

Percepções e práticas no âmbito do compartilhamento e reuso de dados: análise dos pesquisadores vinculados aos Programas de Pós-Graduação em Ciência da Informação do Brasil

Italo Teixeira Chaves¹

<https://orcid.org/0000-0001-7351-9565>

Marcus da Costa Fernandes¹

<https://orcid.org/0009-0005-0819-5310>

Nayra Lorhane Caldeira Bertoleza¹

<https://orcid.org/0000-0002-7540-4697>

Guilherme Ataíde Dias¹

<https://orcid.org/0000-0001-6576-0017>

Marckson Roberto Ferreira de Sousa¹

<https://orcid.org/0000-0003-2001-1631>

¹ Universidade Federal da Paraíba, João Pessoa, PB, Brasil

Resumo: Este estudo investiga as percepções e práticas de docentes doutores vinculados aos Programas de Pós-Graduação acadêmicos em Ciência da Informação no Brasil em relação ao compartilhamento e reuso de dados de pesquisa. Adota uma abordagem exploratória e aplicada, utilizando um questionário eletrônico para coleta de dados. Os resultados indicam que, embora os pesquisadores reconheçam a importância do compartilhamento de dados para a transparência e colaboração científica, a prática ainda enfrenta desafios, como falta de infraestrutura, limitação de tempo e escassez de recursos. A falta de conhecimento sobre gestão de dados e a ausência de políticas institucionais bem estabelecidas também se destacam como barreiras significativas. A pesquisa revela que a maioria dos participantes armazena seus dados em dispositivos pessoais e poucos utilizam repositórios institucionais. Do mesmo modo, a adesão a planos de gestão de dados ainda é limitada, o que pode comprometer a preservação e reutilização dos dados no longo prazo. Conclui-se que, para ampliar as práticas de compartilhamento e reuso de dados, é essencial o fortalecimento das infraestruturas institucionais, a promoção de capacitações e o desenvolvimento de diretrizes claras sobre gestão de dados de pesquisa. Os achados deste estudo contribuem para a compreensão dos desafios enfrentados pelos pesquisadores e podem subsidiar políticas e ações que incentivem a Ciência Aberta no Brasil.

Palavras-chave: gestão de dados de pesquisa; compartilhamento de dados; reuso de dados; ciência aberta; Ciência da Informação

1 Introdução

A gestão de dados de pesquisa (GDP) é um tema pertinente na área de Ciência da Informação (CI), notadamente em um cenário onde a produção, o compartilhamento e o reuso desses dados são essenciais para o avanço do conhecimento científico. A digitalização e a revolução tecnológica provocaram uma nova dinâmica na geração, armazenamento e disseminação de dados, criando tanto oportunidades quanto desafios para pesquisadores e instituições. Sayão, Sales e Carvalho Segundo (2023) enfatizam a importância de uma gestão de dados ampla, envolvendo aspectos como metodologias, software, códigos, algoritmos, *workflow*, entre outros, que tornem a gestão de dados mais transparente.

No Brasil, há um aumento das discussões sobre Ciência Aberta (*Open Science*), constituindo-se de um modelo que promove maior transparência, participação social e o compartilhamento aberto de insumos como dados e algoritmos (Albagli, 2019). Nesse contexto, percebe-se que, todavia, ainda são poucas as pesquisas empíricas que investigam as práticas de GDP e as percepções dos profissionais sobre esse tema. Na Índia, Bunkar e Bhatt (2020) perceberam que os docentes têm dificuldades com a GDP, além disso, ressaltaram o papel potencial que a biblioteca – e por consequência, os bibliotecários - tem em auxiliar os docentes nesse processo.

Considerando estes apontamentos, esta pesquisa investiga as práticas e percepções dos docentes doutores na Ciência da Informação, em cursos de pós-graduação de modalidade acadêmica, com foco no uso e reuso de dados de pesquisa. Considerando o papel central desses profissionais na produção, orientação e disseminação do conhecimento, torna-se fundamental entender como eles gerenciam seus dados de pesquisa e quais são as principais dificuldades que enfrentam.

A **pergunta de partida** que norteia esta investigação é: como caracterizar os aspectos relativos ao compartilhamento e reuso de dados na Gestão de Dados

de Pesquisa no contexto dos pesquisadores em Ciência da Informação no Brasil? A partir dessa questão, busca-se explorar os desafios relacionados à coleta, armazenamento, preservação e compartilhamento de dados, bem como as práticas e ferramentas utilizadas pelos docentes para garantir a segurança e a integridade dos dados.

O estudo adota uma abordagem aplicada e exploratória, dada a necessidade de resolver problemas contextuais com fins práticos. A pesquisa se destina a preencher algumas lacunas existentes sobre o compartilhamento e reuso de dados de pesquisa no Brasil, podendo fornecer informações e conclusões valiosas que possam orientar futuras políticas e práticas no ambiente acadêmico. Ademais, ao focar nas percepções e práticas dos docentes doutores, espera-se contribuir para uma maior compreensão da importância da GDP e dos benefícios do compartilhamento para o avanço científico e a colaboração interdisciplinar.

Assim sendo, define-se como **objetivo geral** analisar as práticas e percepções dos docentes doutores de Ciência da Informação relacionados a gestão de dados científicos, com foco nos aspectos relacionados ao uso e reuso dos dados científicos. Deste modo, os **objetivos específicos** são: (a) verificar as percepções dos docentes acerca da GDP; (b) identificar as práticas dos docentes relativos à gestão de dados científicos; (c) analisar a caracterização da GDP na Ciência da Informação.

Esta investigação tem sua justificativa baseada na importância crescente do tema GDP para todas as áreas do conhecimento, e em especial para a Ciência da Informação, visto que este movimento tornou-se objeto de pesquisa na CI, como destacam Silveira *et al.* (2021). Com a evolução da tecnologia e a crescente digitalização de dados de pesquisa, o volume de dados gerados, armazenados e compartilhados vem aumentando significativamente, criando tanto benefícios quanto barreiras para a comunidade científica.

No Brasil, a discussão sobre Ciência Aberta tem demonstrado cada vez mais força, através da promoção da importância da transparência e do compartilhamento de dados de pesquisa. No entanto, poucas são as pesquisas que têm investigado como os cientistas, em especial os docentes integrantes de

programas de pós-graduação, lidam com a gestão de seus dados de pesquisa. Ressalta-se que estes pesquisadores desempenham um papel decisivo na produção e disseminação de novos conhecimentos, o que faz com que seja essencial a sua compreensão e a plena adoção das práticas relacionadas à GDP.

O mapeamento das práticas e a identificação dos desafios encontrados, fornecerão evidências importantes que permitirão dar subsídios para a orientação de futuras políticas nas instituições de ensino superior, fortalecendo a GDP na Ciência da Informação no Brasil.

2 Gestão de dados de pesquisa

A GDP se apresenta com relevância na ciência contemporânea, sendo um elemento significativo para o avanço do conhecimento, haja vista envolver a compreensão de como os dados primários de pesquisas são geridos, recuperados e preservados, por meio de práticas e políticas que abrangem o ciclo de vida dos dados, desde a sua coleta até a preservação e reutilização. Essas ações visam garantir que os dados sejam organizados, acessíveis e reutilizados, promovendo a transparência e a confiabilidade das pesquisas científicas, uma vez que os dados de pesquisa são informações registradas ou produzidas através de qualquer forma, ou meio durante o decurso de uma pesquisa e todas as evidências que um investigador necessita para validar suas conclusões após uma pesquisa (Silva, 2019). Nesse sentido, a gestão de dados envolve técnica e estratégia, com impactos positivos na eficiência das investigações e na sustentabilidade da produção científica.

Nesse contexto, os dados podem ser classificados de diversas formas. Quanto à origem, incluem primários (brutos e originais, como experimentos e enquetes), secundários (organizados e derivados, como censos) e terciários (estatísticas ou recontagens que garantem confidencialidade). Pelo formato, dividem-se em estruturados (ex.: tabelas), semiestruturados (como XML) e não estruturados (textos ou vídeos). Os dados abertos são acessíveis a todos para uso e compartilhamento. Licenças *Creative Commons* definem as condições de uso,

enquanto formatos como CSV ou PDF facilitam o armazenamento e a reutilização eficiente (Silva, 2019).

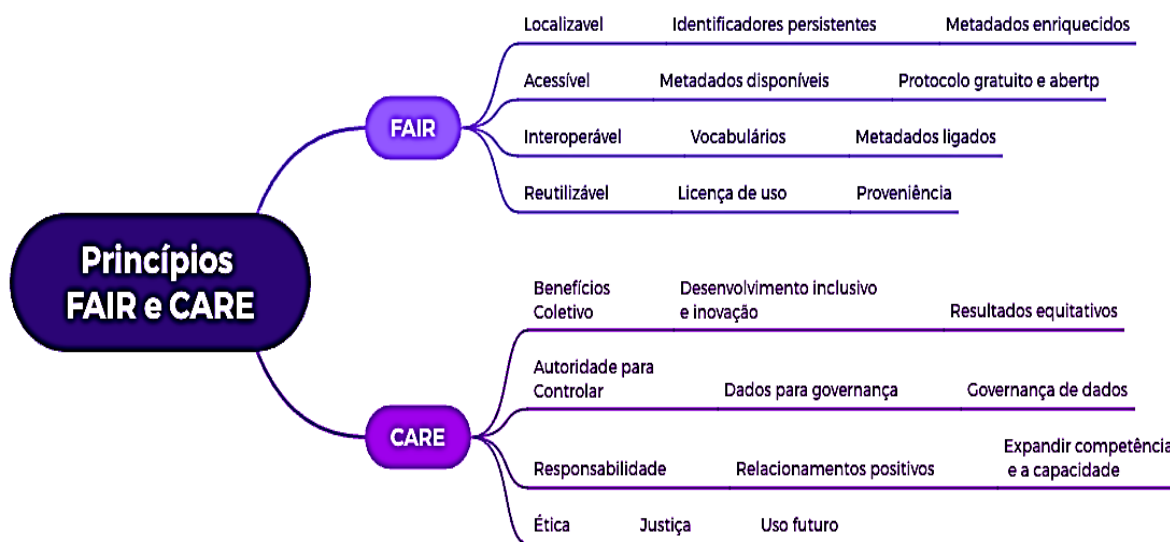
A esse propósito, cabe destacar que a implementação de Planos de Gestão de Dados (PGD) tem se consolidado como uma prática indispensável para pesquisas financiadas por agências de fomento e conduzidas em instituições acadêmicas. Estes planos definem, entre outros aspectos, quais dados serão gerados, como serão armazenados e preservados, além de considerar aspectos éticos, legais e de confidencialidade (Veiga *et al.*, 2022). Silva (2019) observa que o PGD se consubstancia em um documento elaborado para fins de descrever sobre um ciclo de vida dos dados, registrar decisões tomadas em relação aos padrões de metadados, formatos, bases de dados, métodos, segurança e períodos de armazenamento, assim como os custos com a gestão dos dados.

