de experiência em pesquisas acadêmicas. Atua principalmente em universidade do setor público; e a maioria busca fomento à pesquisa junto às entidades nacionais. Na totalidade, 58,2% interagem com outras universidades brasileiras e 32,8% com universidades internacionais, demonstrado relativamente baixa internacionalização. Quanto à formação científica, é clara a participação de graduação em outras áreas do conhecimento.

		N	% Casos
<b>Gênero</b>	Feminino	57	58,8
	Masculino	39	40,2
	Não informou	1	1,0
<b>Titulação acadêmica</b>	Doutorado	85	87,6
	Mestrado	12	12,4
<b>Faixa etária</b>	21 a 30	1	1,0
	31 a 40	24	24,7
	41 a 50	26	26,8
<b>Experiência de pesquisa (anos)</b>	2 a 5	16	16,5
	6 a 10	28	28,9
	11 a 20	22	22,7
	Mais de 20	31	32,0
<b>Lócus de atuação</b>	Setor público	44	46,3
	Setor privado	21	22,1
<b>Lócus de pesquisa</b>	Local de trabalho	80	84,2
	Universidades brasileiras	56	58,9
	Universidade internacional	29	30,5
	Com Institutos de pesquisa	36	37,9
	Org. não governamental	11	11,6
	Entidade governamental	1	1,1
<b>Lócus de financiamento</b>	Fomento nacional	33	34,7
	Fomento internacional	5	5,3
<b>Áreas de conhecimento da graduação</b>	Ciências Exatas	13	13,4
	Ciências Humanas	12	12,4
	Engenharias	9	8,1
	Ciências da Saúde	2	2,1
	Linguística, Letras e Artes	7	7,2
	Ciências Sociais Aplicadas	68	70,1

Tabela 1. Perfil dos respondentes.

Fonte: Elaborada pelos autores (2019).

### Envolvimento nas atividades de curadoria digital de dados científicos

Com o propósito de descrever se o participante possui alusão e conhecimento básico de curadoria digital, 93% afirmam que conhecem a definição geral de CD e quase todos (82%) nunca participaram da elaboração de políticas ou normas nacionais voltadas para CD. Percebe-se que, pelas sugestões apresentadas por alguns pesquisadores, já se tem uma movimentação para utilização de repositórios de dados de pesquisa, principalmente em repositórios institucionais, porém 50,5% informaram nunca ter utilizado esse recurso. Quanto à participação em eventos específicos do assunto, 49% disseram ter participado, mas poucos (18%) pesquisadores já participaram de capacitações ou treinamentos específicos da área.

Atividades básicas de CD abrangem diversas práticas que possibilitam a administração e o gerenciamento da informação digital, dentre elas: coletar, descrever, representar, arquivar e preservar dados (Thompson, Senseney, Baker, Varvel, & Palmer, 2013). Em relação a essas atividades, quase todos (99%) os pesquisadores informaram que já armazenaram, 79,4% acessaram, 60,8% reutilizaram e 62,9% compartilharam seus próprios dados de pesquisa em algum momento.

Em média, 51,9% dos pesquisadores afirmam armazenar seus dados científicos em dispositivos pessoais. Entretanto, a quantidade de pesquisadores que já disponibilizaram seus dados de pesquisa foi de 46,4%, e poucos (1%),

informaram já ter publicado seus dados brutos de pesquisa. Os dados obtidos são apresentados na Tabela 2.

Algumas informações se diferenciam quanto ao quesito da utilização de repositórios digitais para armazenar e reutilizar dados de pesquisa. Cerca de 98% dos pesquisadores confirmaram já ter utilizado repositórios digitais. Essa informação sugere que os pesquisadores têm conhecimento da importância dos repositórios de dados para a pesquisa científica.

Todavia, percebe-se que é preciso desenvolver uma efetiva política nacional para a gestão e compartilhamento de dados científicos, uma vez que 46,4% afirmaram nunca ter utilizado repositório digital para armazenar seus próprios dados de pesquisa, mesmo guardando-os em computador pessoal (70%). Apenas 24,7% afirmam disponibilizá-los e 2% até os descartam quando terminam a pesquisa. A prática de compartilhamento de dados não é uma atividade comum para os pesquisadores brasileiros (Sayão & Sales, 2016; Vanz et al., 2018).

Em relação à reutilização de dados de pesquisa, próprios e de outrem, confirmam igualmente o interesse entre 60,8% dos pesquisadores. Os fatos sugerem o entendimento de que a Ciência da Informação "é ambiente colaborativo, aberto e digital" (Oliveira & Silva, 2016, p. 11).

