

## Gestão de dados em pesquisas científicas: abordagens éticas e práticas de integridade

Ana Paula Lopes Marinho<sup>1\*</sup>; Alex Nunes de Almeida<sup>2</sup>

### SOBRE OS AUTORES

<sup>1</sup> Mestre em Administração. Centro Estadual de Educação Tecnológica Paula Souza. Praça Coronel Fernando Prestes, 30, Bom Retiro, 01124-060, São Paulo, São Paulo, Brasil.

<sup>2</sup> Doutor em Ciências. Pecege. Rua Cezira Giovanoni Moretti, 580, Santa Rosa, 13414-157, Piracicaba, São Paulo, Brasil

\*Autor correspondente: anapaulalopesmarinho@gmail.com

### COMO CITAR

Marinho A.P.L.; Almeida A.N. Gestão de dados em pesquisas científicas: abordagens éticas e práticas de integridade. Revista E&S. 2025; 6: e20240052



Com a crescente disponibilização de dados digitais e o uso de métodos de Inteligência Artificial (IA) e Aprendizado de Máquina (AM) nas ciências de dados, surge um novo paradigma para a gestão e segurança das informações. A produção científica, baseada na coleta, manipulação e compartilhamento de dados por pesquisadores<sup>[1]</sup>, torna a Gestão de Dados (GD) essencial. A qualidade e a transparência das informações devem ser aprimoradas, garantindo técnicas adequadas para replicação<sup>[2]</sup>, especialmente em sistemas que utilizam Grandes Modelos de Linguagem (GML), como o ChatGPT<sup>[3,4]</sup>. Entretanto, a implementação de boas práticas de GD enfrenta desafios significativos, especialmente na necessidade de garantir a conformidade com leis de proteção de dados e a promoção de ações transparentes<sup>[2]</sup>.

A aplicação de IA e AM tem automatizado diversas áreas, mas enfrenta riscos de vieses estatísticos pela falta de diversidade nos modelos, gerando desconfiança em relação aos seus resultados<sup>[4]</sup>. Este estudo teve como objetivo aprimorar boas práticas em GD e integridade na pesquisa, mapeando ferramentas que auxiliam pesquisadores no gerenciamento de informações científicas. Assim, foi apresentada a proposta inicial de um modelo abrangente para um Plano de Gestão de Dados (PGD), integrando métodos e procedimentos globais em uma abordagem unificada.

A implementação de boas práticas de GD é desafiadora, pois pesquisadores enfrentam dificuldades em cumprir legislações que assegurem a reprodutibilidade e integridade dos dados, fundamentais para atender requisitos legais e critérios epistêmicos rigorosos<sup>[5,6]</sup>. Algumas plataformas, pagas ou gratuitas, oferecem sistemas de PGD que permitem aos pesquisadores preencher informações, especialmente em projetos financiados<sup>[7]</sup>. Contudo, para assegurar qualidade e transparência dos dados é preciso ampliar o uso dessas ferramentas<sup>[3]</sup>. Para discutir a importância da GD nas pesquisas, relacionando valores éticos e práticas de integridade, a literatura sugere analisar como os pesquisadores adquirem e valorizam conhecimento, realizam tarefas e atribuem responsabilidades<sup>[1]</sup>.

A questão central do estudo é: como implementar efetivamente um PGD, considerando valores éticos e práticas de integridade? O que busca avançar nas discussões sobre boas práticas de GD, destacando como o PGD pode garantir qualidade e transparência das informações, e incentivando o uso de ferramentas para o gerenciamento de dados científicos, visando a disseminação e reprodutibilidade do conhecimento.

A metodologia foi estruturada em duas etapas principais de modo a equilibrar a abrangência do conhecimento científico para comunicar a essência da pesquisa. Na primeira, realizou-se uma revisão bibliométrica, dividida em três fases (Tabela 1), para identificar tendências, padrões e lacunas em áreas específicas do conhecimento<sup>[8]</sup>. Essa busca preliminar foi fundamental para estabelecer uma base sólida de literatura, refinando os critérios de inclusão e exclusão e identificando palavras-chave e fontes relevantes, garantindo que a pesquisa fosse embasada em estudos com impacto significativo na área. Contudo, a busca na base Scopus<sup>®</sup> revelou que há poucos estudos sobre Data Science com foco em integridade em PGD.