Nesse sentido, os Princípios FAIR (*Findable/Localizável, Accessible/Acessível, Interoperable/Interoperável e Reusable/Reutilizável*), propostos no ano de 2016, tornaram-se um marco fundamental, definindo diretrizes claras para o compartilhamento e a reutilização de dados. De acordo com Wilkinson *et al.* (2016), esses princípios permitem que os dados sejam localizáveis por meio de identificadores persistentes, acessíveis mediante protocolos padrão, interoperáveis por meio de metadados consistentes e reutilizáveis com licenças claras e descrições detalhadas, insertas na compreensão de que os dados passam a ser importantes para o processo científico e para toda a sociedade.

Complementando essa abordagem técnica, os Princípios CARE (*Collective Benefit/Benefício Coletivo, Authority to Control/Autoridade para Controlar, Responsibility e Ethics/Responsabilidade, Ética*) destacam as dimensões éticas e sociais da gestão de dados. Esses princípios são particularmente relevantes em pesquisas que envolvem comunidades específicas ou dados sensíveis, reforçando a importância da soberania, da justiça e do impacto coletivo dos dados (Vidotti; Torino; Coneglian, 2021) e a sua integração com os Princípios FAIR pode contribuir para um maior equilíbrio entre eficiência técnica e responsabilidade social, com a consequente ampliação do alcance e a aceitação

das práticas de gestão de dados. A figura 1 congrega as principais características dos princípios FAIR e CARE.

Figura 1 - Características dos princípios FAIR e CARE



Fonte: Adaptado de Torino, Monteiro e Vidotti (2023).

Com efeito, o ciclo de vida dos dados apresenta-se como uma estrutura importante para a gestão eficiente de dados de pesquisa, abrangendo etapas como planejamento, coleta, processamento, preservação e compartilhamento. Conforme descrito no modelo referencial proposto por Rodrigues *et al.* (2023), resultante da interseção do *Digital Curation Centre*, do *Data Observation Network for Earth* e da *Data Documentation Initiative*, para descrição de dados científicos na Ciência da Informação, com aplicabilidade abrangente nas Ciências Sociais, ofereceu-se uma estrutura detalhada para gerenciar dados científicos e aprimorar a metodologia de pesquisa.

Porquanto, Santos e Sant’Ana (2019) destacam, ao tratar sobre a construção de pontes entre usuários e bases de dados, que no ciclo de vida de dados indispensável se faz a realização de um processo de planejamento e de execução de ações a fim de que os dados estejam disponíveis em um suporte para acesso futuro.

Nessa perspectiva, a integração de fases como curadoria e descrição detalhada dos metadados possibilita assegurar a qualidade e a reusabilidade dos dados ao longo do tempo. O uso de modelos de ciclo de vida, como os propostos pelo *Digital Curation Centre* (DCC) e pelo *Data Observation Network for Earth* (DataONE), tem se mostrado eficaz para a organização e preservação dos dados, otimizando seu impacto na pesquisa (Silva Segundo; Araújo, 2019). Assim, essas práticas contribuem para o avanço científico e promovem abordagens metodológicas alinhadas à Ciência Aberta e ao compartilhamento responsável dos resultados.

No contexto da Ciência da Informação, Sant'Anna (2016) sugere um ciclo de vida dos dados com quatro etapas principais: coleta, armazenamento, recuperação e descarte. Além disso, explica que todas as etapas se fundamentam em aspectos ligados à privacidade, integração, qualidade, direitos autorais, disseminação e preservação. O autor ainda destaca a importância da CI, em conjunto com outras áreas, como Matemática e Ciência da Computação, em traçar soluções estratégicas para lidar com a abundância de dados existentes.

Portanto, a curadoria de dados possui sua relevância na GDP, uma vez que assegura a manutenção dos dados a curto prazo e sua preservação a longo prazo (Dias *et al.*, 2023). Ressalta-se que a curadoria, aliada a práticas de limpeza de dados, nos termos do conceito de *Garbage In, Garbage Out* (GIGO), a fim de impulsionar o processo de tratamento de dados, reforçam a questão da qualidade dos dados em processos computacionais e de informação, no qual dados de baixa qualidade comprometem a validade das análises.

No Brasil, a GDP enfrenta desafios que vão desde a falta de infraestrutura tecnológica até a ausência de políticas nacionais robustas e amplamente implementadas (Costa; Cunha, 2019). Apesar disso, avanços como a disseminação de repositórios institucionais e o fortalecimento de iniciativas de Ciência Aberta oferecem perspectivas promissoras.

Segundo Veiga *et al.* (2022), a capacitação de pesquisadores e o alinhamento com padrões internacionais são passos elementares para superar

essas barreiras e consolidar a gestão de dados como um pilar estratégico da produção científica.

Posto isso, percebe-se que uma GDP, utilizando de tecnologias emergentes e princípios relacionados, fomenta a existência de uma rede de colaboração global de pesquisadores, com avanço científico, resolução de problemas complexos, por meio de uma pesquisa colaborativa, inclusive de maneira multidisciplinar, além de mais transparente e responsável. Dessa forma, a gestão de dados se consolida como uma ferramenta indispensável para a construção de um conhecimento mais acessível e sustentável. Logo, uma das temáticas que emergem sobre a GDP é o compartilhamento e reuso de dados.

Quando dados de pesquisa são disponibilizados no formato aberto (*open data*), tem-se a possibilidade de aproveitamento para aceleração de resultados, novas interpretações e impulsionamento da ciência. O compartilhamento de dados pode ocorrer de diferentes formas, seja entre pesquisadores, depósito em repositórios, publicação de artigos de dados, entre outros. Nesse sentido, dados de pesquisa possuem elementos para mudanças na cultura tradicional da comunicação científica, notadamente diante do movimento da Ciência Aberta, constituindo-se de um movimento global que promove o acesso livre e colaborativo ao conhecimento, visando maior transparência, cooperação e impacto social. (Silva; Prado; Araújo, 2024).

Apesar do compartilhamento de dados oferecer diversos benefícios à sociedade científica, Monteiro e Sant'ana (2023) apontam que preocupações relacionadas ao reconhecimento da autoria, bem como questões pessoais, administrativas ou de recursos, preocupações relacionadas ao tempo despendido e aos custos na geração ou coleta, no tratamento, na análise e na gestão de dados. Além disso, a falta de clareza quanto às normativas, podem inibir o compartilhamento dos dados por parte dos pesquisadores. De toda forma, é válido ressaltar os benefícios existentes no compartilhamento, como destacam Silva, Prado e Araújo (2024, p. 3):

A prática de compartilhamento apresenta vantagens e benefícios para outros pesquisadores que farão o reuso dos dados de pesquisa, bem como para instituições de pesquisa e agências de fomento. Ao compartilhar os dados obtidos em pesquisas científicas, o retorno para quem pesquisa torna-se, de certa forma, mais ágil, e evitam-se gastos desnecessários.

Araújo, Dias e Autran (2021) destacam que o compartilhamento de dados é uma temática importante para todas as áreas do conhecimento. Nesse âmbito, destacam a importância de conscientizar pesquisadores sobre práticas de compartilhamento e disponibilização dos dados, bem como na criação de políticas sobre essas práticas, uma vez que “[...] essas medidas podem colaborar com a credibilidade das investigações científicas, a economia dos recursos públicos e a eliminação do retrabalho na coleta de dados” (Araújo; Dias; Autran, 2021, p. 102).

Compartilhar os dados é, sem dúvidas, uma prática importante, sobretudo quando se pensa no seu uso e reuso. Para tanto, é preciso que os pesquisadores tratem os seus dados. Silva, Prado e Araújo (2024) explicam que os dados devem passar por etapas de catalogação e documentação, envolvendo aspectos ligados diretamente com a representação descritiva e o uso de metadados, além da utilização de identificadores persistentes como o *Digital Object Identifier* (DOI).

Para Pasquetto, Randles e Borgman (2017) é fundamental que se faça a distinção entre uso e reuso de dados. Dados são coletados por um pesquisador ou grupo de pesquisa para um determinado projeto, visando a responder um problema de pesquisa específico. Se o mesmo pesquisador ou grupo de pesquisa utilizar este conjunto de dados novamente, seja para o mesmo projeto, ou para um projeto distinto, considera-se como uso. Quando o conjunto de dados é depositado em um repositório e utilizado por outro pesquisador ou grupo de pesquisa, é considerado reuso. Posto de forma simples, reuso implica no uso de um conjunto de dados por alguém que não o originou.

De acordo com Curty (2016), a sustentabilidade do ciclo de vida da Ciência Aberta depende de condições propícias para a maximização do reuso de dados científicos. Nesse sentido, o reuso de dados científicos é um processo complexo, exige que cientistas consigam traduzir e re-contextualizar os dados primários obtidos por outros pesquisadores, a fim de aplicar para os seus próprios

fins, sem que haja má-interpretação ou má-utilização destes. Reutilizar dados de pesquisa disponíveis de forma pública exige o esforço de lidar com dados produzidos por outra pessoa, o que pode afetar o reuso efetivo desses dados.

É preciso ter a compreensão clara que:

O fato dos dados estarem disponíveis não garante, por si só, sua reutilização, exigem-se competências neste processo. Sem o letramento para o reuso, pesquisadores não têm acesso, não obtêm e não reusam dados, limitando suas fronteiras. Não usufruem dos valores inerentes à Ciência Aberta, dos seus recursos científicos que lhe conferem caráter mais democrático, com disponibilização acessível de dados de alto valor (Estevão; Strauhs, 2020, p. 19).

