

**ANA LÚCIA DA SILVA ARAÚJO**

O desafio da gestão de dados de pesquisa em instituições de ensino superior públicas brasileiras: diretrizes para o desenvolvimento de uma política institucional

Dissertação de mestrado  
Agosto de 2023



**Ana Lúcia da Silva Araújo**

**O desafio da gestão de dados de pesquisa em instituições de ensino superior públicas  
brasileiras:  
diretrizes para o desenvolvimento de uma política institucional**

Dissertação apresentada ao Instituto Brasileiro de Informação em Ciência e Tecnologia e ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

Orientadora: Dra. Luana Farias Sales

Rio de Janeiro  
2023

## CIP - Catalogação na Publicação

A66d Araújo , Ana Lúcia da Silva  
O desafio da gestão de dados de pesquisa em instituições de ensino superior públicas brasileiras: diretrizes para o desenvolvimento de uma política institucional / Ana Lúcia da Silva Araújo . -- Rio de Janeiro, 2023.  
158 f.

Orientadora: Luana Faria Sales.  
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Escola da Comunicação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, 2023.

1. Gestão de dados de pesquisa - Brasil . 2. Gestão de dados de pesquisa. 3. Políticas para gestão de dados de pesquisa . 4. Ciência Aberta - Dados de pesquisa. I. Sales, Luana Faria, orient. II. Título.


Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

ANA LÚCIA DA SILVA ARAÚJO

O DESAFIO DA GESTÃO DE DADOS DE PESQUISA EM INSTITUIÇÕES  
DE ENSINO SUPERIOR PÚBLICAS BRASILEIRAS:  
PROPOSTAS PARA A CONSTRUÇÃO DE UMA POLÍTICA  
INSTITUCIONAL


Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a da Universidade Federal do Rio de Janeiro (UFRJ), Escola de Comunicação (ECO), para obtenção do título de Mestre em Ciência da Informação.

Rio de Janeiro, 05 de setembro de 2023

Documento assinado digitalmente  
 LUANA FARIAS SALES MARQUES  
Data: 22/09/2023 18:47:58-0300  
Verifique em <https://validar.iti.gov.br>


---

Profa Dra Luana Farias Sales-PPGCI/IBICT – ECO/UFRJ  
Orientadora

Documento assinado digitalmente  
 SILVANA APARECIDA BORSETTI GREGORIO VIDC  
Data: 25/09/2023 16:04:49-0300  
Verifique em <https://validar.iti.gov.br>


---

Profa. Dra. Silvana Aparecida Borsetti Gregorio  
Faculdade de Filosofia e Ciências (FFC) – UNESP  
Avaliadora externa

Documento assinado digitalmente  
 MAIRA MURRIETA COSTA  
Data: 27/09/2023 12:40:24-0300  
Verifique em <https://validar.iti.gov.br>

---

Dra. Maira Murrieta Costa  
Ministério da Ciência, Tecnologia e Inovação (MCTI)  
Avaliadora externa

Documento assinado digitalmente  
 NAIRA CRISTOFOLETTI SILVEIRA  
Data: 28/09/2023 21:35:32-0300  
Verifique em <https://validar.iti.gov.br>

---

Profa. Dra. Naira Christofolletti Silveira-PPGCI/IBICT – ECO/UFRJ  
Avaliadora interna

Rio de Janeiro  
2023

## AGRADECIMENTOS

Aos meus pais, José da Silva Araújo e Percília José da Silva Araújo, que, vindos do coração do cerrado mineiro e em meio das adversidades, formaram técnicos, bacharéis, mestres e doutores, ensinando-nos a importância da honestidade, da educação, do trabalho e da superação individual e em família; À tia Geralda, que há 90 anos é um exemplo de força, coragem e determinação; Aos meus irmãos Waguinho, Águida, Gervásio, Liu, Cinésio, Nisa, Dada, Nilma e Rico, onde o conceito de irmandade resplandece em sua forma mais sublime; aos meus sobrinhos Camila, Tom, Carol, Ana Clara, Mateus, Gabriel, Duda, Júlia e Heitor, por compartilham alegrias e desafios, luz de nossa jornada; aos cunhados e cunhadas, que enriquecem os laços familiares. À Sônia, por dispensar sua atenção e generosidade a todos.

Agradeço à UFMG, *incipit vita nova*, marco de um novo começo, que não apenas possibilitou a aquisição de novos conhecimentos e perspectivas, mas, sobretudo, propiciou o meu encontro com pessoas admiráveis. Destaco aqui Gustavo Saldanha, uma notável figura como professor e humanista; meus sinceros agradecimentos. Sua generosidade e suporte foram essenciais para o início deste trabalho. À Shirley, Fernanda de Paula e Viviane Solano, por serem um porto seguro, oferecendo apoio em diferentes momentos, sem julgamentos. Expresso também minha gratidão aos demais amigos - Ronaldo, Aline Coelho, Paulo, Analice, Carla Angelo e Fernanda Almeida - por torcerem por mim à distância. Agradeço especialmente à minha amiga-irmã Leila Anastácio, a figura central em todo este percurso, que acreditou em mim quando me faltou a fé. Sem ela, esta dissertação jamais teria visto a luz do dia.

À UFSJ, agradeço por ter disponibilizado recursos que tornaram minha jornada na pós-graduação mais acessível. Agradeço às minhas brilhantes colegas de trabalho, Cláudia e Cristiane, qual dedicação à biblioteca, juntamente com o apoio e a compreensão ao longo deste período, foram fundamentais para que eu pudesse me dedicar aos estudos.

Manifesto minha profunda gratidão aos professores do PPGCI IBICT UFRJ, bem como a todos os colaboradores, por viabilizarem o ensino e a pesquisa a distância, tornando possível uma educação sem barreiras. À minha professora e orientadora, Luana Sales, expressei minha sincera gratidão por me apresentar novas perspectivas da pesquisa pelo prisma da Ciência da Informação e por guiar-me até a conclusão desta dissertação.

## RESUMO

Este estudo abrangeu a evolução global do compartilhamento de dados de pesquisa, emergindo a necessidade de gerenciar esses dados de maneira organizada e responsável. Com a convicção de que uma política institucional para a gestão de dados de pesquisa, é essencial para estabelecer práticas obrigatórias e padronizadas, esta pesquisa buscou os elementos e as abordagens cruciais para viabilizar tais serviços. A apresentação dos resultados foi embasada em uma investigação na literatura da área, em documentos legais e institucionais, diretrizes e padrões de organizações especializadas nacionais e internacionais, realizada por meio de uma metodologia de pesquisa qualitativa, exploratória e aplicada. Com o intuito de compreender a abrangência e as lacunas no cenário nacional em relação ao tema, foram coletadas informações de instituições de ensino superior brasileiras, bem como de outras instituições com as quais se relacionam, como institutos de pesquisa e agências de fomento. Para atingir o principal objetivo da pesquisa, foram analisadas políticas de gestão de dados de pesquisa em países que já têm políticas instituídas, nas quais instituições de ensino e pesquisa serviram como modelos de referência. Dentre essas instituições, as universidades da Holanda, reconhecidas por seu comprometimento com os princípios da Ciência Aberta, foram escolhidas para extrair elementos e abordagens de seus documentos. Por meio de análises individuais e comparativas, foi elaborada uma planilha para identificar os elementos comuns e singulares dessas políticas. A partir desse levantamento, foi construído um quadro com diretrizes e abordagens possíveis que podem ser aplicados na construção de políticas para as instituições de ensino e pesquisa no Brasil. Foi constatado que há diferenças significativas entre os dois países em relação a apoio de órgãos públicos, infraestrutura, regulamentações de nível nacional e serviços federados para viabilização e apoio aos serviços. Concluiu-se que a elaboração de políticas institucionais e gestão de dados nas instituições de ensino superior do Brasil carecem de apoio dos órgãos públicos, do conhecimento e reconhecimento do valor dos dados de pesquisa de suas lideranças institucionais.