Porém, mesmo sabendo da tendência do acesso aberto a dados científicos, 46,4% disponibilizam abertamente. E quando investigados se já citaram dados de outros pesquisadores, 49% informaram que sim e 47% afirmam que já publicaram seus próprios dados. Os pesquisadores (76,3%) também concordam que estudos que disponibilizam dados de pesquisa podem ser mais citados.

		N	%
<b>Curadoria digital é oportunidade ao profissional da informação?</b>	Discordo fortemente	1	1,0
	Discordo	1	1,0
	Sem opinião	5	5,2
	Concordo	43	44,3
	Concordo fortemente	47	48,5
<b>Participação eventos</b>	Não	47	48,5
	Sim	49	50,5
	Não sei dizer	1	1,0
<b>Participação capacitação</b>	Não	79	81,4
	Sim	18	18,6
<b>Elaboração de políticas normas nacionais</b>	Não	83	85,6
	Sim	13	13,4
	Não sei dizer	1	1,0
<b>Gerência de dados científicos produzidos</b>	Armazeno em PC pessoal	22	22,7
	Armazeno no Google Drive	31	32,0
	Ficam com os alunos	44	46,3
	Armazeno em HD externo	3	3,1
	Descarto	2	2,1
	Plataforma do GP*	1	1,0
	Revista submeti artigo	1	1,0
	Disponibilizo em site próprio	1	1,0
	Disponibilizo Repositórios	1	1,0
	Plataforma parceira trabalho	2	2,1
	Uso Acervos Institucionais	1	1,0
Plataforma do IE do trabalho	21	21,6	
<b>Citou dados científicos</b>	Sim	49	50,5
	Não	46	47,4
	Não sei dizer	3	3,1
<b>Publicou dados científicos</b>	Sim	13	13,4
	Não	12	12,4
	Não sei dizer	9	8,21

		N	%
<b>Dados científicos produzidos pelo pesquisador</b>	Publica junto com os artigos	54	55,7
	Divulga em eventos científicos	48	49,5
	Armazena computador pessoal	68	70,1
	Disponibiliza em repositório	24	24,7
	Descarta os dados	2	2,1
	Outro	8	8,1
<b>Uso de padrão de metadados</b>	DataCite Metadata Schema	4	4,1
	DDI	4	4,1
	Dublin Core	40	41,2
	ISO 19115	2	2,1
	Outro	2	2,1
	Não sei dizer	50	51,5
<b>Compartilhar dados científicos em Plataforma tecnológica de armazenamento</b>	DSpace	44	45,4
	Zenodo	7	7,2
	Dataverse	8	8,2
	FigShare	5	5,1
	Tainacan	2	2,1
	CKAN	5	5,1
	EPrints	15	15,5
	Nunca utilizei	45	46,4
	Outros	6	6,2
<b>AÇÕES sobre dados produzidos na pesquisa</b>	Armazenou	96	99,0
	Acessou	77	79,4
	Compartilhou	61	62,9
	Disponibilizou	45	46,4
	Reutilizou	59	60,8
	Armazeno p/uso próprio	1	1,0
	Compartilho com GP*	1	1,0
Publiquei	1	1,0	
<b>Repositórios brasileiro de dados</b>	BCD/ UFPR	35	36,1
	CIS	14	14,4
	GLOBE	14	14,4
	IBGE	70	72,2
	IBICT	61	62,9
	Outros	16	16,3
<b>REUSO de dados em Plataforma tecnológica</b>	DSpace	41	42,3
	Dryad	1	1,0
	Zenodo	9	9,3
	Dataverse	9	9,3
	FigShare	6	6,2
	Tainacan	4	4,1
	CKAN	6	6,2
	EPrints	14	14,4
	EUDAT	2	2,1
	Nunca utilizei	49	50,5
	BDC/UFPR	1	1,0
	Dadosabertos.info	1	1,0
	Webmuseum	1	1,0
	Wordpress	1	1,0
	Meus próprios artigos	1	1,0
	Comunidade FIOCRUZ	1	1,0
<b>Modelo de ciclo de vida como referência</b>	DCC	17	17,5
	DDI	2	2,1
	UK Data Lifecycle	3	3,1
	CVD-CI	2	2,1
	DataONE	7	7,2
	Research360	1	1
	OAIS	14	14,4
	Cedido IE de trabalho	2	2,1
	Instituição de fomento	11	11,3
	Não sei dizer	46	47,4
	Nenhum modelo	3	3,1
	Outro	3	3,1



	N	%
<b>Pesquisas são mais citadas quando disponibilizam os dados</b>	Concorda fortemente	22 22,7
	Concorda	52 53,6
	Sem opinião	15 15,5
	Discordo	8 8,2
	Discordo fortemente	0 0

**Tabela 2.** Conhecimentos básicos sobre curadoria digital de dados científicos.

Fonte: Elaborada pelos autores (2019).