Dessa forma, é possível inferir que o compartilhamento possibilita o reuso, mas não o garante. É um dos fatores que pode estar garantindo essa reutilização de dados: passa a falta de conhecimento, ou, como os autores mencionam, letramento para o uso de dados. Para minimizar esses desafios é preciso compreender que tipos de dados estão sendo trabalhados, os anseios, as dificuldades dos pesquisadores em compartilhar, entre outros. Tendo por base esse conhecimento é possível desenvolver um plano de ação com vistas a maximizar os efeitos positivos do compartilhamento e reuso de dados de pesquisa.

3 Percurso metodológico

Quanto à natureza, este estudo se configura como uma pesquisa aplicada, a medida em que visa resolver um problema contextual com fins práticos (Gerhardt; Souza, 2009) envolvendo o compartilhamento e reuso de dados de pesquisa. Deste modo, a pesquisa direciona o seu foco para as percepções e práticas desenvolvidas pelos docentes doutores nos Programas de Pós-Graduação acadêmicos na área de Ciência da Informação do Brasil.

Em relação aos procedimentos técnicos, essa pesquisa se classifica como um estudo de caso. O estudo de caso consiste no estudo profundo e exaustivo de um objeto de maneira que permita seu amplo e detalhado conhecimento. É entendido como o delineamento mais adequado para a investigação de um fenômeno contemporâneo dentro de seu contexto real (Gil, 2002).

Nesse contexto, a presente pesquisa se configura como exploratória, visto que os pesquisadores buscam dar maior notoriedade para um tema pouco abordado no cenário brasileiro (Sampieri; Collado; Lúcio, 2013). Embora as discussões no tocante à Ciência Aberta tenham ganhado notoriedade, percebe-se que ainda existem poucos estudos no contexto dos dados abertos, sobretudo no que diz respeito ao seu compartilhamento e reuso.

Para operacionalizar esta pesquisa e delimitar o público estudado, foi utilizado a plataforma Sucupira. A plataforma Sucupira, serviu para selecionar os cursos de pós-graduação em Ciência da Informação do Brasil, suas páginas institucionais e, a partir destas, coletar os dados de contato dos docentes. Os critérios de seleção adotados se baseiam em: programas de Pós-Graduação em Ciência da Informação, modalidade acadêmica e que esteja em funcionamento. A Tabela 1 apresenta os programas selecionados a participarem da pesquisa, conforme os critérios de inclusão estabelecidos.

Tabela 1 - Programa selecionados para pesquisa

Região	Programa De Pós-Graduação	Instituição	Nº de participantes convidados
Norte	Ciência da Informação	Universidade Federal do Pará	22
Nordeste	Ciência da Informação	Universidade Federal da Paraíba	22
	Ciência da Informação	Universidade Federal de Pernambuco	18
	Ciência da Informação	Universidade Federal do Ceará	15
	Ciência da Informação	Universidade Federal de Alagoas	16
	Ciência da Informação	Universidade Federal da Bahia	21
Centro-Oeste	Ciência da Informação	Universidade de Brasília	16
Sudeste	Ciência da Informação	Universidade Federal Fluminense	19
	Ciência da Informação	Universidade Estadual Paulista "Júlio de Mesquita Filho"	20
	Ciência da Informação	Universidade de São Paulo	19
	Ciência da Informação	Universidade de São Carlos	18

Região	Programa De Pós-Graduação	Instituição	Nº de participantes convidados
	Ciência da Informação	Universidade Federal do Rio de Janeiro	31
	Ciência da Informação	Universidade Federal do Espírito Santo	12
	Ciência da Informação	Universidade Federal de Minas Gerais	20
	Organização do Conhecimento		17
Sul	Ciência da Informação	Universidade Federal do Rio Grande do Sul	14
	Ciência da Informação	Universidade Federal de Santa Catarina	22
	Ciência da Informação	Universidade de Londrina	18
Total			340

Fonte: Bertoleza *et al.* (2024).

A pesquisa foi enviada para um total de 340 docentes vinculados a programas de pós-graduação em Ciência da Informação, englobando as cinco regiões do país, dos quais 92 responderam. O questionário foi enviado para o e-mail disponibilizado nas páginas institucionais dos referidos programas. Nos casos em que não havia tal informação disponível, os investigadores realizaram uma pesquisa em publicações desses autores, com vistas a encontrar e-mails utilizados, e, dessa maneira, encaminhar o questionário para resposta. Salienta-se que todos os docentes vinculados aos programas foram convidados a participarem da pesquisa, e que o instrumento de coleta foi estruturado com questões obrigatórias, não havendo a possibilidade de não resposta.

A partir dos dados coletados foi enviado o link do formulário eletrônico Google Forms por meio do e-mail com perguntas em duas categorias, a primeira relativa à percepção dos docentes sobre compartilhamento e reuso de dados de pesquisa e a segunda relativa às práticas que esses mesmos docentes adotam no seu cotidiano enquanto pesquisadores. Salienta-se que o período de coleta dos dados ocorreu entre os meses de setembro e outubro de 2024. Silveira e Cordova (2009) explicam que esse tipo de pesquisa busca informação em um grupo

específico do seu interesse. Salienta-se ainda que “[...] nesse tipo de pesquisa, o respondente não é identificável, portanto, o sigilo é garantido” (Silveira; Cordova, 2009, p. 39).

Para participar desta investigação os pesquisadores tiveram que dar o seu consentimento a partir das disposições presentes no termo de consentimento livre e esclarecido, desenvolvido considerando as resoluções e princípios éticos que regem pesquisas com seres humano¹. Deste modo, compõe a amostra final do estudo aqueles que concordaram em participar e terem seus dados analisados. Os pesquisadores que optarem por não participar da pesquisa não tiveram seus dados coletados. Salienta-se que, mesmo concordando em participar do estudo, os participantes puderam desistir a qualquer momento, sem qualquer restrição.

Considerando isso, a pesquisa teve riscos mínimos, limitados à possibilidade de eventual desconforto ou cansaço ao responder o questionário apresentado. Para evitar tal situação, o formulário foi habilitado para salvar as respostas, possibilitando que o participante as retomasse em momento posterior.

Salienta-se ainda que não houve benefício direto aos participantes desta investigação, sendo o desenvolvimento científico o principal resultado desta proposta investigativa. Assim sendo, os resultados apresentados no tópico seguinte buscam fomentar contribuições científicas no âmbito da Ciência da Informação no tocante ao compartilhamento e reuso dos dados de pesquisa.

O Quadro 1, apresentado a seguir, contém as categorias de análise e as respectivas perguntas constantes no formulário eletrônico enviado aos pesquisadores, desenvolvido por meio do Google Forms. Ademais, destaca-se que as perguntas foram desenvolvidas à luz da literatura científica apresentada anteriormente, bem como com foco no alcance dos objetivos propostos.

Quadro 1 - Categorias de análise e perguntas

Objetivo Específico	Pergunta
Verificar as percepções dos docentes acerca da gestão de dados de pesquisa	Qual o nível de importância do compartilhamento e reuso de dados de pesquisa para o avanço do conhecimento científico?
	Qual sua motivação para compartilhar seus dados de pesquisa?
	Qual seu nível de disposição para reutilizar dados de pesquisas disponibilizados por outros pesquisadores?
	Quais são os principais benefícios que você percebe no compartilhamento dos seus dados de pesquisa com outros pesquisadores?
	Quais são os principais desafios ou barreiras que você enfrenta para disponibilizar seus dados de pesquisa?
	Qual é a sua principal preocupação ao compartilhar seus dados de pesquisa?
Identificar as práticas dos docentes relativos à gestão de dados científicos	Com que frequência você compartilha seus dados de pesquisa?
	Quais recursos de pesquisa você gera e/ou utiliza em seus projetos?
	Qual o seu nível de conhecimento sobre um Plano de Gestão de Dados (PGD)?
	De que maneira você armazena os dados gerados em suas pesquisas?
	Você utiliza algum repositório específico para armazenar e compartilhar seus dados de pesquisa?
	Você já participou de capacitações relacionadas à gestão de dados de pesquisa em suas atividades de ensino e pesquisa?

Fonte: Elaborado pelos autores, com base em Bunkar e Bhatt (2020); Cheung et al. (2022); Lima (2020).

As perguntas presentes no questionário desta investigação foram adaptadas de pesquisas nacionais e internacionais, sobretudo os estudos de Bunkar e Bhatt (2020), Cheung *et al.* (2022) e Lima (2020). Considerando que essas perguntas foram aplicadas e validadas pelas referidas pesquisas, considera-se que o instrumento desenvolvido contém validade científica e está adequado para o estudo no âmbito da Ciência da Informação. A partir da pesquisa bibliográfica realizada, notou-se a existência de lacunas no âmbito do compartilhamento e reuso de dados, dessa maneira, com foco nos estudos empíricos e no aprofundamento da literatura, elegeram-se as perguntas no quadro anterior como pertinentes a esta pesquisa e à temática estudada.