Palavras-chave: Gestão de dados de pesquisa. Dados de pesquisa. Políticas para gestão de dados de pesquisa – Brasil. Ciência Aberta – dados de pesquisa.

## ABSTRACT

This study covered the global evolution of data sharing in research, emerging the need to manage these data in an organized and responsible way. With the conviction that an institutional policy for research data management is essential to establish mandatory and standardized practices, this research sought the crucial elements and approaches to enable such services. The presentation of the results was based on an investigation of the literature in the area, legal and institutional documents, guidelines and standards of national and international specialized organizations, conducted through a qualitative, exploratory and applied research methodology. In order to understand the scope and gaps in the national scenario in relation to the topic, information was collected from Brazilian universities, as well as from other institutions with which they are related, such as research institutes and funding agencies. To achieve the main objective of this research, research data management policies in countries that already have established policies were analyzed, in which education and research institutions served as reference models. Among these, the universities of the Netherlands, recognized for their commitment to the principles of Open Science, were chosen to extract elements and approaches from their documents. Through individual and comparative analyzes, a spreadsheet was developed to identify the common and singular elements of these policies. From this, a framework was constructed with possible guidelines and approaches that can be applied in the construction of policies for teaching and research institutions in Brazil. It was found that there are significant differences between the two countries in terms of support from government institutions, infrastructure, national-level regulations and federated services to enable and support services. It was concluded that the development of institutional policies and data management in Brazilian universities lack support from government institutions, knowledge and recognition of the value of research data from their institutional leaders.

Key-words: Research data management. Research data. Research data management policies – Brazil. Open science – research data.

## LISTA DE FIGURAS

Figura 1 – Guarda-chuva da Ciência Aberta.....	32
Figura 2 – Taxonomia da Ciência Aberta.....	33
Figura 3 – Os dados de pesquisa em evidência.....	34
Figura 4 – Novas práticas: foco nos dados.....	35
Figura 5 – Reflexões sobre a gestão de dados científicos.....	36
Figura 6 – Dados de pesquisa: promessas da RNP e do IBICT.....	41
Figura 7 – Etapas para a construção de um PDA.....	49
Figura 8 – Ações para abertura dos no Brasil.....	52
Figura 9 – A legislação e os dados em nuvem.....	72
Figura 10 – Iniciativas de IES nacionais para a promoção da GDP.....	91
Figura 11 – As universidades holandesas selecionadas.....	95
Figura 12 – Aspirações do Open Science 2030 in the Netherlands.....	98
Figura 13 – Sistema Argos, OpenAIRE.....	99
Figura 14 – Ecossistema do Argos no OpenAIRE.....	100



## LISTA DE QUADROS E GRÁFICOS

Quadro 1 – Estrutura para análise de conceitos.....	60
Quadro 2 – Modelo da planilha para organizar os elementos.....	67
Quadro 3 – Correlação: caminhos para os resultados.....	69
Quadro 4 – Embasamento legal para a GDP no Brasil.....	71
Quadro 5 – Alguns conceitos e as fases na gestão.....	75
Quadro 6 – Conceituando alguns tipos de dados.....	78
Quadro 7 – Classificação: impactos e riscos.....	79
Quadro 8 – Tipos de acessos: relacionando conceitos.....	80
Quadro 9 – A busca por normativas no território nacional.....	81
Quadro 10 – Normativas em institutos de pesquisa nacionais.....	82
Gráfico 1 – O panorama nos institutos de pesquisa nacionais.....	83
Gráfico 2 – O panorama nas agências de fomento nacionais.....	83
Quadro 11 – Padrão FAPESP para planos de gestão de dados.....	84
Quadro 12 – As IES nacionais: uma delimitação do campo.....	86
Gráfico 3 – O panorama nas IES nacionais.....	88
Quadro 13 – Elementos de políticas para GDP em IES nacionais.....	89
Quadro 14 – Os serviços de gestão nas IES nacionais.....	90
Quadro 15 – O suporte para a Ciência Aberta, na Holanda.....	96
Quadro 16 – As políticas das IES holandesas.....	101
Quadro 17 – Os elementos para GDP na Holanda: ocorrências e incidências.....	102
Quadro 18 – Políticas para GDP em IES brasileiras: diretrizes.....	108

## LISTA DE ABREVIATURAS E SIGLAS

ALLEA – Todas as Academias Europeias

BDTD – Biblioteca Digital Brasileira de Teses e Dissertações

BNDES – Banco Nacional de Desenvolvimento Econômico e Social

BOAI – *Budapest Open Access Initiative*

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CCBS – Código de Conduta para Ciências Sociais e Comportamentais

CGU – Controladoria Geral da União

CF – Constituição Federal do Brasil de 1988

CI – Ciência da Informação

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CODATA – Comitê de Dados para Ciência e Tecnologia

CWTS – *Centre for Science and Technology Studies*

DCC – *Digital Curation Centre*

DMPonline – *Data Management Plan Online*

DMPTool – *Data Management Plan Tool*

DOAJ – *Directory of Open Access Journal*

DSpace – Digital Space

EC – *European Commission*

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária

EOSC – *European Open Science Cloud*

ESRC – *Economic and Social Research Council*

FAIR – *Findable, Accessible, Interoperable, Reusable*

FAPEMIG – Fundação de Amparo à Pesquisa do Estado de Minas Gerais

FAPERGS – Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul

FAPERJ – Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo

FEST – Fundação Espírito-Santense de Tecnologia

FINEP – Financiadora de Estudos e Projetos

FIOCRUZ – Fundação Oswaldo Cruz

FioDMP – Plano de Gestão de Dados da Fundação Oswaldo Cruz [modelo FAIR]

FURG – Universidade Federal do Rio Grande

GDP – Gestão de Dados de Pesquisa

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia

ICTs – Instituições Científica, Tecnológica e de Inovação

IES – Instituições de Ensino Superior

IGC – Índice Geral de Cursos

IGI - Índice Global de Inovação

IMPA – Instituto de Matemática Pura e Aplicada

INDA – Infraestrutura Nacional de Dados Abertos

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

INPA – Instituto Nacional de Pesquisas da Amazônia

INPE – Instituto Nacional de Pesquisas Espaciais

INPEC - Instituto Nacional de Pesquisa Educação e Cultura

IFDS - Internet of FAIR Data & Services

IPJB – Instituto de Pesquisas Jardim Botânico

IPT – Instituto de Pesquisas Tecnológicas

ITA – Instituto Tecnológico de Aeronáutica

JBRJ – Jardim Botânico do Rio de Janeiro

KNAW – *Royal Netherlands Academy of Arts and Sciences*

LAI – Lei de Acesso à Informação

LCRDM – *National Coordination Point Research Data Management*

LGPD – Lei Geral de Proteção de Dados

LIMPE – Legalidade, Impessoalidade, Moralidade, Publicidade e Eficiência

LISTA – *Library, Information Science and Technology*

NARCIS – *National Academic Research and Collaborations Information*

NCCRI – Código de Conduta para Integridade em Pesquisa da Holanda

NIH – *National Institute of Health*

NPOS – *National Programme Open Science*

NWO – *Netherlands Organization for Scientific Research*

OAI – *Open Archives Initiative*

OASISBR – *The Brazilian Open Access Publications and Scientific Data Portal*

OECD – *Organization for Economic Co-operation and Development*

OKF – *Open Knowledge Foundation*

ONU – Organização das Nações Unidas;