Poole (2016) argumenta que a citação e a publicação de dados científicos ainda geram muitas discussões. O pesquisador afirma que a citação de dados científicos é importante porque o conjunto de dados gerados nas pesquisas são considerados produtos de pesquisa de primeira classe e que devem estar disponíveis pública e perpetuamente. Citar dados gera outros desafios como granularidade, micro atribuição, identificadores de contribuições e colocação de citações (Poole, 2016).

Os dados obtidos confirmam essa visão que é justificada com argumento de que “no Brasil a falta de instrumentos e diretrizes uniformes para a preservação e segurança de dados de pesquisa – a matéria prima do empreendimento científico” (Vanz et al., 2018, p. 21). Oliveira e Silva (2016) alegam que o avanço no arcabouço normativo-legal brasileiro, referente aos repositórios institucionais, é lento e que fatores restritivos estão além de questões tecnológicas e técnicas; assentam-se, principalmente, em âmbitos políticos, legais, econômicos e culturais.

As evidências em relação à disponibilização dos dados também corroboram a afirmativa de Vanz et al. (2018) pelo baixo percentual de respondentes, 24,7% que afirmam disponibilizar seus dados em repositórios de dados de pesquisa. Esse valor é ainda bastante inferior ao percentual de respondentes que reutilizam dados de pesquisa, 66,6% (Oliveira & Silva, 2016; Vanz et al., 2018).

Percebe-se que falta um entendimento sobre as práticas de CD de fato (compartilhar, disponibilizar, publicar, citar). Sobretudo porque 46,4% afirmam nunca ter utilizado repositórios para compartilhar seus dados de pesquisa, 51,5% não sabem dizer ou não utilizam padrão de metadados para descrever seus dados científicos e nem sabem dizer qual modelo de referência utilizar na elaboração do Plano de Gerenciamento de Dados de Pesquisa (PGP).

Acerca das técnicas relacionadas às atividades de CDDC sobre descrição de dados científicos, Poole (2016) esclarece que estudos apontam que pesquisadores não possuem familiaridade com criação ou documentação de metadados, e que não conseguem prever as necessidades daqueles que reutilizariam seus dados. Ele também reforça que metadados estruturados são essenciais para o compartilhamento e a reutilização dos dados. É importante que pesquisadores de CI compreendam que essas práticas de CDDC desempenham uma função essencial no gerenciamento dos dados científicos e, por meio delas, podem auxiliar outros pesquisadores em suas pesquisas (Poole, 2016; Tripathi, Shukla, & Sonkar, 2017).

Portanto, para avançar com a CDDC, são necessários esforços em várias instâncias, envolvendo, principalmente, “ação e decisão diferenciadas, internas e externas à ciência, que vão desde o pesquisador individual e equipes de pesquisa até o nível macro das políticas públicas e das regulações internacionais, passando pelo nível mesmo das instituições científicas e agências de fomento” (Abagli, 2015, p. 21).

### Preferências e opiniões da curadoria digital de dados científicos

A Tabela 3 apresenta dados dos participantes. Em relação às preferências dos pesquisadores para a área da CDDC, considerou-se que questões éticas nos processos da pesquisa científica são mais importantes no desempenho de atividades de CD (89,7%). Em torno de 75,3% dos pesquisadores identificaram que a garantia de usabilidade e de acessibilidades dos dados científicos é uma atividade de CD importante para o profissional da informação, e que encontrar dados e publicações com finalidade de reuso é a principal atividade que o profissional da informação pode executar para auxiliar outros pesquisadores em suas pesquisas.

Os pesquisadores afirmam que para um profissional da informação atuar em repositórios digitais de dados científicos deve dominar, principalmente, habilidades importantes sobre dados científicos, tais como compreender marcos legais, regulatórios e políticas de direito autoral, as formas de acesso a dados digitais e questões relacionadas à propriedade intelectual, com 86,6%, 87,6% e 84,5% respectivamente das respostas obtidas.

Os participantes consideraram como questão mais crítica para a pesquisa brasileira em CD a sustentabilidade e a manutenção dos dados científicos associadas à evolução, ao desenvolvimento e às mudanças tecnológicas constantes, mas consideraram pouco crítico prestar atenção à educação, ao treinamento e ao desenvolvimento de profissionais que possam dar suporte às atividades de CD.