**Tabela 1.** Resultados consolidados da pesquisa bibliográfica

Fases	Estudos preliminares	Estudos via Scopus	Estudos via Snowball	Total de amostra final
Resultados pesquisa bibliográfica	6	6	5	17

Fonte: Elaborado pelos autores

Diante dessa dificuldade, utilizou-se a estratégia de amostragem em “bola de neve” (*snowball*) para ampliar a base de referências, conectando redes de pesquisadores e identificando pesquisas adicionais<sup>[9]</sup>, facilitando o acesso a dados interconectados conforme as necessidades do projeto. Para a estratégia

*snowball*, utilizou-se a plataforma Research Rabbit® como ferramenta exploratória, conforme descrito no protocolo da revisão (Tabela 2). A escolha atende às exigências éticas e regulamentares.

**Tabela 2.** Protocolo de pesquisa da revisão sistemática – “snowball sampling”

Atividades	Descrição
Escolha das palavras-chaves	A pesquisa teve dois conjuntos de palavras chaves: Integridade em Pesquisa; Plano de Gestão de Dados
Acesso a artigos e leitura dos Abstracts	Identificação de artigos pertinentes ao tema de pesquisa relacionados aos termos de busca.
Exploração de Referências e Citações	Ao escolher as referências e citações dos artigos iniciais, permitiu-se que novas fontes fosse identificada a partir da pesquisa inicial, combinando outras novas a partir de dados representativos.
Ampliação da busca	A partir das novas referências e citações encontradas, ampliou-se a busca na plataforma incluindo novas fontes adicionais.
Análise e síntese	Avaliação das informações e síntese. O processo do “snowballing” repetiu-se para expandir a pesquisa e aumentar a compreensão do tema.
Registro e gestão das referências	Ao longo do processo, as referências encontradas foram salvas na plataforma Research Rabbit®. Isso facilitou a organização e o acompanhamento dos artigos utilizados no TCC.

Fonte: Elaborado pelos autores com base em Naderifar e Ghaljaie<sup>[9]</sup>.

Foram incluídos trabalhos de Cardoso et al.<sup>[7]</sup>, Ministério da Saúde<sup>[10]</sup>, Doorn<sup>[3]</sup>, Engelhardt<sup>[11]</sup>, e Governo Digital<sup>[12]</sup>. Para atualizar os dados, a segunda fase utilizou a base Scopus (Alam et al.<sup>[6]</sup>; Bolland & Grey<sup>[13]</sup>; Bouter<sup>[14]</sup>; Carcausto-Calla et al.<sup>[15]</sup>; Valkenburg et al.<sup>[1]</sup>; Van den Hoven et al.<sup>[16]</sup>). Em seguida, ampliou-se a busca com a técnica de *snowball sampling*<sup>[9]</sup> para obter amostras qualitativas e identificar lacunas sobre integridade na pesquisa (Beugelsdijk & Meyer<sup>[2]</sup>; McGovern et al.<sup>[4]</sup>; Hosseini et al.<sup>[5]</sup>; Wilkinson et al.<sup>[17]</sup>; Van den Eynden et al.<sup>[18]</sup>).

Na segunda etapa, foram mapeados repositórios e plataformas de GP científicos, nacionais e internacionais, regulamentados e recomendados por editores comprometidos com a integridade desta pesquisa. Foram realizadas uma busca no Google Scholar® e consultas em revistas de universidades para verificar políticas de submissão, integridade e transparência de dados em seus sites. Foram analisados estudos de caso empresariais e acadêmicos que evidenciam erros na GD, especialmente quando algoritmos de IA, baseados em fórmulas matemáticas, são erroneamente considerados infalíveis. Esses casos, conduzidos por especialistas, destacam a necessidade do uso de ferramentas de GD para garantir a replicabilidade dos estudos.

O acesso restrito a recursos abertos é um desafio significativo que dificulta a disseminação e validação dos resultados<sup>[5,2]</sup>. A falta de acessibilidade aberta prejudica a replicabilidade, criando barreiras para a validação das pesquisas<sup>[4]</sup>, pois, mesmo com verificações automatizadas, falhas persistem, permitindo práticas irresponsáveis como falsificação ou plágio de informações<sup>[1]</sup>. Na Tabela 3, destaca-se as limitações e potencialidades enfrentadas no acesso e replicabilidade de dados na ciência computacional.