PDA – Plano de Dados Abertos

PGD – Plano de Gestão de Dados

PPBio – Programa de Pesquisa em Biodiversidade no Brasil

PURE – Portal de Resultados de Pesquisa da Holanda

RD – Repositório de Dados

RDA – *Research Data Alliance*

RDP Brasil – Trabalho da Rede de Dados de Pesquisa Brasileira

RI – Repositório Institucional

RNP – Rede Nacional de Pesquisa

SCIELO – *Scientific Electronic Library Online*

SGD – Secretaria de Governo Digital

SEER – Sistema Eletrônico de Editoração de Revistas

SOC – Sistema de Organização do Conhecimento

TEDE – Publicação Eletrônica de Teses e Dissertações

TI - Tecnologia de Informação

TIC – Tecnologia de Informação e Comunicação

UKB – Bibliotecas das Universidades Holandesas e Biblioteca Real da Holanda

VSNU – Associação das Universidades da Holanda

WOS – *Web of Science*

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>16</b>
1.1 OBJETIVO GERAL E OBJETIVOS ESPECÍFICOS.....	20
1.2 JUSTIFICATIVA.....	21
1.3 POTENCIAIS IMPACTOS DA PESQUISA.....	22
<b>2 REVISÃO DE LITERATURA.....</b>	<b>27</b>
2.1 A ERA DOS DADOS: UM MOVIMENTO MUNDIAL.....	27
2.2 OS DADOS DE PESQUISA: NOVAS PERSPECTIVAS.....	29
2.3 O ACESSO ABERTO, CIÊNCIA ABERTO E O BRASIL.....	37
2.4 A GDP NO BRASIL: OS PRIMEIROS PASSOS.....	39
2.5 DOCUMENTOS LEGAIS PARA ORIENTAR A GESTÃO DE DADOS DE PESQUISA 48	
<b>3 METODOLOGIA.....</b>	<b>56</b>
3.1 ETAPAS E DESCRIÇÃO DA PESQUISA.....	57
<b>3.1.1 Mapeamento conceitual da área de Gestão de Dados de Pesquisa.....</b>	<b>59</b>
3.2 PREVISÕES LEGAIS NACIONAIS PARA A GESTÃO DE DADOS.....	61
3.3 PANORAMA DA GESTÃO DE DADOS DE PESQUISA NO BRASIL.....	61
<b>1. 3.3.1 Agências de fomento/Financiadores.....</b>	<b>62</b>
<b>2. 3.3.2 Institutos de pesquisa nacionais.....</b>	<b>63</b>
<b>3. 3.3.3 Instituições de Ensino Superior Nacionais.....</b>	<b>63</b>
3.4 BUSCANDO ELEMENTOS EM IES ESTRANGEIRAS.....	64
<b>4. 3.4.1 Critérios para selecionar as IES holandesas a serem analisadas.....</b>	<b>66</b>
3.5 ORGANIZAÇÃO DAS INFORMAÇÕES COLETADAS NAS INSTITUIÇÕES..... CONSULTADAS.....	67 67
3.6 APRESENTAÇÃO DAS DIRETRIZES PARA UMA POLÍTICA PARA GDP.....	68
<b>4 ANÁLISE DOS RESULTADOS.....</b>	<b>71</b>
4.1 HORIZONTES PARA O COMPARTILHAMENTO DE DADOS NO BRASIL.....	71
4.2 EXPLORANDO OS CONCEITOS.....	74
4.3 A GESTÃO DE DADOS DE PESQUISA: BUSCANDO ABORDAGENS NACIONAIS.. 82	
<b>5. 4.3.1 Iniciativas das agências de fomento e apoio à pesquisa.....</b>	<b>84</b>
<b>6. 4.3.2 Os primeiros acordos das IES nacionais.....</b>	<b>86</b>
4.4 ALÉM DAS FRONTEIRAS: AS UNIVERSIDADES HOLANDEASAS COMO..... REFERÊNCIA.....	94 94
4.5 AS POLÍTICAS DE GESTÃO DE DADOS DE PESQUISA NAS IES DA HOLANDA... 102	
<b>5 PROPOSTAS PARA A CONSTRUÇÃO DE UMA POLÍTICA INSTITUCIONAL. 106</b>	

5.1 ALGUMAS CONSIDERAÇÕES SOBRE A CIÊNCIA DA INFORMAÇÃO E OS.....	118
DADOS DE PESQUISA.....	118
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>120</b>
REFERÊNCIAS.....	123
<b>APÊNDICE A – Plano de Gestão de Dados, modelo FioDMP.....</b>	<b>143</b>
<b>APÊNDICE B – Plano de Gestão de Dados, modelo DMPTool (sem atualização).....</b>	<b>146</b>
<b>APÊNDICE C – A reunião dos conceitos.....</b>	<b>147</b>

## 1 INTRODUÇÃO

No âmbito deste Programa de Pós-Graduação, na especialidade da linha de pesquisa “Comunicação, Organização e Gestão da Informação e do Conhecimento”, há uma oportunidade significativa de contribuir para a formulação de políticas de Gestão de Dados de Pesquisa (GDP) em Instituições de Ensino Superior (IES) públicas no Brasil. A excelência representada pelo IBICT e UFRJ na promoção de pesquisas e soluções para aprimorar o ensino, pesquisa e extensão no país se alinha aos objetivos de desenvolver diretrizes consistentes para a GDP. O conhecimento adquirido nesse campo, incluindo domínio em gestão da informação, estratégias de documentação e considerações éticas relacionadas à pesquisa, oferece uma base sólida para promover a gestão ética e eficiente dos dados de pesquisa em um ambiente acadêmico em constante evolução. Essa interseção de conhecimento desempenha um papel vital na abordagem dos desafios cada vez mais complexos associados à GDP nas instituições acadêmicas.

O movimento do Acesso Aberto aos resultados de pesquisa ganhou impulso nos anos 1990 e se consolidou no início dos anos 2000, com a Declaração Internacional sobre o Acesso Aberto, publicado pelo *Budapest Open Access Initiative* (BOAI) (2002) e a Declaração de Berlim sobre o Acesso Aberto ao Conhecimento nas Ciências e Humanidades (2003). Na comemoração dos vinte anos da declaração de Budapeste, o BOAI (2022) enalteceu as conquistas quanto ao desenvolvimento de novas infraestruturas, ferramentas e serviços, novos métodos de avaliação e revisão por pares e, também, a elaboração de políticas de Acesso Abertos implantados por financiadores e universidades para promover e padronizar as novas práticas. Na primeira declaração de Budapeste, percebe-se uma ênfase aos resultados de pesquisa publicados em periódicos. Na manifestação de 2022, foi destacada a necessidade de incluir a Gestão e o Compartilhamento dos Dados de Pesquisa (GDP), metadados, códigos, protocolos, licenças, padrões e outros elementos que estão sob o guarda-chuva da Ciência Aberta.

Com o avanço das Tecnologias de Informação e Comunicação (TIC), os pesquisadores e as organizações começaram a explorar melhores práticas para compartilhar resultados de pesquisa de forma mais ampla e ágil. Com recursos computacionais e internet cada vez mais eficientes, surgiram novas formas de fazer ciência e de disponibilizar os resultados. Assim,



ormo e-science<sup>1</sup> emergiu como um novo paradigma, definindo com maior precisão a pesquisa realizada de forma colaborativa, instantânea e sem fronteiras, bem como o uso intensivo de dados gerados/coletados durante todo o processo da investigação.

No contexto desse novo paradigma, as Instituições de Ensino Superior (IES) se adaptaram aos princípios do Acesso Aberto e da *e-science* e passaram a adotar ferramentas e serviços para os *stakeholders*<sup>2</sup>. No Brasil, o Acesso Aberto foi impulsionado no início dos anos 2000, com o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) que customizou e disponibilizou Sistemas de Organização da Informação e do Conhecimento (SOCs), como exemplos: a Publicação Eletrônica de Teses e Dissertações (TEDE), a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o *The Brazilian Open Access Publications and Scientific Data Portal* (Oasisbr) e o Sistema Eletrônico de Editoração de Revistas (SEER), e Repositórios Institucionais (RDs), por meio do *Digital Space* (DSpace).