		N	% Casos
Atividades que o profissional da informação pode auxiliar o pesquisador	Armazenar e preservar dados	52	53,6
	Elaborar um plano de gerenciamento de dados	45	46,4
	Sobre questões legais relacionadas aos dados de minhas pesquisas	42	43,3
	Compartilhar dados	45	46,4
	Criação de metadados	47	48,5
	Encontrar dados e publicações para reuso	53	54,6
	Citações de dados	45	46,4
	Não sei dizer	3	3,1
	Concordo com todas as anteriores	43	44,3
	Manipulação dos dados e geração de visualizações	1	1,0
	Marcos éticos e legais da CD	2	2,1
	Gerenciamento do ciclo de vida / continuum dos objetos digitais, talvez até mesmo a criação do sistema de manutenção de registros desses materiais	61	62,9
	Atividades de CD importantes para o profissional da informação	Envolvimento ativo ao longo do tempo e a longo prazo dos criadores dos dados científicos	48
Avaliação e seleção dos dados científicos para disponibilidade em repositórios digitais		68	70,1
Desenvolvimento e provisão do acesso aos dados científicos		53	54,6
Garantir a preservação (usabilidade e acessibilidade) dos dados científicos		73	75,3
Análise de dados e visualização		1	1,0
Não sei		2	2,1
Armazenar e preservar dados		87	89,7
Habilidades para atuação em repositório digital	Processos ética da pesquisa científica	79	81,4
	Métodos de pesquisa	80	82,5
	Formas de Comunicação científica	82	84,5
	Propriedade intelectual	82	84,5
	Formas de acesso a dados digitais	85	87,6
	Padrões de metadados	81	83,5
	Marcos legais e regulatórios e Políticas de direito autoral	84	86,6
	Todos as opções anteriores	72	74,2
	Organização da informação e do conhecimento	1	1,0
	Conhecimento das práticas e dos modelos internacionais	1	1,0

Tabela 3. Preferências e Opiniões da CDDC.

Fonte: Elaborada pelos autores (2019).

### O cenário da curadoria digital de dados científicos na opinião dos pesquisadores da CI

Sobre o estágio de evolução da CDDC no Brasil, os pesquisadores afirmaram que a CDDC está na sua forma embrionária, dando os primeiros passos. Afirmações de que a CDDC está “precisando de mais estudos que a desenvolvam” constituem uma amostra desse cenário.

Um dos respondentes menciona que ainda não há uma política nacional para gestão e compartilhamento de dados científicos. Mas que há movimentos governamentais para discutir e desenvolver uma minuta de Decreto para a ciência aberta no Brasil, por meio da qual se tem discutido o papel das agências de fomento, o acesso a recursos internacionais para desenvolvimento de pesquisa, a disponibilização dos dados em repositórios internacionais.

Portanto, de acordo com os pesquisadores da CI, parece longo o caminho a ser percorrido para que a área de CI assuma a CD como parte de sua missão. Contudo as respostas indicaram que a CDDC no Brasil, e não apenas na CI, está “em sua fase de expansão, apesar de não ser amplamente discutida dentro dos espaços institucionais, como também em disciplinas específicas” e “se tornando um assunto que necessita de atenção dos pesquisadores, pois a organização e acesso aos dados abertos de objetos digitais é uma necessidade eminente”.

## CONSIDERAÇÕES FINAIS

Esta pesquisa apresentou o *Survey* para levantar e descrever o cenário atual da Curadoria Digital de Dados Científicos (CDDC) para o campo da Ciência da Informação (CI) no Brasil, por meio de atitudes, opinião e preferências de seus representantes pesquisadores.

Contribui com a apresentação do cenário de atuação profissional e acadêmica dos pesquisadores da CI em CDDC. Analisa como o campo da CI, por meio de seus representantes pesquisadores, percebe o papel importante que tem frente a este fenômeno e também como ela está acompanhando e interagindo com o movimento internacional de CDDC. A partir dessa apresentação pode-se iniciar uma discussão mais aprofundada acerca de estratégias para incrementar a relevância das atividades CDDC realizadas por esses profissionais, e, conseqüentemente, pela própria CI.

Por meio dos dados coletados e da opinião dos pesquisadores, infere-se que há uma movimentação inicial para adesão à curadoria digital de dados científicos. Mas, percebe-se um entendimento de a CDDC tem como missão muito além de simplesmente disponibilizar e preservar dados, mas saber administrar e gerenciar a forma de como os dados podem ser reutilizados, contribuindo com novas descobertas científicas. O levantamento nos permite concluir que há baixa atuação dos pesquisadores sobre o tema.