A educação sobre integridade enfrenta uma lacuna no entendimento de ética e replicabilidade, evidenciando a necessidade de treinamentos específicos<sup>[2,6,14]</sup>. Estudos de Carcausto-Calla et al.<sup>[15]</sup> e Van den Hoven et al.<sup>[16]</sup> apontam resultados positivos com treinamentos estruturados que utilizam metodologias ativas, como oficinas, análise de casos, jogos de simulação e até a pedagogia freiriana, promovendo autonomia crítica e integridade na pesquisa.

**Tabela 3. Limitações e potencialidades de acesso a recursos de acesso aberto**

Limitações	Autor(es)
Acesso a infraestruturas gratuitas e de alto desempenho para realização de pesquisas e criação de PGD's	McGovern et al. <sup>[4]</sup> ; Beugelsdijk et al. <sup>[2]</sup>
Desenvolvimento de aplicações simples, transparentes de AM e com infraestrutura de verificação de prevenção de plágio	Valkenburg et al. <sup>[1]</sup>
O uso de LLMs possui dilemas quanto ao emprego da transparência. A responsabilidade é essencial para garantir a integridade, rigor legal para a replicabilidade dos dados.	Hosseini et al. <sup>[5]</sup> ; Beugelsdijk et al. <sup>[2]</sup>
Conhecimento científico falsos e plagiados que entram em bases de dados.	Valkenburg et al. <sup>[1]</sup>
A cultura de conhecimento compartilhado (Creative Commons) cresce, mas ainda enfrentamos desafios com licenciamento de domínio público e custos para liberar o acesso à pesquisa.	Beugelsdijk et al. <sup>[2]</sup>
A principal limitação à compreensão pública do processo científico é a falta de treinamento, infraestrutura adequada, relatórios consistentes, procedimentos claros e atenção à ética de publicação.	Alam et al. <sup>[6]</sup> ; Bouter <sup>[14]</sup>
Os pesquisadores precisam de ajuda de suas instituições de pesquisa para otimizar o funcionamento de sua bússola moral.	Bouter <sup>[14]</sup>
Há poucas oportunidades comunidades interessadas possam interagir no sentido de buscar colaborativamente soluções desafiadoras aos problemas de integridade na pesquisa.	Alam et al. <sup>[6]</sup>
Potencialidades	Autor(es)
A utilização do LLMs como ferramenta para facilitar a redação ou edição de textos é uma prática que promove a inclusão e diversidade.	Hosseini et al. <sup>[5]</sup>
A maioria das revistas científicas adotou boas práticas de pesquisa, incorporando processos de revisão e códigos de ética que promovem o acesso aos dados e a transparência da pesquisa, conhecidos em inglês como Data Access and Research Transparency (DART).	Beugelsdijk et al. <sup>[2]</sup>
Comitês de Ética são responsáveis por criar, disseminar e aprovar os protocolos de pesquisa. Prática disseminada em muitos países, incluindo o Brasil.	Beugelsdijk et al. <sup>[2]</sup> ; Ministério da Saúde <sup>[10]</sup>
Lei de Proteção de dados propõe e estimula que os pesquisadores elaborem um PGD com os detalhes dos tipos de dados coletados, como foram analisados e armazenados.	Beugelsdijk et al. <sup>[2]</sup>
Programas de treinamento desenvolvidos com uso de metodologias ativas para desenvolvimento de integridade e pesquisa para comunidade científica.	Carcausto-Calla et al. <sup>[15]</sup>
Evidências mostram que os treinamentos fortaleceram conceitos normativos, promoveram autonomia crítica, abertura, autoconfiança e coragem para lidar com questões de integridade na prática.	Van den Hoven et al. <sup>[16]</sup>

Fonte: Elaborado pelos autores

Nas publicações acadêmicas, há uma crescente adoção de políticas de transparência e acesso a dados, com muitas revistas integrando o Comitê de Ética em Publicação (COPE). No entanto, como apontam Beugelsdijk et al.<sup>[2]</sup> e Bolland et al.<sup>[13]</sup>, ainda que o COPE não defina padrões metodológicos específicos, essa mudança visa uma cultura de pesquisa mais ética, mas o uso de recursos "Creative Commons" pode enfrentar desafios financeiros. A Lei Geral de Proteção de Dados no Brasil e a legislação correspondente na União Europeia, sancionadas em 2018, reforçam a necessidade de um PGD, que é avaliado pelos Comitês de Ética em Pesquisa (CEP). O PGD deve respeitar as políticas específicas da revista, CEP e leis de proteção de dados.