Durante a implantação dos Repositórios Institucionais (RI), as organizações enfrentaram um grande desafio para substituir o arquivamento e o acesso de documentos físicos por documentos em formato digital. Isso exigiu planejamento, análise de custos, atribuição de responsabilidades, treinamento de equipes e conscientização dos *stakeholders* da instituição sobre o Acesso Aberto por meio de TICs modernas. Na ocasião, foi necessário o estabelecimento de políticas institucionais para padronizar e tornar obrigatórias práticas consideradas atualmente como rotineiras.

Em cada período histórico, no processo de evolução dos modos de fazer ciência e formar cientistas, houve o desenvolvimento e a adaptação a novos métodos, tecnologias e paradigmas. A *e-science*, como novo paradigma, viabilizou investigações científicas por meio de recursos computacionais eficientes, gerando um número massivo de pesquisas e, conseqüentemente, uma quantidade ainda maior de dados. O termo “*e-Science*” passou a ser empregado para designar as pesquisas feitas em larga escala, por meio de recursos computacionais e colaborativos, sendo apontada por Jim Gray (2007), como o quarto paradigma da ciência.

Com as tecnologias e recursos cada vez mais avançados para atender à necessidade da verificação de integridade, confiabilidade e sustentabilidade da pesquisa, eis que uma nova

---

<sup>1</sup> Termo criado em 1999, por John Taylor, Diretor-Geral do *Joint Information Systems Committee* para denominar o uso de TICs que viabilizou a pesquisa colaborativa, geração/coleta de dados de forma intensiva.

<sup>2</sup> *Stakeholder* é uma pessoa, grupo ou organização que tem um interesse ou participação em um negócio, projeto ou organização. São stakeholders internos aqueles que fazem parte da mesma organização, ou externos. Quando se trata de interações entre as partes, usa-se apenas a expressão *stakeholder*.

demanda entrou em evidência: os **dados gerados/coletados** no processo da investigação científica. Os olhares de agências de financiamento, públicas e privadas, editoras de periódicos e outras organizações, penderam-se para o Acesso Aberto também dos dados de pesquisa. Com o reconhecimento do seu valor e importância, passou-se a exigir um Plano de Gestão de Dados (PGD), com justificativas e objetivos conforme a natureza da organização solicitante.

Na definição de Borgman e Bourne (2021), um PGD é um documento que descreve o gerenciamento, a preservação e o compartilhamento de dados científicos e dos metadados que os acompanham, onde deve-se informar como os dados serão gerados/coletados; onde serão armazenados e arquivados; e como serão compartilhados e reusados. A GDP chegou nas agendas de instituições que realizam pesquisas, nas que financiam e também nas que publicam resultados de pesquisa. As IES, como componentes fundamentais desse ecossistema, estão ligadas a essas organizações e dependem significativamente delas para apoiar e promover suas atividades acadêmicas e científicas. A integração das IES ao movimento da GDP tem ocorrido de forma gradual e progressiva, mas sua adesão, tanto em âmbito nacional quanto internacional, ainda não é universal. Isso se deve ao fato de que as IES se encontram em estágios distintos de desenvolvimento, possuem diferentes níveis de disponibilidade orçamentária e apresentam variações na conscientização e compreensão do tema.

É importante ressaltar que em países onde o poder público oferece infraestrutura e apoio financeiro, a GDP vem sendo adotada com maior adesão. No Brasil, por exemplo, foi identificado que a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), em 2017, começou a solicitar o PGD aos pesquisadores. Não por acaso, as IES do estado de São Paulo estão à frente na disponibilização de recursos para a promoção da GDP. Alguns “focos” foram encontrados em outros estados, mas com abrangência menor.

Para os pesquisadores e as instituições de pesquisa, as responsabilidades vão além da elaboração do PGD. Para as IES, especificamente, é preciso o envolvimento de lideranças institucionais e uma gama de categorias profissionais para efetivar o gerenciamento dos dados. Esta missão, como na ocasião da implantação do RI, nos primeiros anos do século XXI, somente será bem-sucedida com a construção de ferramentas e políticas claras e objetivas, que abordem processos sistemáticos para a GDP, formalizando e monitorando as práticas das IES. Entretanto, ainda não estão claros quais elementos, procedimentos e

responsabilidades devem constar no documento em questão.

**Problematização:** Para atender às demandas relacionadas aos dados de pesquisa, as IES precisam de investimentos, capacitação de profissionais e de documentos regulatórios para a promoção da governança, tais como legislação, códigos de conduta e declarações de integridade da pesquisa. Esses documentos podem ser de alcance institucional, nacional e regional, como é o caso da União Europeia (UE). A menção da obrigatoriedade do PGD em diferentes documentos fortalece a ideia do compartilhamento e reuso de dados cada vez mais presente na prática acadêmica. A UNESCO, a UE e o Banco Mundial estão se manifestando a favor do gerenciamento e compartilhamento de dados, reconhecendo sua importância para a transparência, a colaboração, o acesso e o reuso.

Independentemente de pressões externas, as IES públicas do Brasil precisam reconhecer a importância de planejar, organizar e compartilhar os dados de pesquisa, e seguir procedimentos formalizados para garantir a integridade, a preservação e o acesso responsável. É fundamental que as IES viabilizem esse novo serviço, não uma opção, mas como uma obrigação, atuando proativamente, independentemente de estímulos externos.

Nas agências de fomento que exigem o PGD, não foram encontradas informações sobre os procedimentos ou abordagens adotadas para monitorar e assegurar a efetiva gestão de dados durante e após a realização de pesquisas. A Fapesp financia cerca de 10 mil projetos de pesquisa por ano e, em 2017, passou a exigir a apresentação de um PGD para conceder benefícios. Se os números indicados no Metabusador de dados de pesquisa da Fapesp estiverem atualizados, entre 2020 e 2023 foram arquivados 458 conjuntos de dados. Isso sugere que há uma lacuna entre o volume de dados gerados/coletados pelos milhares de projetos que financia e o apresentado no Metabusador. Ressalta-se que a GDP não deve ser realizada apenas para atender às demandas de *stakeholders* externos, mas também por uma questão moral, considerando o compromisso assumido em gerenciar os dados no momento da aceitação dos termos, pelo pesquisador.

Não é possível compartilhar os dados sem que sejam gerenciados. Como um recurso de informação valioso, deve ser gerenciado conforme suas especificidades, garantindo a preservação e acessos futuros. Os processos para gestão que viabilizam o compartilhamento e reuso devem estar em conformidade com as leis e princípios éticos, e assegurar o alinhamento com os interesses de todas as partes envolvidas, incluindo pesquisadores, IES, financiadores e demais *stakeholders*. Torna-se imperativo, então, reunir um conjunto de informações para

formalizar as práticas criteriosamente, e elaborar uma política de gestão que viabilize o compartilhamento e reuso dos dados de pesquisa de forma ampla, mas também responsável.

Parte-se do pressuposto de que, dada a diversidade e as peculiaridades da rede de elementos necessários para a GDP, a ausência de documentação formal adequada pode resultar em erros, inconsistências e omissões no processo. Compreender as demandas existentes é fundamental para a formulação de uma política institucional com procedimentos que incluam elementos e abordagens tanto padronizados quanto flexíveis, além de acordos e medidas de contingência.

Embora algumas IES nacionais já estejam promovendo a GDP, seus termos não abrangem alguns elementos para se alinharem aos padrões internacionais. Assim, surge a seguinte **questão de pesquisa**: Quais são os elementos necessários para compor uma política para gestão de dados de pesquisa nas instituições de ensino superior públicas brasileiras?

## 1.1 OBJETIVO GERAL E OBJETIVOS ESPECÍFICOS

Esta pesquisa tem como **objetivo geral** investigar quais são os elementos necessários para compor uma política para a gestão de dados de pesquisa nas instituições de ensino superior públicas brasileiras.