Constatou-se claramente que ainda não há uma efetividade de práticas de CD dos próprios dados produzidos por esses pesquisadores. Claramente observa-se que não há suficiente envolvimento com a CDDC por parte dos pesquisadores para poder afirmar que a CI nacional tomará a CDDC como um pilar ou uma parte fundante de sua missão. Para isso, é necessário um domínio dessas atividades e dos conceitos envolvidos para realização das mesmas.

Assim, concluiu-se, de forma geral, ser ainda necessária uma mudança evolutiva considerável na formação disciplinar teórica, prática e técnica desses pesquisadores. Tal conclusão é ainda mais relevante para o fortalecimento da CI enquanto área de conhecimento como um todo, pois são esses mesmos pesquisadores que formam os futuros profissionais da informação que irão atuar nas práticas profissionais e em serviços técnicos de informação.

Uma limitação desta pesquisa é que a participação dos pesquisadores da CI nacional poderia ter sido ainda maior e também mais abrangente. Acredita-se que os objetivos inicialmente propostos, mesmo que preliminarmente, foram atingidos.

Como trabalhos futuros, sugere-se novas pesquisas que procurem gerar informações ainda mais detalhadas para subsidiar e aprofundar esta discussão, complementando, corroborando ou refutando os resultados alcançados até aqui. O levantamento detalhado das informações descritivas do cenário da CI frente à CDDC permitirá que se estabeleçam estratégias de evolução do tratamento deste tema pela área da CI no país. Permitirá ainda discutir, de forma embasada empiricamente, sobre lacunas da formação disciplinar do profissional da informação, tanto em nível de graduação quanto pós-graduação.

Outra sugestão poderia ser a realização de pesquisas que aprofundem em temas mais específicos no domínio da CDDC, tais como a integração, o reuso, a publicação e a preservação de dados científicos; o uso de ferramentas de repositórios de dados científicos e não apenas textos (artigos, monografias, teses e dissertações). O levantamento de informações mais detalhadas sobre tais atividades e ferramentas permitirá que a CI reflita sobre estratégias e meios para conscientizar a comunidade acadêmica para a importância da CDDC para a Ciência da Informação, bem como sobre a importância de se disponibilizar os dados produzidos em suas pesquisas. Acredita-se que tal discussão permitirá que a CI contribua, de forma definitiva, para melhorar as taxas de reuso de dados e condições de reprodução dos resultados de pesquisas científicas em todas as áreas de conhecimento, assim como de seus próprios pesquisadores.

## APÊNDICE 1 - INSTRUMENTO DE PESQUISA DA COLETA DE DADOS

### 1. CARACTERIZAÇÃO DO RESPONDENTE

Trabalha em qual Instituição:

Qual Formação acadêmica de graduação?

Titulação acadêmica

- Mestrado
- Doutorado

Idade?

- 21 a 30
- 31 a 40
- 41 a 50
- Mais de 50

Gênero?

- Feminino
- Masculino
- Não quero informar

Tempo de atuação como pesquisador?

- 2 a 5 anos
- 6 a 10 anos
- 11 a 20 anos
- Mais de 20 anos

Entidades onde realiza suas pesquisas? Selecione todas as aplicáveis.

- Universidade onde trabalho
- Universidade no Brasil
- Universidade internacional
- Instituto de pesquisa
- Organização não governamental
- Instituição de fomento à pesquisa nacional
- Instituição de fomento à pesquisa internacional
- Setor público
- Setor privado
- Outro:

## 2. INFORMAÇÕES SOBRE CURADORIA DIGITAL

A Curadoria Digital é um conceito abrangente que inclui o gerenciamento de dados científicos ao longo ciclo de vida das pesquisas. Inclui atividades de preservação de dados para acesso, compartilhamento, uso e reuso futuro. Este campo está se tornando uma oportunidade de carreira para o profissional da informação. Qual sua opinião mediante esta afirmativa?

- Concordo fortemente
- Concordo
- Sem opinião
- Discordo
- Discordo fortemente

Atualmente no Brasil, algumas instituições de pesquisa utilizam repositórios digitais de dados científicos de experimentos, medições ou entrevistas, para colaboração entre pesquisadores. Quais dos repositórios de dados científicos você já ouviu falar. Marque todas as opções aplicáveis.