Na Tabela 4, observa-se que a integridade na pesquisa é vista como uma responsabilidade individual dos pesquisadores<sup>[1]</sup> o que requer cuidado em todas as etapas, ressaltando a transparência e segurança na gestão de informações, independentemente das tecnologias empregadas<sup>[2,6,16]</sup>, pois o compromisso com a integridade exige conduta responsável dos pesquisadores.

**Tabela 4.** Síntese teórica sobre a integridade na pesquisa

Autores	Conceito de integridade na pesquisa
Valkenburg et al. <sup>[1]</sup>	A integridade na pesquisa é vista pelos autores como uma questão de responsabilidade dos pesquisadores em termos individuais em relação as suas respectivas atividades na pesquisa.
Beugelsdijk et al. <sup>[2]</sup>	A transparência é condição básica na pesquisa científica que decorre desde a etapa de produção de dados até o resultado das análises, desde a produção de dados até os resultados das análises, acumulando credibilidade no desenvolvimento da teoria sobre o fenômeno.
Alam et al. <sup>[6]</sup>	De acordo com os autores a integridade na pesquisa consiste em agregar o valor e benefícios sociais a pesquisa. A aplicação da integridade na pesquisa é marcada pela adesão de forma consistente de honestidade, responsabilidade, cortesia profissional, justiça e boa gestão.
Van den Hoven et al. <sup>[16]</sup>	Compreende que integridade na pesquisa é o efeito de estimular o empoderamento em direção à conduta responsável de pesquisa.

Fonte: Elaborado pelos autores

A maioria dos periódicos científicos exige que os autores enviem documentos separados, como tabelas e figuras, detalhando a base de dados. Também recomendam disponibilizar os dados em repositórios legais, acessíveis via “Digital Object Identifier” (DOI), permitindo citações como bibliografia formal<sup>[2]</sup>. A identificação por DOI de um artigo conecta-o ao pesquisador ou à organização de pesquisa. Algumas plataformas cruzam dados exigindo login com o identificador “Open Researcher and Contributor ID” (ORCID), uma iniciativa sem fins lucrativos que fornece um identificador único e permanente para pesquisadores e colaboradores acadêmicos<sup>[19]</sup>. Utilizado para validar a inclusão de informações associadas ao DOI, essa abordagem promove transparência na pesquisa, embora nem todos os sistemas ofereçam essa integração<sup>[6]</sup>.

### Plataformas e repositórios pesquisados

Foram consultados repositórios que fornecem DOI, com destaque para os recomendados pelos revisores e que oferecem uma interface de metadados padronizada e acessível globalmente para exportação de dados<sup>[8,20]</sup>. Para cada repositório listado, é apresentada uma breve explicação sobre sua função principal no contexto da replicabilidade de dados (Tabela 5).

**Tabela 5.** Repositórios indicados por revisores que fornecem DOI

Nome do repositório	Dados de acesso	Função
DataCite Commons	<a href="https://repositoryfinder.datacite.org">https:// repositoryfinder.datacite.org</a>	Permite ao pesquisador realizar pesquisas, acessar e reutilizar conjuntos de dados
Dryad	<a href="https://datadryad.org/stash">https://datadryad.org/stash</a>	Repositório de dados que oferece armazenamento e compartilhamento de dados científicos
Harvard Dataverse	<a href="https://dataverse.org/">https://dataverse.org/</a>	Permite que os pesquisadores compartilhem, preservem e citem seus dados
Mendeley	<a href="http://www.elsevier.com/solutions/mendeley">www.elsevier.com/solutions/mendeley</a>	Gerenciamento de referências e rede social acadêmica. Apresenta funções de organização de documentos e redes de pesquisa
ROR - Research Organization Registry	<a href="https://ror.org/">https://ror.org/</a>	Registro global de identificadores únicos para organizações de pesquisa

Fonte: Elaborado pelos autores

Diversas ferramentas para a elaboração de PGD, ou Data Management Plans (DMP) em inglês, foram investigadas. A seguir, listamos as ferramentas identificadas nesta pesquisa, acompanhadas de uma breve explicação (Tabela 6).