No decorrer da pesquisa, apresentam-se os seguintes **objetivos específicos**:

- conhecer os conceitos utilizados na área para compreender o processo da GDP;
- identificar as instituições nacionais que abordam a gestão de dados de pesquisa para conhecer o panorama nacional;
- analisar os elementos, abordagens e procedimentos presentes nas políticas de gestão de dados de pesquisa em instituições de ensino superior estrangeiras, para extrair seus elementos e utilizá-los como referência;
- propor elementos, abordagens e procedimentos para a formulação de uma política de gestão de dados de pesquisa em instituições de ensino superior públicas no Brasil.

## 1.2 JUSTIFICATIVA

A gestão e o compartilhamento dos dados de pesquisa pode ser a primeira oportunidade em que os autores e as IES possam ter domínio mais amplo sobre o nível de acesso e reuso de ativos produzidos em suas pesquisas, considerando o controle e a rastreabilidade possíveis, se forem atribuídas aos dados, as ferramentas pertinentes. Em alguns artigos, livros e outras produções científicas, o conteúdo passava (e ainda passa) a ser propriedade da editora, que detinha (e detém) os direitos e ganhos financeiros com os acessos dos documentos. Já as teses e dissertações foram disponibilizadas em Acesso Aberto, exceto nos casos de embargos temporários e nas que contêm informações confidenciais. No que diz respeito aos dados de pesquisa, os tipos de acesso são semelhantes, mas, no entanto, são acompanhados de metadados adicionais que registram os acessos e rastreiam as atividades relacionadas.

A relevância desta pesquisa reside na necessidade inevitável de gerenciamento e compartilhamento dos dados de pesquisa nas IES brasileiras, mesmo que de forma tardia. A pressão virá não apenas de agências de fomento, editoras e outras organizações, mas também de instituições de ensino e pesquisa estrangeiras. Além disso, as IES devem entrar em conformidade com diversas disposições legais e, principalmente, com o Decreto nº 8.777, que institui a Política de Dados Abertos do Poder Executivo Federal (BRASIL, 2016), e abrange o gerenciamento e a abertura de dados, incluindo os de pesquisa. Portanto, as IES, por estarem subordinadas ao Poder Executivo Federal, devem obedecer diretamente às disposições deste decreto.

O principal objetivo da gestão de dados é viabilizar o reuso. Para alcançar esse objetivo, deve-se elaborar o PGD que apresente os critérios de acesso e/ou embargo e/ou restrição, observando as leis, a ética e integridade que já preveem normativas quanto ao sigilo, a proteção e o descarte. Em algumas instituições o pesquisador tem liberdade para definir, conforme interesses baseados em critérios e justificativas, a abertura ou restrição.

A infraestrutura e a tecnologia atuais representam desafios significativos para a gestão de dados nas IES públicas no Brasil, considerando a escassez de recursos financeiros que dificultam a implementação de atividades e procedimentos necessários para esse serviço. Além disso, é comum encontrar resistências e desconfianças por parte dos investigadores em relação à abertura dos dados. Para o docente, conciliar as demandas da gestão de dados de

pesquisa com as práticas de ensino e orientação também será um grande desafio; para essas questões, programas de conscientização, workshops e outros eventos serão necessários para o envolvimento dos acadêmicos para demonstrar os benefícios que o compartilhamento de dados pode trazer, quando há uma gestão eficiente.

Em 2014, um grupo multidisciplinar de pesquisadores se reuniu em Leiden, na Holanda, e identificou que, para uma gestão eficaz, os dados de pesquisa devem atender aos princípios FAIR: *Findable* (Encontráveis), *Accessible* (Acessíveis), *Interoperable* (Interoperáveis) e *Reusable* (Reutilizáveis). Desde então, essa sigla se tornou uma bandeira na gestão de dados, e cada um desses princípios se desdobram em outras orientações específicas.

Para que o Brasil possa se alinhar aos propósitos do FAIR e se juntar ao movimento global, é essencial estabelecer normativas institucionais por meio de critérios claros e abordar aspectos como padronização de metadados, seleção de repositórios apropriados, adoção de licenças e definição de políticas de compartilhamento. Ao seguir essas normativas e implementar os princípios FAIR, o Brasil poderá avançar na gestão de dados de pesquisa, fomentando a colaboração e o progresso científico, além de promover o reuso de dados de forma ampla e impactante.

### 1.3 POTENCIAIS IMPACTOS DA PESQUISA

Para serem compartilhados, os dados devem ser gerenciados. Para que sejam gerenciados corretamente, deve haver uma política institucional. Para que a política seja eficiente e eficaz, deve-se aplicar estratégias para definir as diretrizes que normatizem as práticas na casa.

Esta pesquisa tem o potencial de oferecer benefícios significativos para a ciência, a sociedade e o meio ambiente. O Acesso Aberto aos dados devidamente gerenciados promove a colaboração entre pesquisadores, instituições e órgãos governamentais, tornando a ciência mais acessível, confiável e contribuindo para que os processos sejam mais eficientes; além disso, contribui na luta contra o negacionismo e a descrença nos resultados da pesquisa. O Brasil sofre preconceitos por ser um país latino-americano, mas tem uma forte tradição de pesquisa. A GDP pode colaborar para reduzir esse preconceito por meio de uma política que

promova a divulgação das pesquisas brasileiras e o intercâmbio científico entre o Brasil e outros países.

Embora o foco desta subseção seja destacar o potencial impacto da pesquisa para o cenário científico, social e ambiental, considerou-se importante apresentar a motivação pessoal da pesquisadora, considerando seu anseio em explorar esse domínio no contexto nacional, para uma contribuição em prol da sociedade brasileira.

- **Importância para a pesquisadora:** *“Em meio às páginas desta trajetória, como bibliotecária, testemunhei a metamorfose infinitesimal no tempo, em que os antigos formatos físicos cederam lugar à explosão cósmica digital, e a fusão é sentida dentro de nosso universo. Nesta expansão, vi-me desafiada a adaptar-me como espécie em uma seleção natural para que o sistema desta profissão mantivesse seu significado na sociedade. Nesta jornada no tempo espaço, cuidou-se para preservar a filosofia da ciência em prateleiras forjadas em metais pesados e em transição para fibras ópticas e ondas eletromagnéticas. Uma das missões dos Cientistas da Informação é, também, ser a ponte que une aqueles que criam histórias, que fazem perguntas e formulam hipóteses, aos amantes que voam em palavras, buscadores de respostas que abrem caminhos para as galáxias. Os pulsares desta supernova (os dados de pesquisa), como guardiões deste universo em expansão, garantir a democratização do acesso aberto é um princípio, proteger com restrições é uma exceção. Os bibliotecários, com mãos cobertas de poeira física e digital, guardamos a memória e derrubamos fronteiras, erguendo pilares para o futuro”.*
- **Importância para a ciência:** **viabilizar** o reuso dos dados de pesquisa; **contribuir** para conformidade as perspectivas da *e-science*, tornando o acesso aos dados de forma mais colaborativa, eficiente e acessível nacionalmente e internacionalmente; **formalizar e padronizar** uso de ferramentas e procedimentos para a GDP em nível institucional; incentivar a **interação entre os pesquisadores** nacionais e estrangeiros, de áreas afins e transdisciplinares; promover a **sustentabilidade ambiental**, reduzindo os impactos associados a pesquisas que utilizam recursos naturais, substâncias corrosivas, radioativas, poluentes, cobaias e outras, evitando a realização de

nova geração/coleta de dados; garantir a **transparência** e **integridade**, permitindo que outros pesquisadores e qualquer indivíduo possa verificar as práticas metodológicas e o processamento dos dados, contribuindo assim para a validação e avanço do conhecimento científico; reduzir os **custos da pesquisa**, uma vez que a reuso evita o retrabalho na geração/coleta dos dados, bem como nos custos associados à própria gestão de dados duplicados, resultando em economia financeira que pode ser direcionada a outros fins; **agilizar a pesquisa**, pois acelera o progresso da investigação, uma vez que elimina a necessidade de gastar tempo na geração/coleta de dados, permitindo que os estudos sejam concluídos de forma mais rápida; **reduzir** a necessidade de **dispêndio e/ou o desgaste** de materiais de escritório, equipamentos de segurança, maquinários, instrumentos de laboratórios, reagentes, consumos de energia elétrica, água, dentre outros recursos.