- BDC/UFPR - Base de Dados Científicos da UFPR
- CIS – Consórcio de Informações Sociais
- GLOBE – Global Collaboration Engine
- IBGE – Instituto Brasileiro de Geografia e Estatística
- IBICT Dataverse Network
- Outro:

Como você faz o gerenciamento dos dados produzidos em suas pesquisas?

- Armazeno e mantenho no meu computador pessoal
- Armazeno e mantenho no Google Drive, Box ou outro similar
- Utilizo a plataforma fornecida da instituição em que trabalho
- Utilizo uma plataforma específica de armazenamento parceira da instituição em que trabalho
- Descarto quando termina a pesquisa
- Outro:

Em relação aos dados produzidos em suas pesquisas, marque as ações que você já realizou. Selecione todas as opções aplicáveis.

- Armazenou
- Acessou
- Compartilhou
- Disponibilizou
- Reutilizou
- Nenhuma das acima

Quais dessas plataformas tecnológicas você já utilizou para reutilizar dados de suas pesquisas? ( ) DSpace

- Dryad
- Zenodo
- Dataverse
- FigShare
- Tainacan
- CKAN
- EPrints
- EUDAT
- Outro:
- Nunca utilizei plataforma de gerenciamento de dados

O IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia) lançou em 2016 o Manifesto de Acesso Aberto a Dados de Pesquisa Brasileira para a Ciência Cidadã, para apoiar movimentos e iniciativas para a Ciência Aberta no Brasil, traduzidos pelo amplo e irrestrito acesso a fontes primárias de pesquisa possibilitando compartilhamento e reutilização de dados de pesquisa. Você tem conhecimento desse Manifesto?

- Sim, mas não o conheço bem
- Sim, e o conheço bem
- Não sei do que se trata

Não sei dizer

Já participou da elaboração de políticas e normas para o desenvolvimento das práticas de Curadoria Digital de dados científicos?

- Sim  
 Não  
 Não sei dizer

Já participou de algum evento que abordou como tema principal a Curadoria Digital de dados científicos?

- Não  
 Sim  
 Não sei dizer

Já participou de treinamento técnico (curso, capacitação) em práticas de Curadoria Digital de dados científicos?

- Sim  
 Não  
 Não sei dizer

### 3. PREFERÊNCIA, VISÃO E EXPERIÊNCIA

Identifique qual a importância das disciplinas/componentes curriculares que poderiam ser incluídos no currículo de graduação dos profissionais da informação (bibliotecários, arquivistas, cientistas da informação) para atuarem na área de Curadoria Digital? (1 – menos importante a 5 – mais importante)

	5 Extremamente importante	4 Muito importante	3 Moderadamente importante	2 Ligeiramente importante	1 Nem um pouco importante
Teoria de metadados					
Preservação digital					
e-Science					
Aprendizado de máquina					
Banco de dados					
Bibliotecas digitais					
Modelagem da Informação					
Análises estatísticas					
Big data					
Visualização de dados					
Linguagem de programação					
Ontologias					
Organização da informação					
Gerenciamento e análise de sistemas					
Recuperação da informação					

Quais habilidades você considera importante para um profissional da informação deve dominar em repositórios digitais de dados científicos?

- Processos ética da pesquisa científica  
 Métodos de pesquisa  
 Formas de Comunicação científica  
 Propriedade intelectual  
 Formas de acesso a dados digitais  
 Padrões de metadados  
 Marcos legais e regulatórios e Políticas de direito autoral  
 Todos as opções anteriores

- Nenhuma das opções anteriores
- Outros:

Das plataformas tecnológicas de armazenamento de dados abaixo, qual ou quais você já utilizou para compartilhar dados de suas pesquisas?

- DSpace
- Dryad
- Zenodo
- Dataverse
- FigShare
- Tainacan
- CKAN
- EPrints
- EUDAT
- Nunca utilizei plataforma de gerenciamento de dados
- Outro:

Qual padrão de metadados você utiliza para representar seus dados científicos?

- DarwinCore
- DataCite Metadata Schema
- DDI (Data Documentatio Initiative)
- Dublin Core
- ISO 19115
- Outro:
- Não sei dizer

Que modelo de ciclo de vida de dados você tem como referência para elaborar seu próprio Plano de Gerenciamento de Dados de Pesquisa?