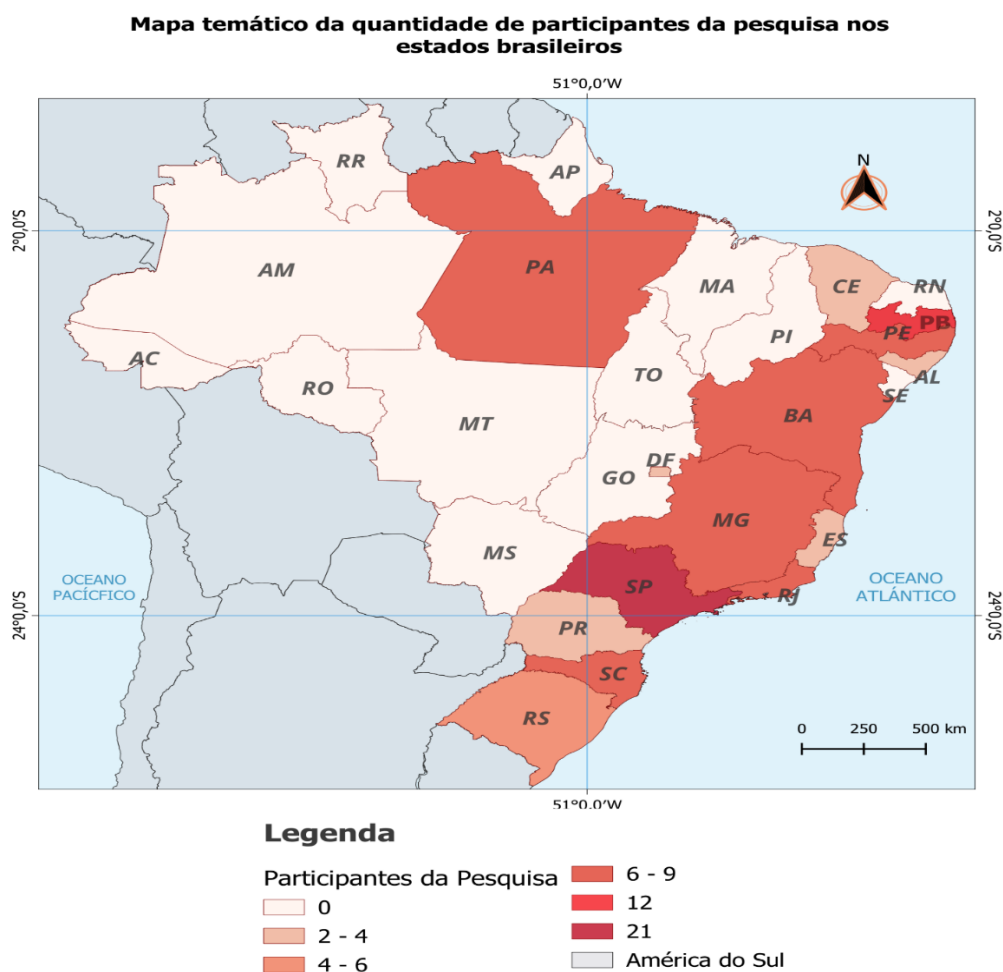
Além das perguntas dispostas no Quadro 1 foram desenvolvidas três perguntas para traçar minimamente o perfil dos participantes: unidade federativa da sua principal instituição de pesquisa, anos de atuação na pós-graduação e faixa

etária do pesquisador. A partir das respostas coletadas, os dados foram tabulados e representados por meio de quadros e gráficos, e posteriormente analisados em uma abordagem mista, a partir do cruzamento das respostas dos participantes com a literatura, utilizando para isso estatística descritiva básica, oportunizando uma análise capaz de consolidar o objetivo geral deste estudo.

4 Resultados e discussões

A pesquisa contou com a participação de 92 pesquisadores, obtendo respostas de todas as regiões do Brasil. A maior concentração de respondentes foi, respectivamente, dos estados de São Paulo (n=21), Paraíba (n=12) e Santa Catarina (n=8). A Figura 2 apresenta um mapa com a distribuição dos participantes da pesquisa.

Figura 2 - Distribuição de participação da pesquisa

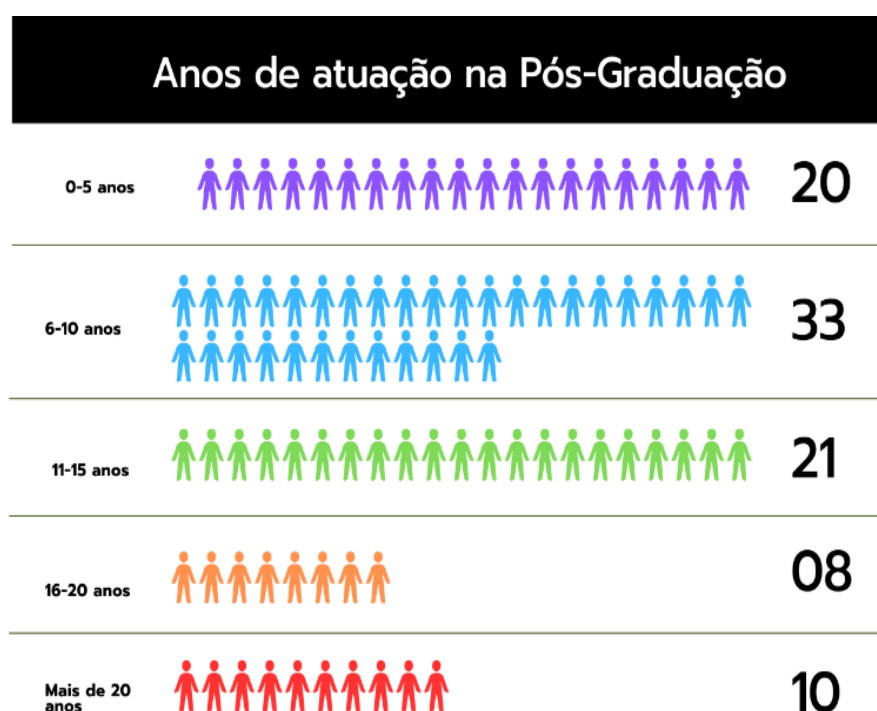


Fonte: Bertoleza *et al.* (2024).

Salienta-se que alguns estados não dispõem de Programas acadêmicos de Pós-Graduação em Ciência da Informação, como Acre, Amazonas, Mato Grosso, entre outros. E, por esta razão, não estão presentes no escopo desta pesquisa.

Para traçar melhor o perfil dos participantes da pesquisa, optou-se por questionar qual o período de atuação dos pesquisadores na pós-graduação. Os resultados dessa pergunta foram representados na Figura 3.

Figura 3 - Anos de atuação na pós-graduação

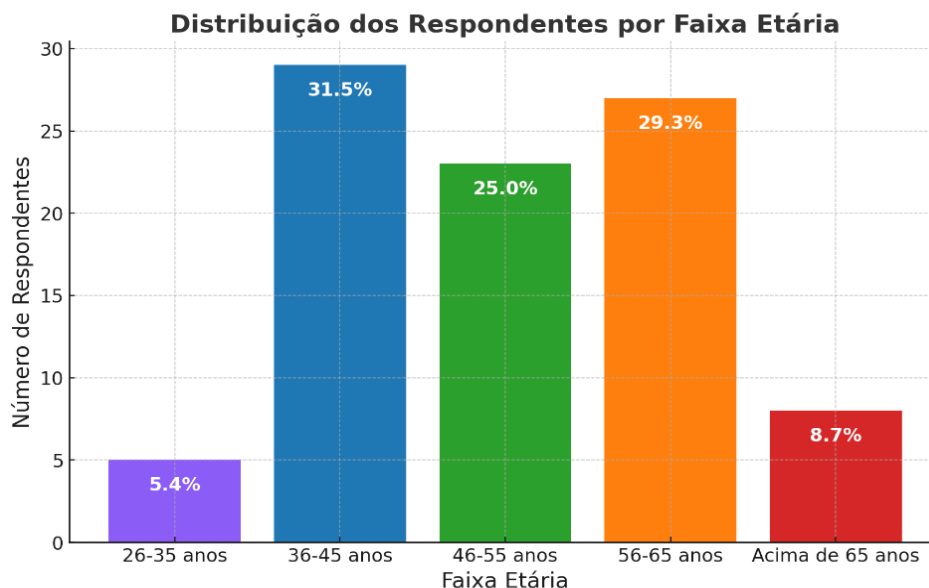


Fonte: Bertoleza *et al.* (2024).

Os participantes apresentam uma diversidade de tempos de atuação na pós-graduação, com predominância na faixa de seis a dez anos (35,9%), indicando que muitos ainda estão em fases intermediárias de suas carreiras acadêmicas. Isso pode significar que alguns pesquisadores podem ter tido poucos contatos com práticas relativas ao compartilhamento de dados, sobretudo quando se considera que essa temática está sendo discutida e implementada nas instituições e agências de fomento no Brasil.

Por fim, para finalizar o perfil dos pesquisadores participantes, questionou-se a faixa etária destes. Os resultados dessa questão estão dispostos no Gráfico 1.

Gráfico 1 - Faixa etária do pesquisador



Fonte: Bertoleza *et al.* (2024).

Em relação à idade, a maioria dos pesquisadores está entre 36 e 45 anos, o que sugere um grupo ativo e em consolidação profissional na pesquisa científica. Ademais, nota-se que a faixa etária 26 e 35 anos representa o menor percentual ($n=5$), contudo, parte destes participantes podem se configurar como nativos digitais e terem uma maior familiaridade com as tecnologias.

Como pode ser observado no perfil brevemente levantado, boa parte dos programas de pós-graduação em Ciência da Informação participaram da pesquisa. Assim sendo, com o intuito de perceber as percepções dos pesquisadores, questionou-se, logo de início, sobre a frequência, importância e disposição dos pesquisadores em compartilhar e reutilizar dados. Os dados foram sintetizados na Tabela 2, com a quantidade de participantes e sua respectiva porcentagem. Com vistas a identificar os maiores e menores índices de respostas, optou-se por realizar um destaque em negrito.

Tabela 2 - Frequência, importância e disposição para compartilhamento e reuso de dados de pesquisa

FREQUÊNCIA				
Com que frequência você compartilha seus dados de pesquisa?				
Sempre	Frequentemente	Ocasionalmente	Raramente	Nunca
11 (12%)	16 (17,4%)	38 (41,3%)	18 (19,6%)	9 (9,8%)
IMPORTÂNCIA				
Qual o nível de importância do compartilhamento e reuso de dados de pesquisa para o avanço do conhecimento científico?				
Muito Importante	Importante	Moderadamente Importante	Pouco Importante	Nada Importante
60 (65,2%)	30 (32,6%)	1 (1,1%)	0 (0,0%)	0 (0,0%)
DISPOSIÇÃO				
Qual seu nível de disposição para reutilizar dados de pesquisas disponibilizados por outros pesquisadores?				
Muito disposto	Disposto	Neutro	Pouco disposto	Nada disposto
29 (31,5%)	49 (53,3%)	13 (14,1%)	1 (1,1%)	0 (0,0%)

Fonte: Bertoleza *et al.* (2024).

O primeiro questionamento, que buscou compreender a frequência sobre compartilhamento de dados, o que acabou evidenciando que, a maioria dos respondentes compartilha ocasional ou raramente os dados de pesquisa. Diante disso, é possível notar que existem outros cenários semelhantes ao investigado, como evidência Bunkar e Bhatt (2020) ao descobrirem que aproximadamente 60% dos pesquisadores da Universidade de Parul, na Índia, compartilham dados de pesquisa. Percebeu-se que pesquisadores com mais de 20 anos de carreira, ocasional ou raramente compartilham seus dados de pesquisa, em contraponto, estes se mostraram dispostos ou muito dispostos a compartilharem seus dados. Isso leva a necessidade de refletir sobre ações de formação para possibilitar que haja entendimento das possibilidades envolvendo o compartilhamento. Em conexão com essa ideia, Estevão e Strauhs (2020) chamam atenção para o

letramento envolvendo o reuso de dados, aspectos que precisa ser fortalecido na CI.