**Tabela 6.** Ferramentas para elaboração de Plano de Gestão de Dados (DMPtools)

Nome	Dados de acesso	Função
DCC – Digital Curation Centre	<a href="https://www.dcc.ac.uk/">https://www.dcc.ac.uk/</a>	Centro de curadoria digital que oferece recursos e orientações para a GD de pesquisa
DMPtool	<a href="https://dmptool.org/auth">https://dmptool.org/auth</a>	Ferramenta online que auxilia os pesquisadores na elaboração de planos de GD. Ferramenta utilizada por várias universidades
DMPTuuli	<a href="https://www.dmptuuli.fi/">https://www.dmptuuli.fi/</a>	Similar ao DMPtool, porém específica para pesquisadores finlandeses
Mendeley	<a href="https://data.mendeley.com/">https://data.mendeley.com/</a>	Embora mais conhecido como serviço de gerenciamento de referências. Oferece recursos para a GD de pesquisa
USGS – Science for a Changing World	<a href="https://www.usgs.gov/data-management">https://www.usgs.gov/data-management</a>	Agência governamental dos Estados Unidos que se dedica à gestão e análise de dados geoespaciais e ambientais

Fonte: Elaborado pelos autores

## Modelo de Plano de Gestão de Dados

O PGD é um documento essencial que organiza e detalha o desenvolvimento da pesquisa. Ele deve incluir a descrição dos protocolos de coleta de dados, a avaliação de riscos e benefícios, e os procedimentos necessários para garantir a confidencialidade e privacidade das informações, além das práticas de retenção<sup>[2]</sup>. A gestão eficaz de dados garante sua curadoria, beneficiando a comunidade científica e cumprindo as diretrizes de governança estabelecidas por órgãos reguladores<sup>[12]</sup>.

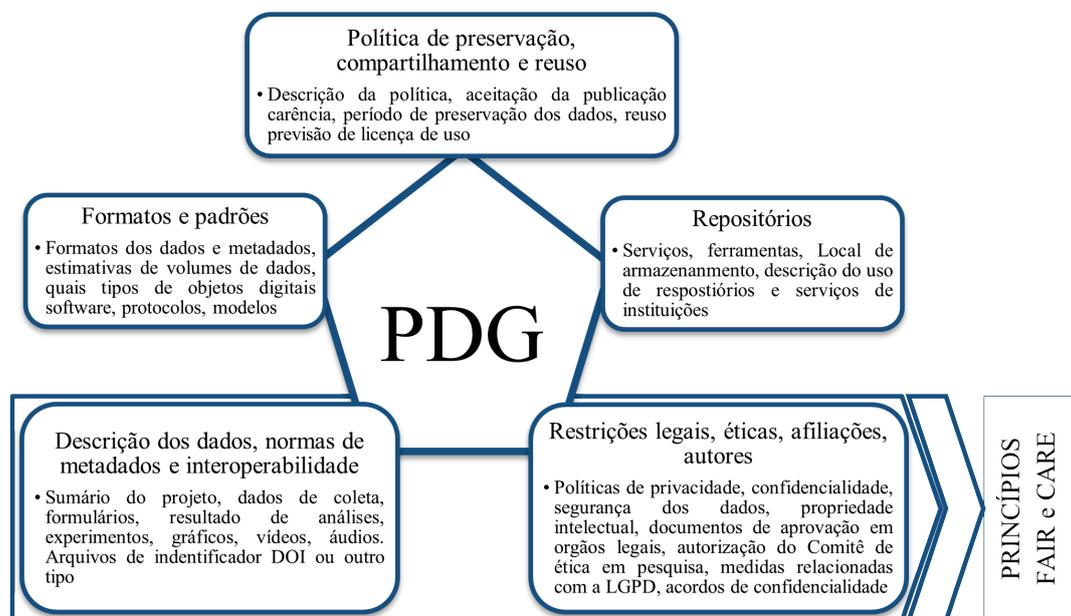
Um PGD bem elaborado garante o arquivamento adequado dos dados, definindo diretrizes para sua preservação e facilitando o uso por futuros pesquisadores. Conforme Wilkinson et al.<sup>[17]</sup>, um PGD converte metadados em conjuntos de dados aptos a receber um identificador DOI. Para tanto, decisões estratégicas, como o uso de ferramentas computacionais, são essenciais.