- **Importância para a sociedade:** **incluir** a sociedade civil como parte interessada para que possam visualizar e/ou utilizar os dados para **desenvolver soluções** para necessidades específicas (construção civil, agricultura, por exemplo); oportunizar **aprendizado e capacitação** de pessoas e grupos na disseminação do conhecimento; combater a **desigualdade e injustiça**, permitindo uma maior compreensão da consciência de classes sociais, história, raças, gêneros, dentre outras extensões, promovendo uma sociedade mais equitativa e justa; promover o **acesso aos dados** de pesquisa **financiados**, direta ou indiretamente, pela sociedade, como forma de garantir o direito de acesso à informação; fortalecer a confiança na ciência, combatendo as *fake news*, uma vez que os dados podem ser verificados também pela sociedade.
- **Importância para as IES:** aumentar a produção científica estimulando e facilitando a produção de pesquisa de alta qualidade e **eleva a reputação e a confiabilidade** da instituição, contribuir para a expansão do conhecimento em diversas áreas; **aumentar a visibilidade** da IES no cenário acadêmico nacional e internacional por meio de sua contribuição na abertura dos dados gerados/coletados; facilitar a **extensão universitária**,



pois os dados podem ser compartilhados e aplicados em ações sociais e programas de extensão, promovendo assim o envolvimento da instituição com a comunidade; promover a **cultura dos dados de pesquisa abertos**, assimilando seus princípios na graduação e pós-graduação, incentivando as práticas de transparência, compartilhamento e colaboração; estímulo ao **investimento empresarial**, pois os dados são ativos valiosos que podem atrair investimentos privados para projetos de desenvolvimento e transferência de tecnologia; impulsionar o **desenvolvimento econômico e social** por meio da colaboração entre as IES, micro e macroempresas privadas, órgãos públicos e organizações sociais não governamentais.

- **Importância para os pesquisadores:** possibilitar a **preservação a longo prazo** dos dados gerados/coletados em suas pesquisas; demonstrar seu compromisso com a **transparência e integridade** da pesquisa disponibilizando o plano de gestão de dados; fortalecer a **credibilidade** dos resultados ao permitir que outros pesquisadores verifiquem e reutilizem os dados; contribuir para o **desenvolvimento de outras pesquisas** em áreas correlatas e interdisciplinares; obter **reconhecimento e créditos** por seu trabalho, elevando sua notoriedade no campo de estudo; participar do **movimento colaborativo** entre pesquisadores nacionais e estrangeiros, expandindo suas redes de relacionamento e criando oportunidades para pesquisas conjuntas.
- **Importância para o Brasil:** como nação democrática, permite a todas as pessoas físicas e jurídicas o acesso aos dados, fortalecendo a **democracia**; transparência na **prestação de contas** e decisões baseadas em **evidências**; desenvolvimento econômico; **melhoria** da educação e da **inovação**; assume efetivamente seu **compromisso** com as recomendações da OCDE e Unesco sobre a Ciência Aberta; **colaboração** para o desenvolvimento social, da ciência e da economia entre as nações.

Conforme o relatório da *Science-Metrix*<sup>3</sup> (2018), o Brasil está em 13º lugar entre os **países que mais produzem artigos científicos** no mundo, em primeiro lugar entre os que

---

<sup>3</sup> [Science Metrix](#) é uma empresa norte-americana dedicada a avaliar atividades ligadas à ciência e tecnologia.

publicam com o **Acesso Aberto** na *Web of Science* e **líder na América Latina** na produção científica.

Em 2023, o país subiu cinco posições no Índice Global de Inovação (IGI) em comparação com o ranking de 2022. Dentre 132 países, ocupa o 49º lugar, passando a ser o líder da América Latina e Caribe (Site da Serra, 2023). Este **protagonismo mundial é reconhecido** pelo desenvolvimento de Tecnologias de Informação e Comunicação; Energias Renováveis; Agricultura de Precisão; Indústria Aeroespacial; Ciências da Saúde e Biotecnologia; Mineração e Metalurgia; Petróleo e Gás; Biocombustíveis e diversas outras atividades tecnológicas e industriais (BNDES, 2018). O país exerce uma influência significativa na economia na política global, e também se destaca por sua riqueza cultural, conquistas esportivas e contribuições de destaque para a pesquisa. Nas novas formas da prática científica, as instituições devem aderir às tendências da GDP para se manterem relevantes no cenário internacional.

Para que todos esses benefícios possam se concretizar, é essencial que uma política seja desenvolvida para que a gestão seja viabilizada. As normativas do documento devem organizar os processos de gestão, sendo formuladas de maneira clara e objetiva, para garantir a eficiência e a eficácia de sua aplicação. O crescente destaque dos dados de pesquisa trouxe à tona uma série de novos serviços, termos, conceitos, atributos e procedimentos, nos quais o tratamento dos recursos de informação difere substancialmente das práticas tradicionalmente associadas aos recursos de informação convencionais.

Nesta pesquisa apresentada para obtenção do título de mestre pelo Programa de Pós-Graduação em Ciência da Informação, consubstanciada pela excelência do IBICT e da UFRJ na promoção e no suporte ao desenvolvimento científico e social, espera-se contribuir para o fortalecimento da gestão e compartilhamento dos dados de pesquisa.

Este trabalho está estruturado da seguinte forma: na seção 2, apresenta-se a literatura que contextualiza a GDP no cenário acadêmico atual, as redes de apoio que oferecem suporte e serviços, os desafios e as perspectivas da GDP nas IES. Na seção 3, descreve-se a metodologia utilizada para conduzir a pesquisa. Na seção 4, expõe-se os resultados e discute-se as informações levantadas durante a pesquisa. Na seção 5, são apresentadas as considerações finais da pesquisa.

## 2 REVISÃO DE LITERATURA

Este capítulo apresenta a literatura que fundamenta o aporte teórico e contextual para o desenvolvimento desta pesquisa. Apresenta o contexto dos dados de pesquisa e sua gestão, o compartilhamento e reuso, o contexto dos dados de pesquisa nas instituições de ensino superior no Brasil, os desafios e as perspectivas da gestão e uma apresentação dos termos e conceitos utilizados na área.

Em uma pesquisa realizada em abril de 2023 no Portal de Periódicos Capes, empregando os termos "*research data management*" e "*data sharing*", foram recuperados 3.573 resultados. Ao aplicar filtros de intervalo de tempo, recuperaram-se 179 trabalhos publicados entre 1970 e 2011, e entre 2012 e 2023, foram listados 3.394 trabalhos. Os números indicam que as iniciativas para a abertura e o compartilhamento dos dados de pesquisa não são muito recentes, mas nos últimos dez anos houve um aumento significativo, revelando uma diversidade de questões e buscas para aprimorar os serviços e viabilizar o gerenciamento. A literatura consultada nesta pesquisa será, majoritariamente, de estudos publicados entre 2012 e 2023.

### 2.1 A ERA DOS DADOS: UM MOVIMENTO MUNDIAL

No contexto desta pesquisa, a expressão “era dos dados” se refere principalmente aos dados abertos, os quais são definidos pelo *Open Knowledge Foundation* e *Open Data Institute* (2013) como dados disponíveis para acesso, uso, modificação e compartilhamento por qualquer pessoa, sem restrições e para qualquer finalidade. A noção de abertura ou restrição de acesso, independentemente de sua natureza, existe desde antes da era digital, do surgimento do movimento da Ciência Aberta e da *e-science*. No entanto, com o avanço das tecnologias de informação e comunicação e no contexto da *e-science*, tornou-se necessário desenvolver novas formas para promover um acesso que permitisse o Acesso Aberto aos dados de pesquisa como uma regra, e a restrição, como exceção. Essa transformação teve um impacto significativo em organizações de natureza pública e privada, educacional, comercial e

outras, obrigando-as a repensar suas abordagens em relação à disponibilização e utilização de dados.