Embora a prática de compartilhamento e reuso de dados ainda esteja em período de desenvolvimento e consolidação na Ciência da Informação brasileira, fica notório que é majoritária a percepção dos pesquisadores que esta prática é importante ou muito importante, conforme pontuam mais de 90% dos participantes. A respeito disso, existem barreiras, como falta de infraestrutura tecnológica e limitação de tempo, que, mesmo com os participantes considerando a importância do compartilhamento de dados, ainda existe dificuldade em consolidar esta prática. Lima (2020) corrobora esse ponto em sua pesquisa ao notar que pesquisadores não consideram problema o compartilhamento de dados, embora pontuem a importância de saber como fazer isso e destacarem uma atenção especial para as questões éticas que devem nortear o referido compartilhamento.

Por fim, obteve-se que mais de 80% dos participantes estão dispostos a reutilizarem dados de pesquisa disponibilizados por terceiros. Mesmo que haja disposição em reutilizar dados de pesquisa, torna-se um desafio à medida que os pesquisadores não têm o hábito de compartilhá-los. Sobre essa característica, Lima (2020) notou a existência de uma barreira visto que alguns pesquisadores compartilham dados somente com grupos parceiros, alunos ou colegas. Compreendendo a existência dessas problemáticas, Monteiro e Sant'Ana (2023) enfatizam a urgência de existirem diretrizes e recomendações de boas práticas para o compartilhamento e reuso de dados, para evitar um estado de anomia quanto a essas práticas.

Buscou-se compreender quais as motivações envolvidas no compartilhamento e reuso de dados por parte dos entrevistados, bem como os benefícios percebidos por eles. Além disso, foi questionado quais os principais desafios ou barreiras percebidas pelos respondentes para os mesmos disponibilizarem seus dados de pesquisa. As respostas, com possibilidade de múltipla escolha, estão sintetizadas na Tabela 3:

Tabela 3 - Motivação, benefícios e desafios ou barreiras para compartilhamento e reuso de dados de pesquisa

MOTIVAÇÃO		BENEFÍCIOS		DESAFIOS OU BARREIRAS	
Qual a sua motivação para compartilhar seus dados de pesquisa?	n (%)	Quais são os principais benefícios percebidos no compartilhamento dos seus dados de pesquisa com outros pesquisadores	n (%)	Quais são os principais desafios ou barreiras que você enfrenta para disponibilizar seus dados de pesquisa?	n (%)
Promover a transparência na pesquisa científica	56 (60,9%)	Maior impacto e reconhecimento da pesquisa	46 (50%)	Falta de infraestrutura tecnológica	53 (57,6%)
Contribuir para a ciência aberta	67 (72,8%)	Possibilitar a validação da pesquisa	49 (53,3%)	Falta de recursos financeiros e/ou pessoal	38 (41,3%)
Aumentar a visibilidade e impacto da pesquisa	55 (59,8%)	Colaborações interdisciplinares, intra e extra pares	54 (58,7%)	Falta de conhecimento em gestão de dados de pesquisa	42 (45,7%)
Exigências de agências de fomento ou instituições	21 (22,8%)	Evita duplicação de esforços	58 (63%)	Dificuldade em garantir a segurança dos dados	29 (31,5%)
		Contribui para o avanço científico	56 (60,9%)	Limitações de tempo	47 (51,1%)
		Dar suporte ao acesso aberto e aos dados abertos	43 (46,7%)	Não tenho hábito de disponibilizar os dados de pesquisa	19 (20,7%)
		Precaução contra fraude dos dados	25 (27,2%)		

Fonte: Bertoleza *et al.* (2024).

No que diz respeito a suas motivações quanto ao compartilhamento e reuso de dados, “Contribuir para a Ciência Aberta” foi o fator com maior número de

respostas (72,8%), seguido de “Promover a transparência na pesquisa científica” (60,9%) e “Aumentar a visibilidade e impacto da pesquisa” (59,8%), podendo-se assim inferir que o compartilhamento de dados pode ser motivado principalmente por razões de validação acadêmica e científica, bem como o comprometimento como livre acesso ao conhecimento. Nesse sentido, Curty (2016) afirma que dados disponíveis para reutilização costumam ser considerados credíveis e fiáveis por estarem sujeitos ao escrutínio e verificação. Além disso, a autora ressalta a importância do compartilhamento de dados para a sustentabilidade da Ciência Aberta.

Dos respondentes, 22,8% mencionaram exigências institucionais ou de agências de fomento como motivação para compartilhar seus dados de pesquisa. Assim, percebe-se que o compartilhamento e reuso de dados parece ser motivado, principalmente, por razões de prestígio científico e acadêmico, sendo as obrigações institucionais e motivações éticas fatores secundários. Nessa perspectiva, Costa e Cunha (2019) afirmam que o Brasil ainda carece de um posicionamento uniformizado para todas as instituições de pesquisa e pesquisadores quanto a GDP, o que só é possível a partir do momento em que o país possuir uma diretriz nacional, levando as instituições de pesquisa a elaborar políticas institucionais alinhadas, bem como as agências de fomento poderem cobrar o PGD dos pesquisadores. Essa afirmação é refletida nas respostas dadas pelos participantes desta pesquisa.

Quanto aos benefícios percebidos do compartilhamento e reuso de dados de pesquisa, as respostas indicam que evitar a duplicação de esforços (63%) e a promoção de colaborações interdisciplinares (58,7%) são os aspectos mais valorizados pelos pesquisadores participantes. Além disso, a validação da pesquisa (53,3%), o aumento do impacto acadêmico (50%) e dar suporte ao acesso aberto (46,7%) também tiveram um número alto de respostas, em alinhamento com a discussão em relação à motivação dos docentes para compartilharem seus dados de pesquisa.

Pesquisas como as de Curty (2016), Monteiro e Sant’Ana (2023), Araújo, Dias e Autran (2021), Lima (2020), Pasquetto, Randles e Borgman (2017) e Silva,

Prado e Araújo (2024) apontam para o mesmo caminho quanto aos benefícios do compartilhamento e reuso de dados de pesquisa. Porém, é importante ressaltar que os dados de pesquisas depositados em repositórios só têm o potencial de gerar valor a novas pesquisas se seu acesso estiver liberado e estiverem formatados em condições que viabilizem o seu uso (Estevão; Strauhs, 2020). Além disso, é a partir do reuso efetivo dos dados que se pode mensurar os benefícios do seu compartilhamento (Curty, 2016).

Quanto aos desafios enfrentados pelos pesquisadores para o compartilhamento de dados de pesquisa, a falta de infraestrutura tecnológica se mostrou como maior obstáculo, sendo 57,6% das respostas, podendo indicar que muitos pesquisadores não têm acesso a ferramentas adequadas para armazenamento e compartilhamento de dados. Limitações de tempo foi o segundo obstáculo com maior número de respostas (51,1%), o que pode ser entendido pela sobrecarga de trabalho que os docentes enfrentam, visto que a pesquisa já é uma atividade que exige muito esforço, além das atividades administrativas e de docência que a profissão acumula. Assim, o tempo necessário para organizar e disponibilizar os dados pode ser um obstáculo significativo.

Na sequência, falta de conhecimento em gestão de dados de pesquisa (45,7%) e falta de recursos financeiros e/ou pessoal (41,3%) aparecem como as maiores dificuldades ou barreiras para o compartilhamento de dados de pesquisa. Entende-se que muitos pesquisadores não possuem treinamento ou suporte suficiente para organizar e compartilhar seus dados corretamente. Além disso, a falta de recurso financeiro dificulta contratar especialistas e a aquisição de ferramentas necessárias para a gestão de dados. Nesse sentido, Lima (2020) destaca que, no Brasil, bibliotecas universitárias têm se preocupado com os dados de pesquisa, instituindo aos poucos serviços em GDP, a exemplo da Universidade de São Paulo (USP), Universidade Federal do Paraná (UFPR) e da Universidade Federal do Rio Grande do Sul (UFRGS). O investimento em capacitação de bibliotecários, bem como a oferta de serviços voltados para a GDP nas bibliotecas, é um passo fundamental para que barreiras como falta de conhecimento, de recursos financeiros e de pessoal sejam superadas.

Quanto a garantir a segurança dos dados, 31,5% dos respondentes afirmaram ser uma dificuldade, assim a preocupação com proteção contra acessos indevidos pode ser um motivo para o não compartilhamento dos dados, especialmente para dados sensíveis. Por outro lado, a precaução contra fraudes aparece como o benefício menos reconhecido pelos respondentes, sendo 27,2% das respostas. Em relação ao hábito de compartilhar dados de pesquisa, 20,7% dos pesquisadores não têm o hábito de disponibilizar seus dados. É importante destacar que, conforme Monteiro e Sant’Ana (2023), a adaptação dos pesquisadores a novas práticas, como o compartilhamento e reuso de dados de pesquisa, pode levar a tensões e enfraquecimento das relações sociais ou da ordem estabelecida nas comunidades científicas conforme a cultura em cada área (Monteiro; Sant’Ana, 2023).

Considerando que os pesquisadores percebem tanto as motivações e benefícios quanto o que pode ser considerado uma barreira no compartilhamento e reuso de dados, é preciso, então, compreender as principais preocupações. Algumas pesquisas, como de Lima (2020), Monteiro e Sant’Anna (2023) e Silva, Prado e Araújo (2024) pontuam que as práticas de compartilhamento de dados ainda estão em estado inicial no Brasil, e, naturalmente, é normal surgirem certas incertezas por parte dos pesquisadores em compartilhar amplamente seus dados.

Diante disso, as principais preocupações notadas neste estudo se relacionam com o **uso indevido ou distorção dos dados** (n=52) e com **questões éticas e legais** (n=51). Curty (2016) ressalta que o uso indevido dos dados se configura como uma preocupação recorrente dos pesquisadores, estando presentes em outras pesquisas. Ademais, salienta-se que existem dimensões éticas e técnicas para o reuso de dados, como pontuam Silva, Prado e Araújo (2024, p. 2), ao destacarem que os “[...] aspectos técnicos como a forma de estruturação e organização dos dados de pesquisa, e os aspectos éticos, como a disponibilidade do consentimento de uso, o anonimato e o reconhecimento do trabalho de terceiros”.