- Digital Curation Centre (DCC)
- Modelo de Atividade de Pesquisa Idealizada I2S2
- DDI Combined Lifecycle Model
- UK Data Archive Data Lifecycle
- CVD-CI
- DataONE Data Lifecycle
- Research360
- OAIS
- Modelo padrão oferecido pela Instituição de Fomento à pesquisa
- Não sei dizer
- Outro:

Você já citou dados científicos produzidos por outros pesquisadores? (Cadernos de anotações de Laboratório ou campo, diários e transcrições textuais, dados de entrevistas e de questionários, Fluxos de trabalho e metodologias, registros de projetos de pesquisa, Fotografias, Arquivos de áudio, software, dentre outros).

- Não
- Sim
- Não sei dizer

Você concorda que estudos científicos que disponibilizam os dados produzidos podem ser mais citados do que estudos semelhantes que não o fazem.

- Concordo fortemente
- Concordo
- Sem opinião
- Discordo
- Discordo fortemente

Você já publicou dados científicos, por exemplo cadernos de anotações de Laboratório ou campo, diários e transcrições textuais, dados de entrevistas e de questionários, Fluxos de trabalho e metodologias, registros de projetos de pesquisa, Fotografias, Arquivos de áudio, software, dentre outros?

- Não
- Sim
- Não sei dizer

Em relação aos dados produzidos em suas pesquisas, normalmente o que você faz?

- Publica junto com os artigos
- Divulga em eventos científicos
- Armazena no computador pessoal
- Disponibiliza em repositórios de dados de pesquisa
- Descarta os dados
- Outro
- Não sei dizer

Qual(ais) necessidade(s) de formação técnica curricular que você precisa para praticar a Curadoria Digital de dados científicos?

#### 4. ACOMPANHAMENTO DE TENDÊNCIAS

Em quais atividades de Curadoria Digital um profissional da informação pode auxiliar um pesquisador? Selecione todas as opções aplicáveis.

- Armazenar e preservar dados
- Elaborar um plano de gerenciamento de dados
- Sobre questões legais relacionadas aos dados de minhas pesquisas
- Compartilhar dados
- Criação de metadados
- Encontrar dados e publicações para reuso
- Citações de dados
- Não sei dizer
- Concordo com todas as anteriores
- Outra:

Quais das atividades de Curadoria Digital abaixo são mais importantes para que um profissional da informação possa interagir junto aos pesquisadores no desenvolvimento de suas pesquisas científicas?

- Gerenciamento do ciclo de vida / continuum dos objetos digitais, talvez até mesmo a criação do sistema de manutenção de registros desses materiais
- Envolvimento ativo ao longo do tempo e a longo prazo dos criadores dos dados científicos
- Avaliação e seleção dos dados científicos para disponibilidade em repositórios digitais
- Desenvolvimento e provisão do acesso aos dados científicos
- Garantir a preservação (usabilidade e acessibilidade) dos dados científicos
- Outro:

Na sua opinião, qual das preocupações relacionadas abaixo é mais crítica para a pesquisa científica brasileira?

- A sustentabilidade e manutenção dos dados científicos em ambientes tecnológicos a longo prazo
- A criação ou avaliação, seleção, aquisição e ingestão e manutenção dos dados científicos para uso futuro
- Evolução, desenvolvimento e mudanças tecnológicas constantes
- Disponibilização dos dados científicos (publicação)
- Gerenciamento de acesso e outros controles sobre os dados científicos, primando pela base ética e legal desses controles
- A capacidade de usar, combinar, recombinar, interagir, processar, anotar, discutir e revisar os dados científicos ao longo do tempo (reuso para novas descobertas do conhecimento)
- Gerar ligação, contexto e metadados relacionados a dados científicos em softwares específicos
- Manter as informações de autenticidade, integridade, proveniência e contexto computacional relacionadas aos dados científicos
- Preservação digital dos dados científicos, incluindo preservar o acesso a estados passados dos dados científicos
- Descarte e destruição deliberada e / ou acidental dos dados científicos, em períodos de tempo potencialmente ampliados, embora os prazos possam ser comparativamente de curto ou médio prazo
- Reconhecer os impactos dos orçamentos financeiros necessários e potenciais mudanças futuras na política do país
- Prestar atenção à educação, treinamento e desenvolvimento de profissionais que possam dar suporte às

atividades necessárias de Curadoria Digital.

Na sua opinião, a Curadoria Digital de dados científicos, atualmente, no Brasil está ...