A governança dos dados deve seguir os princípios FAIR — Findable (localizáveis), Accessible (acessíveis), Interoperable (interoperáveis) e Reusable (reutilizáveis) — para fortalecer a transparência e integridade dos dados, além de promover a colaboração acadêmica<sup>[11, 8]</sup>. Complementarmente, os princípios CARE (Coletivo, Autoridade, Responsabilidade e Ética), introduzidos por Souza et al.<sup>[21]</sup>, focam na gestão ética de dados, especialmente para populações vulneráveis. Criado na Conferência Internacional de Dados, o princípio CARE foi originalmente voltado ao tratamento de dados indígenas, mas se estende a todas as populações nas quais estudos científicos são baseados.

Segundo Van den Eynden<sup>[21]</sup> e Beugelsdijk et al.<sup>[2]</sup>, a implementação de boas práticas de GD exige uma abordagem metódica e organizada. O modelo de PGD deve alinhar-se aos princípios FAIR e CARE (Figura 1) e ser desenvolvido seguindo as seguintes etapas:

- 1) Descrição dos dados, normas de metadados e interoperabilidade: o processo exige informações detalhadas sobre os dados utilizados na pesquisa, garantindo a interoperabilidade. Os dados devem estar disponíveis em uma plataforma compatível que permita sua interação.
- 2) Restrições legais, éticas, afiliações e autores: pesquisadores devem cumprir leis e regulamentos sobre uso e compartilhamento de dados, como proteção de dados, direitos autorais e propriedade intelectual, considerando também limitações de comitês de ética e questões éticas relacionadas à privacidade e consentimento dos participantes.
- 3) Política de preservação, compartilhamento e reuso dos dados: envolve a preservação, compartilhamento e reutilização dos dados, baseada em práticas de armazenamento que garantem acesso a longo prazo. Isso inclui sistemas digitais, backups regulares, documentação adequada e padrões de preservação para evitar perdas e acessos inadequados.

- 4) Formatos e padrões: um PGD bem estruturado deve indicar os formatos de arquivo e a estrutura dos dados, informações essenciais para garantir a interoperabilidade e portabilidade entre sistemas e plataformas<sup>[3, 2, 10]</sup>.
- 5) Repositórios: locais específicos armazenam e organizam dados de forma sistemática e segura, permitindo que os autores os estruturam de forma a facilitar o acesso, compartilhamento e preservação permanente.



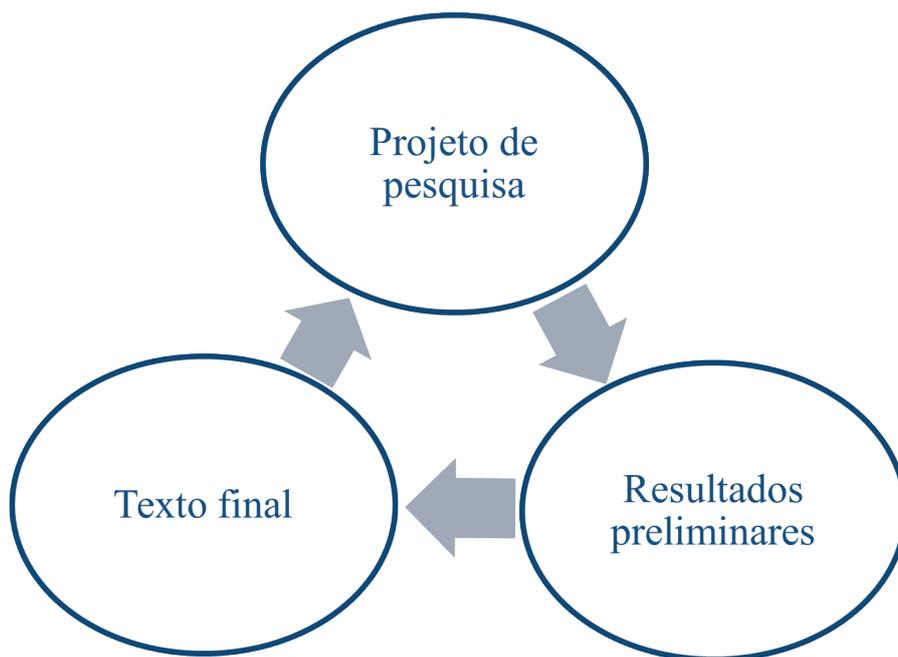
**Figura 1.** Etapas para desenvolvimento do Plano de Gestão de Dados