Somado a estas preocupações está envolvida a **falta de reconhecimento ou crédito** (n= 33). Esta preocupação atravessa também questões legais, uma vez

que perpassa o direito autoral de quem publicar e disponibilizar os dados, ou seja, espera-se que os autores sejam creditados, tanto por uma questão ética quanto legal. Ademais, considerando que uma boa prática em que há a devida citação “o compartilhamento dos conjuntos de dados, possibilita aumentar o número de citações e, conseqüentemente, o aumento das métricas e reconhecimento entre seus pares” (Monteiro; Sant’Anna, 2023, p. 15).

Um quantitativo menor considera como preocupação o **risco de perder a vantagem competitiva** (n=8). Isto posto, é válido salientar que “[...] em termos dos fatores que desestimulam o compartilhamento de dados de pesquisa, destacaram-se: a necessidade de publicar primeiro para garantir o caráter inédito”, entre outros, conforme pontuam Silva, Prado e Araújo (2024, p. 3). Atrelado a essas preocupações, existe ainda a possibilidade de dados estarem incompletos e terem seu compartilhamento inviável (Cheung *et al.*, 2022).

Interligando todas essas nuances sobre o compartilhamento e reuso de dados está o desenvolvimento do PGD. Silva (2019) esclarece sobre a importância desse documento para pesquisa, considerando tanto aspectos iniciais, quanto no seu desenvolvimento, envolvendo todo o ciclo de vida dos dados. Logo, questionou-se aos pesquisadores sobre o nível de conhecimento envolvendo o PGD. Os resultados estão expostos na Tabela 4. Com vistas a identificar os maiores e menores índices de respostas, optou-se por realizar um destaque em negrito.

Tabela 4 -Conhecimento dos participantes sobre PGD

PLANO DE GESTÃO DE DADOS				
Qual o seu nível de conhecimento sobre um Plano de Gestão de Dados (PGD)?				
Nenhum conhecimento	Conhecimento muito limitado	Conhecimento moderado	Bom conhecimento	Conhecimento avançado
15 (16,3%)	28 (30,4%)	30 (32,6%)	15 (16,3%)	4 (4,3%)

Fonte: Bertoleza *et al.* (2024).

Nota-se que boa parte dos pesquisadores tem entre nenhum conhecimento e conhecimento moderado sobre PGD (aproximadamente 80%), o que pode resultar em fragilidades com o ciclo de vida dos dados, e, conseqüentemente, nas práticas de compartilhamento e reuso. De maneira geral, os pesquisadores que têm conhecimento bom ou avançado nessa temática estão entre o grupo de faixa etária 36-45 anos (n=4), 46-55 anos (n= 4) e 56-65 anos (n = 7), ou seja, pesquisadores com mais experiência tem conhecimento nessa seara. Além disso, é perceptível que a maioria dos pesquisadores detém conhecimento moderado ou muito limitado, evidenciando a necessidade de aperfeiçoamento no tocante ao desenvolvimento de um PGD. Lima (2020, p. 78) pontua que:

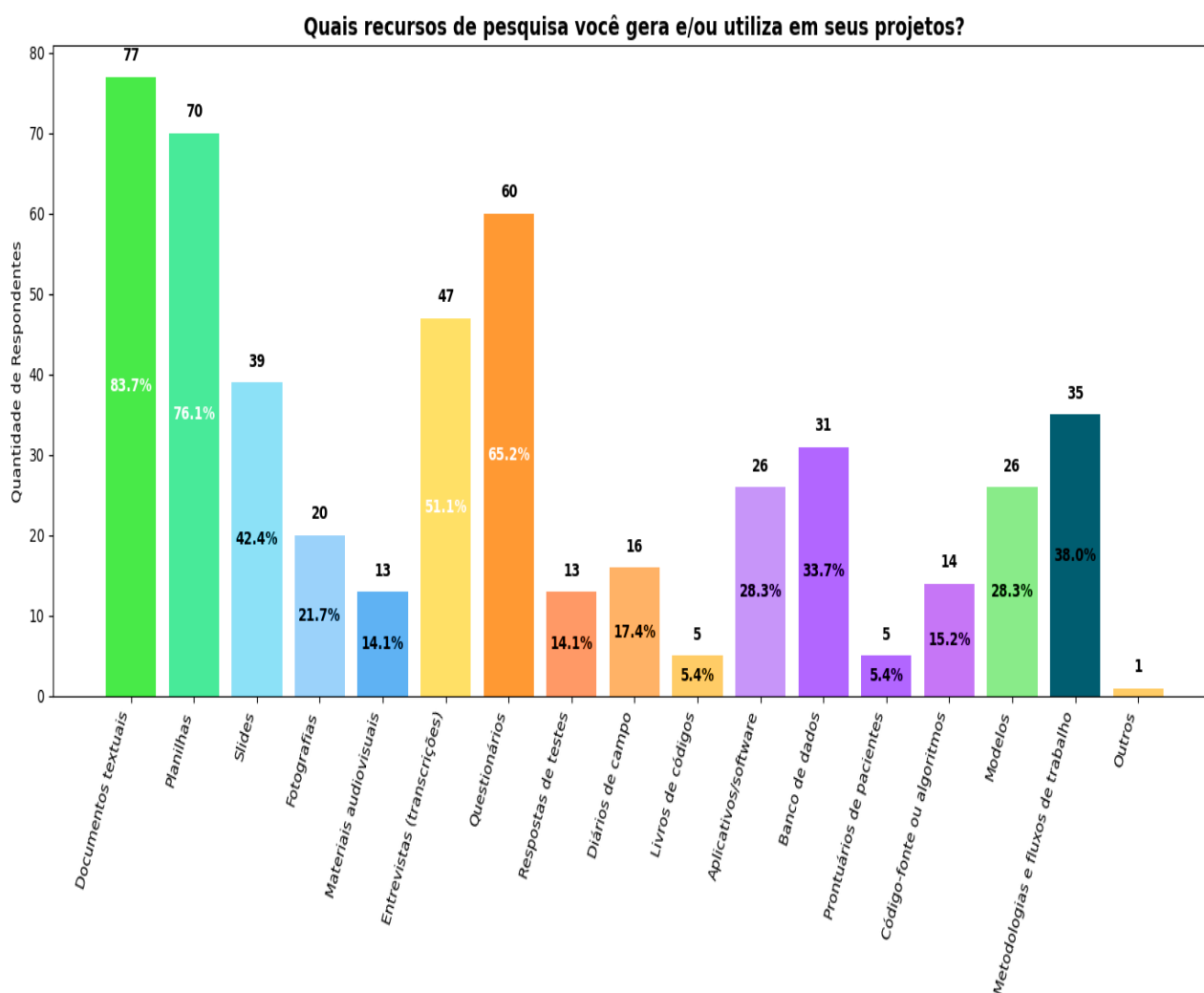
O PGD deve incluir informações suficientes que permita aos revisores das agências de fomento avaliarem os dados que serão gerados, sua adequação a pesquisa, além disso, deve refletir as melhores práticas em sua área de pesquisa. Assim sendo, não é exagero afirmar que o PGD é uma extensão da metodologia descrita em trabalhos acadêmicos ou em projetos de pesquisa, porém, de uma forma diferente, sucinta, contudo, extremamente detalhada e que representa o caráter da pesquisa e dos dados que serão produzidos.

Para além de uma boa prática de pesquisa, alinhada à Ciência Aberta, o PGD tem se configurado como um documento requerido por algumas agências de fomento, tanto nacionais quanto internacionais (Bunkar; Bhatt, 2020; Silva; Prado; Araújo, 2024). Isso requer, portanto, que os pesquisadores desenvolvam letramento informacional para dados (Estevão; Strauhs, 2020), de modo a terem habilidade e competência para o uso de dados em todo o seu ciclo de vida, inclusive na etapa de desenvolvimento do PGD. Ainda nesse respeito, Araújo, Dias e Autran (2021) destacam ferramentas que podem ser úteis para montar um PGD, como o *DMPonline*, que pode ter seu uso em todas as áreas do conhecimento.

Como o PGD refere-se à gestão de dados em todo o seu ciclo de vida, um dos questionamentos necessário de se fazer é a respeito dos dados que estão sendo coletados pelos pesquisadores da Ciência da Informação. Isto é, ao falar de PGD considera-se a organicidade deste documento, ou seja, ele pode se transformar e

sofrer alterações com o desenvolver de uma pesquisa (Lima, 2020). Assim sendo, buscou-se compreender quais tipos de dados estão sendo gerados pelos pesquisadores da CI. Os resultados deste questionamento estão expostos no Gráfico 2.

Gráfico 2 - Principais dados gerados pelos pesquisadores da CI



Fonte: Bertoleza *et al.* (2024).