No âmbito governamental, com a crescente reivindicação da sociedade por maior transparência, prestação de contas e acesso aos dados gerados/coletados, os setores públicos começaram a procurar meios para atender as demandas, a exemplo do governo dos Estados Unidos (EUA) que lançou em 2009, o *data.gov*. No ano seguinte, o Reino Unido (UK) disponibilizou o *data.gov.uk*, com os mesmos propósitos do governo dos EUA (Patrice, 2010). A necessidade de abertura aos dados e informações produzidos por órgãos públicos têm sido amplamente reconhecida e abraçada por diversos países.

O Brasil, que embora já dispusesse de algumas normativas municipais, estaduais e de órgãos públicos que atuavam isoladamente, somente em 2011 foi aprovada a “Lei de Acesso à Informação” (LAI) (BRASIL, 2011). Como parte dos programas para promover o Acesso Aberto de modo a “[...] garantir e facilitar o acesso pelos cidadãos, pela sociedade e, em especial, pelas diversas instâncias do setor público aos dados e informações produzidas ou custodiadas pelo Poder Executivo federal” (BRASIL, 2012) criou-se a Infraestrutura Nacional de Dados Abertos (INDA); e em 2015, a Controladoria Geral da União (CGU) e a Secretaria de Governo Digital (SGD) do Ministério da Economia lançaram o Portal Brasileiro de Dados Abertos (BRASIL, 2019).

Ainda atendendo a demandas específicas para acesso e proteção dos dados, foi promulgada a Lei Geral de Proteção de Dados Pessoais (LGPD) para regulamentar o “tratamento de dados pessoais, inclusive em meios digitais, por pessoas físicas ou jurídicas, de direito público ou privado” (BRASIL, 2018). Esta lei define o dado pessoal como “informação relacionada a uma pessoa natural identificada ou identificável”, e de dado pessoal sensível como “origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dados referentes à saúde ou à vida sexual, dados genéticos ou biométricos, quando vinculados a uma pessoa natural”. Segundo Mulholland (2018), a LGPD foi influenciada pelo Regulamento Geral de Proteção de Dados (GDPR) da União Europeia, também publicado em 2018.

Para evidenciar o valor dos dados, foram comparados com um novo petróleo, ou um novo ouro (Dijk; Doorn, 2015; Ball, 2017; The Economist, 2017). Assim como ocorre com qualquer ativo valioso, existem diversos interessados em se tornarem proprietários, motivados por diferentes razões, como, por exemplo, ganhos financeiros. Seja visando os lucros, seja por

um dever de instituições públicas e privadas, ou para fins relacionados às demandas científicas, é necessário investir em infraestrutura para a preservação e proteção dos dados, estabelecer termos e condições para o compartilhamento e uso, garantir uma documentação regulatória adequada e tomar outras medidas necessárias. A comercialização passou a ser motivo de grande preocupação internacional por parte da sociedade e do poder público. As leis criadas especificamente para esse fim ajudam a regular e proteger quanto ao uso indevido. Entretanto, mesmo com a proteção legal, os dados continuam sendo disponibilizados e comercializados sem a autorização ou o conhecimento de pessoas físicas e jurídicas (Borgman, 2018).

Quanto à relação comercial entre as IES e o setor privado, a Lei nº 13.243, de 2016 dispõe sobre o desenvolvimento científico, a pesquisa, a capacitação científica e tecnológica e a inovação no ambiente empresarial, acadêmico e governamental. A lei não trata especificamente do uso comercial de dados entre empresas privadas e públicas, mas destacam-se as previsões de acordos e contratos prévios para direitos de uso, exploração, propriedade intelectual e patentes.

Os dados estão moldando o futuro da humanidade com o avanço das tecnologias digitais em um ritmo sem precedentes. Sales (2014, p. 71), alerta para o fato de que o “acesso aos dados no futuro vai depender de como vamos equacionar todas as variáveis que se sobrepõem”.

## 2.2 OS DADOS DE PESQUISA: NOVAS PERSPECTIVAS

Conforme Wallis, Rolando e Borgman (2013), os dados de pesquisa são como a pedra angular da ciência. Para viabilizar o compartilhamento, é fundamental promover um esforço de trabalho coletivo que envolva pesquisadores, instituições, organizações e até mesmo colaborações internacionais (Borgman; Bourne, 2021).

O compartilhamento e reuso de dados de pesquisa não é uma atividade recente, mas a diversificação de usos, propósitos e conceitos se tornou mais abrangente. Em meio à era dos dados, o conceito de dados de pesquisa mais amplamente aceito é de que são “registros factuais usados como fonte primária para a pesquisa científica e que são comumente aceitos pelos pesquisadores como necessários para validar os resultados do trabalho científico”

(OCDE, 2007, p. 7). Para Chawinga e Zinn (2019, p. 2) “compartilhamento de dados é inevitável em um mundo moderno, onde parte do novo conhecimento é descoberto pela reanálise de dados gerados anteriormente” e devem estar à disposição para o público com o objetivo de fornecer acesso para reutilização. Quanto a fornecer acesso, Borgman (2018) alerta para que se devem observar os casos previstos pelos aspectos éticos e legais para proteger a privacidade dos participantes ou confidencialidade das informações.

A OCDE (2007) aponta também que o valor dos dados implica o seu uso e ainda que o acesso total e aberto aos dados de pesquisa deve ser adotado internacionalmente para o intercâmbio, definições comuns, repositórios nacionais ou multinacionais, padrões de segurança, processos de certificação e diversas outras recomendações para cooperação internacional para o progresso das pesquisas científicas.

Identificar os pioneiros em reconhecer os benefícios e adotar o compartilhamento de dados de pesquisa é uma tarefa difícil, pois houve várias iniciativas individuais e institucionais que surgiram há muito tempo, especialmente nas áreas biomédicas, físicas e astronômicas (Borgman; Bourne, 2021). No entanto, nos últimos vinte anos, as publicações e os incentivos para a promoção da gestão e do compartilhamento desses dados ganharam a atenção de todas as áreas do conhecimento. Algumas iniciativas realizadas por organizações proeminentes, como a OCDE contribuíram de forma significativa na divulgação e promoção da GDP. O relatório *Princípios e Diretrizes para o Acesso a Dados de Pesquisa com Financiamento Público*, publicado em 2007, enfatiza o reconhecimento da organização quanto ao compartilhamento dos dados de pesquisa para aumentar a participação dos países no sistema científico mundial, facilitando o desenvolvimento social e econômico e a formação de novos pesquisadores. O relatório também acabou fortalecendo a ideia da GDP ao recomendar que os dados de pesquisa provenientes de financiamento público deveriam ser de acesso aberto (OCDE, 2007). Esta recomendação continua a ser uma referência importante e valiosa na elaboração de políticas nesta área cada vez mais importante. Em 2020, foi publicado o relatório *Acesso Aprimorado a Dados com Financiamento Público para Ciência, Tecnologia e Inovação* que trouxe atualizações sobre o tema, apresentando novas diretrizes para promover os dados abertos, incluindo os de pesquisa, os quais também abordam impactos, desafios, práticas em benefício da gestão de dados (OCDE, 2020).

Sobre a gestão de dados de pesquisa, Chawinga e Zinn (2019) e Sales e Sayão (2019) definem como um conjunto de ações para organizar, armazenar, proteger, preservar e

compartilhar os dados gerados ou coletados durante o processo de pesquisa. Esse processo antecede à própria pesquisa, pois requer a elaboração de um PGD, no qual é informado como os dados serão gerados/coletados, processados, métodos e padrões, condições para compartilhar e outras definições para o gerenciamento (Sales; Sayão, 2015).