Fonte: Elaborado pelos autores com base em Van den Eynden<sup>[18]</sup> e Beugelsdijk et al.<sup>[18,2]</sup>

O modelo de PGD proposto neste estudo abrange cinco etapas essenciais para seu desenvolvimento, alinhadas às três fases do ciclo de desenvolvimento de um Trabalho de Conclusão de Curso (TCC). Embora voltado para estudantes de Data Science Analytics, ele pode ser aplicado em outras disciplinas que exijam um PGD. Considerando a complexidade das bases de dados utilizadas, é indispensável um sistema computacional robusto que assegure agilidade, eficiência e segurança. O uso de plataformas ou repositórios integrados para coleta, criação e compartilhamento de dados também é fundamental. As diretrizes do modelo, ilustradas na Figura 2, buscam orientar os pesquisadores a adotar práticas cuidadosas e responsáveis, promovendo a credibilidade dos estudos e contribuindo para o avanço do conhecimento em suas áreas.

#### Ciclo da pesquisa

- 1) Projeto de pesquisa: envolve as atividades iniciais para a concepção do projeto, garantindo que a integridade e a confiabilidade dos dados sejam definidas desde o início.
- 2) Resultados preliminares: representam um marco crucial, no qual é fundamental assegurar a aplicação da metodologia e dos materiais estabelecidos no projeto.
- 3) Texto final: embora seja a etapa final, essa atividade é crucial para revisar padrões e formatos, estimar o volume de dados gerado e especificar as ferramentas para seu processamento.



**Figura 2.** Ciclo da pesquisa

Fonte: Elaborado pelos autores.

Em cada uma das três etapas, foram incorporadas atividades essenciais para a gestão abrangente de dados. Essa integração ao ciclo de pesquisa garantiu a aplicação dos princípios CARE e FAIR, promovendo um modelo que alinha ações éticas e responsáveis. Assim, o pesquisador pode conduzir suas atividades de forma transparente, descrevendo detalhadamente a base de dados e comunicando com clareza a confiabilidade dos resultados do estudo. Este estudo apresentou um modelo para facilitar a implementação do PGD em pesquisas de GD e outros trabalhos acadêmicos, enfatizando a importância da integridade nas três fases do ciclo de pesquisa: concepção, desenvolvimento e redação final. Ao incorporar práticas éticas desde o início até a conclusão do projeto, reforça-se a conduta responsável dos pesquisadores e a necessidade de uma gestão rigorosa dos dados. Essa abordagem não só eleva a qualidade dos trabalhos acadêmicos, mas também promove o diálogo sobre transparência na disseminação e reutilização de dados.