## REFERÊNCIAS

- Abagli, S. (2015). Ciência aberta em questão. In *Ciência aberta, questões abertas* (p. 09–25). Brasília: IBICT.
- Cox, A., Searle, S., Wolski, M., Simons, N., & Richardson, J. (2015). Librarians as partners in research data service development at griffith university. *Program: Electronic Library and Information Systems.*, 49(4), 440–460. doi: 10.1108/PROG-02-2015-0013.
- Gil, A. C. (2008). *Métodos e técnicas de pesquisa social*. (6a. ed.). São Paulo: Atlas.
- Gray, J., Szalay, A. S., Thakar, A. R., & Stoughton, C. (2002). Online scientific data curation, publication, and archiving. In *Virtual observatories* (v. 4846, p. 103–107). Redmond, WA: International Society for Optics and Photonics. Recuperado de <https://arxiv.org/pdf/cs/0208012.pdf>.
- Higgins, S. (2018). Digital curation: The development of a discipline within information science. *Journal of Documentation*, 74(6), 1318–1338. doi: 10.1108/JD-02-2018-0024.
- Kouper, I. (2016). Professional participation in digital curation. *Library and Information Science Research*, 38(3), 212–223. doi: 10.1016/j.lisr.2016.08.009.
- Lee, D. J., & Stvilia, B. (2017). Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PlosOne*, 12(3), e0173987. doi: 10.1371/journal.pone.0173987.
- Marconi, M. D. A., & Lakatos, E. M. (2003). *Fundamentos de metodologia científica*. (5a. ed.). São Paulo: Atlas.
- National Research Council. (2015). *Preparing the workforce for digital curation*. Washington, DC: National Academies Press. Recuperado de <https://www.ncbi.nlm.nih.gov/books/NBK293663/>.
- Nielsen, H. J., & Hjørland, B. (2014). Curating research data: the potential roles of libraries and information professionals. *Journal of Documentation*, 70(2), 221–240. doi: 10.1108/JD-03-2013-0034.
- Oliveira, & Grácio, M. C. C. (2005). Análise a respeito do tamanho de amostras aleatórias simples: uma aplicação na área de ciência da informação. *DataGramaZero: Revista de Ciência da Informação*, 6(3), 01–11.
- Oliveira, & Silva, E. M. d. (2016). Ciência aberta: dimensões para um novo fazer científico. *Informação & Informação*, 21(2), 05–39. doi: 10.5433/1981-8920.2016v21n2p5.
- Poole, A. H. (2016). The conceptual landscape of digital curation. *Journal of Documentation*, 72(5), 961–986. doi: 10.1108/JD-10-2015-0123.
- Sales, L. F., & Sayão, L. F. (2012). O impacto da curadoria digital dos dados de pesquisa na comunicação científica. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 7(2), 118–135. doi: 10.5007/1518-2924.2012v17nesp2p118.
- Sayão, L. F., & Sales, L. F. (2016). Curadoria digital e dados de pesquisa. *AtoZ: novas práticas em informação e conhecimento*, 5(2), 67–71. doi: 10.5380/atoz.v5i2.49708.
- Shintaku, M., Duque, C., & Suaiden, E. J. (2015). Análise da adesão às tendências da ciência pelos repositórios institucionais brasileiros. *InCID: Revista de Ciência da Informação e Documentação*, 6(2), 148–169. doi: 10.11606/issn.2178-2075.v6i2p148-169.
- Thompson, C. A., Senseney, M., Baker, K. S., Varvel, V. E., & Palmer, C. L. (2013). Specialization in data curation: Preliminary results from an alumni survey. *Proceedings of the American Society for Information Science and Technology*, 50(1), 01–04. doi: 10.1002/meet.14505001151.
- Tibbo, H. R., & Hank, C. (2015). Digital data curation essentials for data scientists and data curators and librarians. In *Proceedings of the 15th acm/ieee-cs joint conference on digital libraries* (p. 293–294). New York: Association for Computing Machinery. doi: 10.1145/2756406.2756928.
- Tripathi, M., Shukla, A., & Sonkar, S. K. (2017). Research data management practices in university libraries: a study. *DESIDOC Journal of Library and Information Technology*, 37(6), 417–424.
- Vanz, S. A. d. S., Passos, P. C. S. J., Caregnato, S. E., Pavão, C. G., Borges, N., Rocha, R. P. d., ... Azambuja, L. A. B. (2018). *Acesso aberto a dados de pesquisa no brasil: práticas e percepções dos pesquisadores: relatório 2018*. Recuperado de <http://hdl.handle.net/10183/185195>.
- Resende, L. C. de & Bax, M. P. (2020). A curadoria de dados científicos na Ciência da Informação: levantamento do cenário nacional. *AtoZ: novas práticas em informação e conhecimento*, 9(1), 94 – 103. Recuperado de: <http://dx.doi.org/10.5380/atoz.v9i1.69190>