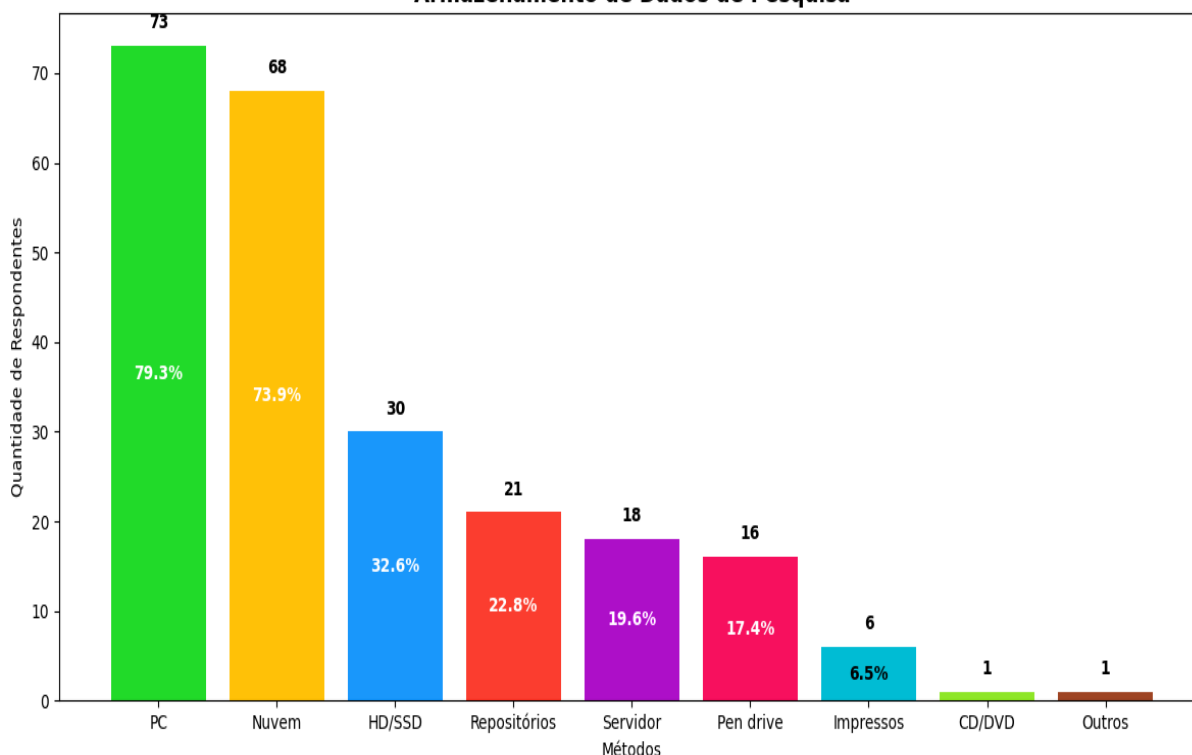
Os dados produzidos em maior incidência na CI foram documentos textuais (n=77), planilhas (n=70), questionários (n=60), entrevistas (transcrições) (n=47) e slides (n=39). Compreender o tipo de dado é essencial para pensar em sua preservação, seja em ambiente físico ou digital e, a posteriori, o seu

compartilhamento para possível reuso. Independente do tipo de dado, é preciso considerar “[...] as questões de segurança e integridade dos dados, verificação de falhas; utilização de padrões consolidados para os dados e metadados, tais como a criação de glossários, ontologias, taxonomias etc.” (Lima, 2020, p. 79).

Os dados da pesquisa evidenciam que os pesquisadores da Ciência da Informação utilizam diferentes métodos para armazenar seus dados de pesquisa, sendo o computador pessoal e o armazenamento em nuvem as opções predominantes. Outros dispositivos físicos, como *pen drives* e *HDs/SSDs* externos, continuam sendo amplamente utilizados, o que pode estar relacionado às preocupações com acessibilidade, privacidade ou hábitos consolidados. O Gráfico 3 apresenta as principais formas de armazenamento utilizadas na CI.

Gráfico 3 - Forma de armazenamento

Armazenamento de Dados de Pesquisa



Fonte: Bertoleza *et al.* (2024).

Esse comportamento é semelhante ao identificado por Cheung *et al.* (2022) em seu estudo no Canadá, no qual a maioria dos pesquisadores também

armazenavam seus dados em dispositivos pessoais, apresentando baixa adesão a repositórios institucionais. A pesquisa realizada por Bunkar e Bhatt (2020) na Índia reforça essa tendência, mostrando que o armazenamento em computadores pessoais e dispositivos removíveis é uma prática comum, evidenciando uma lacuna na implementação de políticas institucionais para gestão de dados.

Outro aspecto relevante identificado foi o baixo uso de repositórios institucionais para armazenamento e compartilhamento de dados. A maioria dos pesquisadores (n=64) afirmou não utilizar repositórios para essa finalidade, enquanto apenas 29 indicaram fazê-lo. Esse resultado sugere desafios como desconhecimento sobre a existência de repositórios, ausência de infraestrutura adequada ou baixa exigência institucional para o depósito de dados. De acordo com Cheung *et al.* (2022), muitos pesquisadores expressam interesse em utilizar repositórios no futuro, mas não o fazem atualmente devido à falta de diretrizes claras e suporte técnico. Essa mesma barreira foi identificada no estudo de Bunkar e Bhatt (2020), que apontaram que a ausência de treinamento e a falta de políticas bem estabelecidas nas universidades dificultam a adoção dessas práticas.

Além disso, os dados indicam que a maioria dos pesquisadores da Ciência da Informação nunca participou de capacitações sobre gestão de dados, com 63 registros apontando a ausência de treinamentos formais. Esse dado corrobora os achados de Cheung *et al.* (2022), que destacam a necessidade de maior investimento em treinamentos para auxiliar os pesquisadores a adotarem boas práticas de gestão de dados, especialmente na elaboração do PGD. Bunkar e Bhatt (2020), destarte, identificaram que muitos pesquisadores desconhecem a importância de capacitações para organização, preservação e compartilhamento de dados, reforçando a necessidade de iniciativas institucionais para promover o letramento informacional nesse contexto.

Porquanto, esses achados apontam para a necessidade de estratégias institucionais que incentivem o uso de repositórios e a capacitação dos pesquisadores para a adoção de boas práticas na gestão de dados.

5 Considerações finais

Este estudo buscou analisar as percepções e práticas relacionadas ao compartilhamento e reuso de dados de pesquisa entre docentes doutores vinculados aos Programas de Pós-Graduação em Ciência da Informação no Brasil, considerando a relevância da Gestão de Dados de Pesquisa no contexto contemporâneo da Ciência Aberta. A partir dos objetivos específicos delineados, verificou-se que os pesquisadores reconhecem amplamente a importância do compartilhamento e reuso de dados para o avanço do conhecimento científico, destacando benefícios como a transparência, a colaboração interdisciplinar e a otimização de recursos científicos.

Contudo, apesar dessa compreensão positiva sobre o tema, identificou-se que a prática regular do compartilhamento de dados ainda é incipiente e enfrenta barreiras significativas. A frequência com que os pesquisadores disponibilizam dados de pesquisa demonstrou-se de forma ocasional, sendo influenciada principalmente pela ausência de infraestrutura tecnológica adequada, limitações de tempo, escassez de recursos financeiros e pessoal capacitado. Esses resultados corroboram estudos anteriores no contexto brasileiro e internacional, indicando que desafios semelhantes são enfrentados por pesquisadores de diferentes áreas do conhecimento.

Quanto às motivações para compartilhar dados, destacou-se principalmente a adesão ao movimento da Ciência Aberta e a promoção da transparência científica. Entretanto, preocupações sobre o uso indevido dos dados, questões éticas, legais e a falta de reconhecimento ou crédito também emergiram como importantes fatores que inibem a plena adesão a essa prática. Esse cenário evidencia a necessidade de políticas institucionais e diretrizes que orientem e apoiem os pesquisadores em todo o ciclo de vida dos dados.

Outro achado relevante refere-se ao baixo nível de conhecimento e à pouca utilização prática dos Planos de Gestão de Dados entre os docentes investigados, o que sugere lacunas na formação e capacitação em GDP. Esse resultado reforça a relevância de investimentos perenes em ações formativas voltadas à

conscientização e capacitação de pesquisadores para fins de implementação e consolidação de práticas de gestão de dados.

Considerando os resultados obtidos, percebe-se a necessidade do fortalecimento das infraestruturas tecnológicas das instituições de ensino superior brasileiras, a criação e disseminação de políticas institucionais explícitas sobre GDP, bem como o incentivo à adoção sistemática dos PGD nas práticas acadêmicas cotidianas. Ademais, infere-se ser essencial o desenvolvimento de programas regulares de capacitação e sensibilização sobre a gestão de dados, visando a suprir as lacunas de conhecimento identificadas neste estudo.

A pesquisa trouxe novos *insights* acerca das percepções e práticas relacionadas ao compartilhamento e reuso de dados de pesquisa na Ciência da Informação, porém, entende-se que o assunto ainda é um campo que precisa ser aprofundado. Nesse sentido, espera-se e encoraja-se que os dados utilizados na presente pesquisa, através do compartilhamento em repositório, possam responder a novos questionamentos mediante futuras análises, destes e de outros pesquisadores que se debruçam sobre o tema.

Destaca-se que este estudo apresenta limitações decorrentes do recorte específico dos docentes vinculados aos Programas de Pós-Graduação em Ciência da Informação, sendo recomendadas futuras pesquisas que possam incluir outras áreas do conhecimento e diferentes perfis acadêmicos, permitindo uma visão mais ampla e comparativa sobre o cenário brasileiro da GDP.

Em suma, este trabalho reforça o entendimento de que a gestão, o compartilhamento e o reuso de dados são componentes fundamentais para o fortalecimento da Ciência Aberta no Brasil, contribuindo para a construção de uma ciência mais transparente, participativa e colaborativa.

Referências

ALBAGLI, Sarita. Ciência aberta: movimento de movimentos. *In*: SHINTAKU, Milton; SALES, Luana Farias (org.). **Ciência aberta para editores científicos**. Botucatu: ABEC, 2019. p. 15-20.

ARAÚJO, Débora Gomes; DIAS, Guilherme Ataíde; AUTRAN, Marynice Medeiros Matos. Compartilhamento de dados no contexto da ciência brasileira:

um estudo integrativo. **Informação & Informação**, Londrina, v. 26, n. 3, p. 100-124, 2021. Disponível em:
<https://doi.org/10.5433/1981-8920.2021v26n3p100>. Acesso em: 17 fev. 2025.

BERTOLEZA, Nayra Lorhanne Caldeira.; CHAVES, Italo Teixeira.;
FERNANDES, Marcus da Costa.; DIAS, Guilherme Ataíde.; SOUSA,
Marckson Roberto Ferreira de. **Percepções e práticas sobre dados abertos na
Ciência da Informação**. DataPB, 2024. Disponível em:
<https://doi.org/10.48472/DATAPB/LEYJR7>. Acesso em: 18 jul. 2025.

BUNKAR, Anjana R.; BHATT, Dhaval D. Perception of researchers &
academicians of Parul University towards research data management system &
role of library: a study. **Journal of Library & Information Technology**, Delhi,
v. 40, n. 3, p. 139-146 2020. Disponível em:
<https://doi.org/10.14429/djlit.40.03.15302>. Acesso em: 16 fev. 2025.