# REFERÊNCIAS

- [1] Valkenburg, G.; Dix, G.; Tjeldink, J.; Rijcke, S. 2021. Expanding research integrity: A cultural-practice perspective. *Science and engineering ethics* 27(10):1-23. <https://doi.org/10.1007/s11948-021-00291-z>.
- [2] Beugelsdijk, S.; Van Witteloostuijn, A.; Meyer, K.E. 2020. A new approach to data access and research transparency (DART). *Journal of International Business Studies* 51:887-905. <https://doi.org/10.1057/s41267-020-00323-z>.
- [3] Doorn, P.K. 2018. *Science Europe Practical Guide to the International Alignment of Research Data Management*. <https://doi.org/10.5281/zenodo.4915862>.
- [4] McGovern, A.; Ebert-Uphoff, I.; Gagne, D.J.; Bostrom, A. 2022. Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science. *Environmental Data Science*, 1, p.e6. DOI: 10.1017/eds.2022.5.
- [5] Hosseini, M.; Resnik, D.B.; Holmes, K. 2023. The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics* 19(4):449-465. <https://doi.org/10.1177/17470161231180449>.
- [6] Alam, S.; Burley, R.; Graf, C.; Meadows, A.; Mejias, G.; Pattinson, D. 2024. (Re?) Building trust in research integrity. *Information Services & Use* 44(1):1-4. DOI: 10.3233/ISU-230200.
- [7] Cardoso, J.; Pereira, F.; Moreira, J.M. 2024. Serviço de Repositório Digital de Dados de Investigação. *Revista Científica Da UEM: Série Letras E Ciências Sociais*, 4(1). Disponível em: <http://196.3.97.23/revista/index.php/lcs/article/view/223>. Acesso em: 21 jan. 2025.
- [8] Gottardi, T.; Medeiros, C.B.; Reis, J.C., 2021. Semantic Search to Foster Scientific Findability: A Systematic Literature Review. *Journal of Information and Data Management* 12(5). <https://doi.org/10.5753/jidm.2021.1919>.
- [9] Naderifar, M.; Goli, H.; Ghaljaie, F. 2017. Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in development of medical education*, 14(3). <https://doi.org/10.5812/sdme.67670>.
- [10] Ministério da Saúde [MS]. 2023. Conselho Nacional de Saúde. Comitê de Ética em Pesquisa. Disponível em: < <https://conselho.saude.gov.br/comites-de-etica-em-pesquisa-conep?view=default> >. Acesso em: 11 de abril de 2024.
- [11] Engelhardt, C. 2022. D7. 4 How to be FAIR with your data. Disponível em: <https://library.oapen.org/handle/20.500.12657/54460>. Acesso em: 21 jan. 25.
- [12] Governo Digital. 2024. Governança de Dados. Disponível em: <https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/governancadedados/governanca-de-dados>. Acesso em: 11 de abril de 2024.
- [13] Bolland, M.J.; Avenell, A.; Grey, A. 2024. Publication integrity: what is it, why does it matter, how it is safeguarded and how could we do better?. *Journal of the Royal Society of New Zealand*, 55(1541):1-20. DOI: 10.1080/03036758.2024.2325004.
- [14] Bouter, L. 2020. What research institutions can do to foster research integrity. *Science and engineering ethics*, 26(4):2363-2369. DOI: 10.1007/s11948-020-00178-5.
- [15] Carcausto-Calla, W.; Zapata, N.A.; Cueva, F.E.I.; Morales, S.A.G.; Rios, A.R. 2023. Mapping of Empirical Studies on Research Integrity in University Institutions 12(5):9-22. <https://doi.org/10.36941/ajis-2023-0122>.
- [16] Van den Hoven, M.; Lindemann, T.; Zollitsch, L.; Prieß-Buchheit, J. 2023. A Taxonomy for Research Integrity Training: Design, Conduct, and Improvements in Research Integrity Courses. *Science and engineering ethics*, 29(14):1-21. <https://doi.org/10.1007/s11948-022-00425-x>.
- [17] Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Mons, B. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1):1-9. <https://doi.org/10.1038/sdata.2016.18>.
- [18] Van den Eynden, V.; Corti, L.; Woollard, M.; Bishop, L.; Horton, L. 2011. Managing and sharing data; a best practice guide for researchers. Disponível em: <https://dam.ukdataservice.ac.uk/media/622417/managingsharing.pdf>. Acesso em: 21 jan. 2025.
- [19] ORCID. (2024). ORCID - Connecting Research and Researchers. Disponível em: <https://orcid.org/>. Acesso em: 11 de abril de 2024.
- [20] Laender, A.H.F.; Medeiros, C.M.B.; Cendes, I.L.; Barreto, M.L.; Van Sluys, M.A.; Almeida, U.B.D. 2020. Abertura e gestão de dados: desafios para a ciência brasileira. Disponível em: <https://www.abc.org.br/wp-content/uploads/2020/09/ABC-Abertura-e-Gest%C3%A3o-de-Dados-desafios-para-a-ci%C3%Aancia-brasileira.pdf>. Acesso em: 21 jan. 2025.
- [21] Souza, L.P.; Sousa, R.S.C.; Löw, M.M.; Barros, T.H.B. 2023, August. Promovendo a justiça epistêmica: uma análise dos princípios CARE na gestão de dados de pesquisa em relação aos povos indígenas. In *Anais do Workshop de Informação, Dados e Tecnologia-WIDaT* 6. <https://doi.org/10.22477/vi.widat.73>.