COSTA, Máira Murrieta; CUNHA, Murilo Bastos. A necessidade de uma
política nacional para a gestão de dados de pesquisa no Brasil. **Liinc em
Revista**, Rio de Janeiro, v.15, n. 2, p. 287-309, 2019. Disponível em:
<https://doi.org/10.18617/liinc.v15i2.4763>. Acesso em: 14 jan. 2025.

CHEUNG, Melissa. *et al.* Practices before policy: research data management
behaviours in Canada. **Partnership: The Canadian Journal of Library and
Information Practice and Research**, Ontário, v. 17, n. 1, p. 1-80, 2022.
Disponível em: <https://doi.org/10.21083/partnership.v17i1.6779>. Acesso em: 24
fev. 2025.

CURTY, Renata Gonçalves. As diferentes dimensões do reuso de dados
científicos. **Tendências da Pesquisa Brasileira em Ciência da Informação**,
Rio de Janeiro, v. 9, n. 2, p. 1-28, 2016.

DIAS, Guilherme Ataíde *et al.* Garbage In, Garbage Out (GIGO): enfrentando
esta máxima nos conjuntos de dados associados ao Programa Dinheiro Direto na
Escola (PDDE). *In: WORKSHOP DE INFORMAÇÃO DADOS E
TECNOLOGIA*, 6, 2023, Brasília. **Anais [...]**. Brasília: Ibict, 2023. p. 1-11.

ESTEVÃO, Janete Saldanha Bach; STRAUHS, Faimara Rocio. Letramento
informacional para reuso de dados nas ciências sociais: requisitos e
competências. **Informação & Informação**, Londrina, v. 25, n. 2, p. 1-25, 2020.
Disponível em: <https://doi.org/10.5433/1981-8920.2020v25n2p1>. Acesso em: 17
fev. 2025.

GERHARDT, Tatiana Engel; SOUZA, Aline Corrêa. Aspectos teóricos e
conceituais. *In: GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo (org.).
Métodos de pesquisa*. Porto Alegre: Ed. UFRGS, 2009. p. 11-30.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

LIMA, Juliana Soares. **Gestão de dados de pesquisa no contexto da Ciência Aberta**: percepção dos pesquisadores da Universidade Federal do Ceará. 2020. Dissertação (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Centro de Humanidades, Universidade Federal do Ceará, Fortaleza, 2020.

MONTEIRO, Elizabete Cristina Souza Aguiar; SANT'ANA, Ricardo César Gonçalves. Desafios no compartilhamento de dados: anomia relacionada aos direitos autorais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 28, p. 1-24, 2023. Disponível em: <https://doi.org/10.1590/1981-5344/41032>. Acesso em: 17 fev. 2025.

PASQUETTO, Irene V.; RANGLES, Bernadette M.; BORGMAN, Christine L. On the reuse of scientific data. **Data Science Journal**, Paris, v. 16, p. 1-9, 2017. Disponível em: <https://doi.org/10.5334/dsj-2017-008>. Acesso em: 17 fev. 2025.

RODRIGUES, Débora Gomes Araújo *et al.* Modelo referencial de ciclo de vida dos dados científicos: uma proposta metodológica. **Informação & Informação**, Londrina, v. 28, n. 3, p. 55-80, 2023. Disponível em: <https://doi.org/10.5433/1981-8920.2023v28n3p55>. Acesso em: 16 fev. 2025.

SAMPIERI, Roberto Hernández; COLLADO, Carlos Fernández; LUCIO, Pilar Baptista. **Metodologia de pesquisa**. 5. ed. Porto Alegre: Penso, 2013.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da Ciência da Informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116-142, 2016. Disponível em: <https://doi.org/10.5433/1981-8920.2016v21n2p116>. Acesso em: 16 fev. 2025.

SANTOS, Plácida Leopoldina Ventura Costa; SANT'ANA, Ricardo César Gonçalves. Camadas de representação de dados e suas especificidades no cenário científico. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire (org.). **Dados científicos: perspectivas e desafios**. João Pessoa: Ed. UFPB, 2019. p. 53-66.

SAYÃO, Luís Fernando; SALES, Luana Farias; CARVALHO SEGUNDO, Washington Luís R. Objeto Digital Fair: ampliando as fronteiras da gestão de dados. **Informação & Sociedade: Estudos**, João Pessoa, v. 33, p. 1-27, 2023. Disponível em: <https://doi.org/10.22478/ufpb.1809-4783.2023v33.68646>. Acesso em: 27 jan. 2025.

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

SILVA, Luciana Candida; PRADO, Cássia Braga; ARAÚJO, Eliany Alvarenga. Fatores que influenciam o processo de compartilhamento e reuso de dados de pesquisa: revisão sistemática de literatura a partir da base de dados BRAPCI.

AtoZ: Novas Práticas em Informação e Conhecimento, Curitiba, v. 13, p. 1-15, 2024. Disponível em: <https://doi.org/10.5380/atoz.v13i0.91352>. Acesso em: 17 fev. 2025.

SILVA SEGUNDO, Sanderli José; ARAÚJO, Wagner Junqueira. Curadoria e ciclo de vida dos dados. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire (org.). **Dados científicos: perspectivas e desafios**. João Pessoa: Ed. UFPB, 2019. p. 113-152.

SILVEIRA, Denise Tolfo; CORDOVA, Fernanda Peixoto. A pesquisa científica. In: GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo (org.). **Métodos de pesquisa**. Porto Alegre: Ed. UFRGS, 2009. p. 31-42.

SILVEIRA, Lucia *et al.* Ciência aberta na perspectiva de especialistas brasileiros: proposta de taxonomia. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 26, p. 1-27, 2021. Disponível em: <https://doi.org/10.5007/1518-2924.2021.e79646>. Acesso em: 27 fev. 2025.

TORINO, Emanuelle; MONTEIRO, Elizabete Cristina Souza Aguiar; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Plano de gestão de dados de pesquisa de povos indígenas: considerações acerca dos princípios FAIR e CARE. **Revista Brasileira de Preservação Digital**, Campinas, v. 4, p.1-23, 2023. Disponível em: <https://doi.org/10.20396/rebpred.v4i00.17936>. Acesso em: 17 fev. 2025.

VEIGA, Viviane Santos Oliveira *et al.* **Plano de gestão de dados de pesquisa: guia de elaboração**. Rio de Janeiro: Fiocruz, 2022.

VIDOTTI, Silvana Aparecida Borsetti Gregorio; TORINO, Emanuelle; CONEGLIAN, Caio Saraiva. #SejaJUSTOeCUIDADOSO: princípios FAIR e CARE na gestão de dados de pesquisa. In: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 201-214.

WILKINSON, Mark D. *et al.* The FAIR guiding principles for scientific data management and stewardship. **Scientific Data**, New York, v. 3, n. 160018, p. 1-9, 2016. Disponível em: <https://doi.org/10.1038/sdata.2016.18>. Acesso em: 16 fev. 2025.

Perceptions and practices in data sharing and reuse: analysis of researchers affiliated with Graduate Programs in Information Science in Brazil

Abstract: This study aims to investigate the perceptions and practices of PhD teacher affiliated with academic Postgraduate Programs in Information Science in Brazil regarding the sharing and reuse of research data. It adopts an exploratory and applied approach, using an electronic questionnaire for data collection. The results indicate that, although researchers recognize the importance of data sharing for transparency and scientific collaboration, the practice still faces challenges, such as lack of infrastructure, time constraints, and scarcity of resources. The lack of knowledge about data management and the absence of well-established institutional policies also stand out as significant barriers. The research reveals that most participants store their data on personal devices and few use institutional repositories. Likewise, adherence to data management plans is still limited, which may compromise the preservation and reuse of data in the long term. It is concluded that, in order to expand data sharing and reuse practices, it is essential to strengthen institutional infrastructures, promote training, and develop clear guidelines on research data management. The findings of this study contribute to the understanding of the challenges faced by researchers and can support policies and actions that encourage Open Science in Brazil.

Keywords: research data management; data sharing; data reuse; open science; Information Science

Financiamento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Processo 311563/2018-0).

Declaração de autoria

Concepção e elaboração do estudo: Italo Teixeira Chaves, Marcus da Costa Fernandes, Nayra Lorhanne Caldeira Bertoleza, Guilherme Ataíde Dias, Marckson Roberto Ferreira de Sousa

Coleta de dados: Italo Teixeira Chaves, Marcus da Costa Fernandes, Nayra Lorhanne Caldeira Bertoleza

Análise e interpretação de dados: Italo Teixeira Chaves, Marcus da Costa

Fernandes, Nayra Lorhanne Caldeira Bertoleza

Redação: Italo Teixeira Chaves, Marcus da Costa Fernandes, Nayra Lorhanne Caldeira Bertoleza

Revisão crítica do manuscrito: Guilherme Ataíde Dias, Marckson Roberto Ferreira de Sousa

Declaração de disponibilidade de dados

Todo o conjunto de dados que dá suporte aos resultados deste estudo está disponível em: BERTOLEZA, Nayra Lorhanne Caldeira.; CHAVES, Italo Teixeira.; FERNANDES, Marcus da Costa.; DIAS, Guilherme Ataíde.; SOUSA, Marckson Roberto Ferreira de. Percepções e práticas sobre dados abertos na Ciência da Informação. **DataPB**, 2024. Dados de pesquisa. Disponível em: <https://doi.org/10.48472/DATAPB/LEYJR7>.

Autoria para correspondência

Italo Teixeira Chaves
italochaves55@hotmail.com

Editor-chefe

Thiago Henrique Bragato Barros

Como citar

CHAVES, Italo Teixeira; FERNANDES, Marcus da Costa; BERTOLEZA, Nayra Lorhanne Caldeira; DIAS, Guilherme Ataíde; SOUSA, Marckson Roberto Ferreira. Percepções e práticas no âmbito do compartilhamento e reuso de dados: análise dos pesquisadores vinculados aos Programas de Pós-Graduação em Ciência da Informação do Brasil. **Em Questão**, Porto Alegre, v. 31, e-146102, 2025. DOI: <https://doi.org/10.1590/1808-5245.31.146102>.

Parecer(es) aberto(s):

<https://doi.org/10.1590/1808-5245.31.146102A>

Recebido: 28/02/2025

Aceito: 03/07/2025



¹ Esta investigação foi aprovada no Comitê de Ética da Pesquisa (CEP), sob o seguinte protocolo: 83897824.1.0000.5188.