Existe uma pressão cada vez maior para que os dados estejam acessíveis ao público tanto quanto possível (Dillo; Doorn, 2014), mas como a ciência é fortemente orientada por dados e processada por máquinas, o acesso pode se tornar crítico, pois os formatos de arquivos mudam constantemente e softwares são substituídos (Mons, 2019). Essas questões demandam serviços de curadoria digital para garantir o acesso e reuso ao longo do ciclo de vida dos dados, sendo este um dos principais desafios na gestão. Conforme Sayão e Sales (2019, p. 4.160) “existem poucos especialistas em curadoria de dados para assisti-los” e Heidorn (2008, p. 291) citado por Sayão e Sales (2019, p. 4.160), informa que está “faltando número suficiente de indivíduos familiarizados com as instituições e tecnologias disponíveis para tornar os dados visíveis para um grande número de pessoas”. Ou seja, o problema apontado por Heidorn, quinze anos atrás, ainda ecoa nos dias de hoje.

Sales (2014, p. 54), esclarece que:

Tanto a gestão quanto o estabelecimento de políticas de preservação fazem parte de um leque de atividades que são englobadas pelo conceito de curadoria digital. Por ser uma área emergente, ainda há algumas divergências sobre o seu conceito, sendo muitas vezes confundido com preservação digital, arquivo digital, entre outros.

As considerações de Sales (2014, p. 54) são reforçadas por Almeida (2019, p. 45) quando informa que “Apesar de os conceitos de gestão de dados de pesquisa e curadoria estarem bastante imbricados, acreditamos que a curadoria e a preservação de dados são componentes da gestão de dados de pesquisa”, e “engloba as infraestruturas, políticas, questões legais, éticas e propriedade intelectual, seleção e coleção, descrição, organização, acesso, compartilhamento, reuso, segurança, custos, capacitação, dentre outros”.

Para o *National Coordination Point Research Data Management* (LCRDM) a curadoria de dados é uma a atividade de gestão para promover o uso desde a sua criação, a fim de garantir que esteja disponível para descoberta e reutilização no futuro, seguindo os processos de seleção, anotação, manutenção, arquivamento e rastreamento de dados.

Os novos serviços e novas especialidades profissionais, como administradores de dados e cientistas de dados, estão surgindo e se integrando às equipes compostas por

pesquisadores, técnicos de laboratório, assistentes de pesquisa e analistas (Sayão; Sales, 2021). A especialização desses profissionais está sendo delineada e será um desafio atribuir funções e definir as qualificações e os critérios para sua nomeação. Para Sales e Sayão (2015, 2018) será necessário um profissional por disciplina, pois precisam ser especialistas em suas respectivas áreas, uma vez que o serviço deve estar presente em todo o ciclo de vida dos dados, ou seja, na ocasião em que são gerados/coletados até o seu descarte.

Antes do reconhecimento mundial das potencialidades de uso e reuso, estes eram considerados como subprodutos e, frequentemente, seu tratamento era inadequado ou negligenciado, correndo o risco de serem perdidos e/ou descartados (Sayão; Sales, 2014). Essas reflexões são compartilhadas por Torino, Roa-Martínez e Vidotti (2020, p. 184), ao afirmarem que os dados gerados ao longo dos séculos foram “muitas vezes esquecidos ou relegados a uma relevância inferior à dos artigos científicos”. A evolução gradual reflete o reconhecimento cada vez maior das potencialidades dos dados de pesquisa e a necessidade de tratá-los de maneira adequada e valorizada (Torino; Roa-Martínez; Vidotti, 2020; Sayão; Sales, 2021).

O termo *e-Science* foi introduzido por John Taylor<sup>4</sup>, em 2001, para designar “a ciência crescentemente desenvolvida por meio de colaborações globalmente distribuídas viabilizadas pela Internet, usando grandes coleções de dados, recursos computacionais em escala tera e visualização de alto desempenho” (Maciel; Appel; Albagli, 2013). A Unesco (2022, p. 7-9) define Ciência Aberta e Dados de Pesquisa Abertos, respectivamente, como:

[...] um construto inclusivo que combina vários movimentos e práticas que têm o objetivo de disponibilizar abertamente conhecimento científico multilíngue, torná-lo acessível e reutilizável para todos, aumentar as colaborações científicas e o compartilhamento de informações para o benefício da ciência e da sociedade, e abrir os processos de criação, avaliação e comunicação do conhecimento científico a atores da sociedade, além da comunidade científica tradicional

[...] dados digitais e analógicos, tanto brutos como processados, e os metadados que os acompanham, assim como pontuações numéricas, registros textuais, imagens e sons, protocolos, códigos de análise e fluxos de trabalho que podem ser usados, reutilizados, retidos e redistribuídos abertamente por qualquer pessoa, sujeitos a reconhecimento.

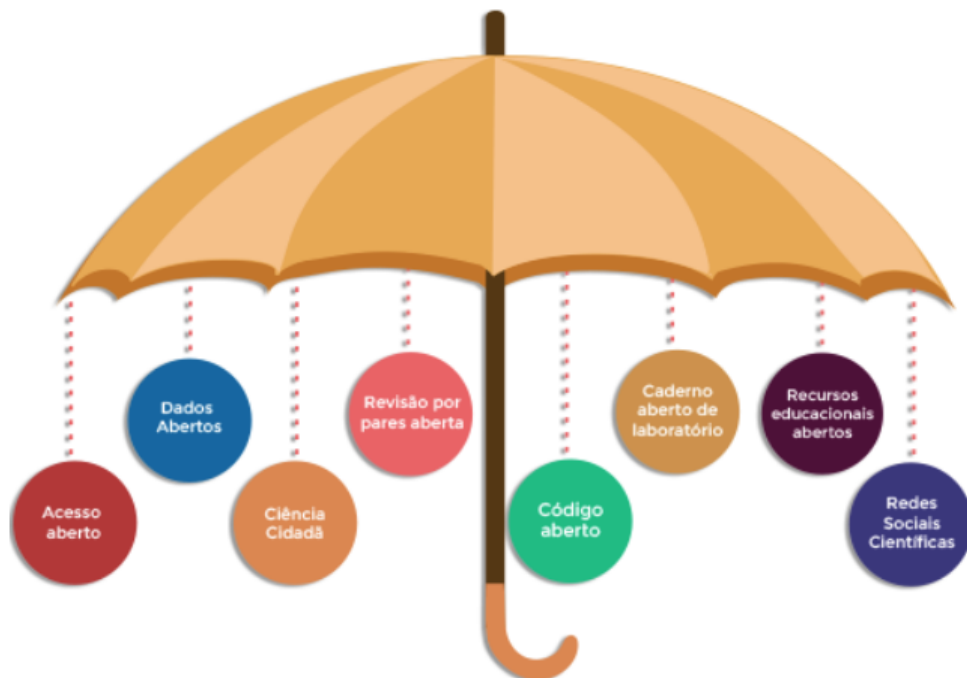
---

<sup>4</sup> John Taylor foi diretor geral do Conselho de Pesquisa do *Office of Science and Technology* do Reino Unido, em 2001.



Dentro do “guarda-chuva”<sup>5</sup> da Ciência Aberta (uma alegoria comumente usada na área), estão contemplados alguns eixos, ilustrados na imagem Figura 1.

Figura 1 – Guarda-chuva da Ciência Aberta



Fonte: Autoria incerta. A imagem foi extraída no site da Fiocruz (2021)

Neste novo paradigma, os elementos propulsores e interdependentes da ciência moderna encontram-se em um esquema indissociável. Em um modelo mais abrangente, Pontika *et al.* (2015, p. 3) apresentaram um modelo de taxonomia da área para “mapear o campo e oferecer uma representação aprofundada dos conceitos em torno do sistema operacional” e “fornecer uma terminologia estruturada e consistente que reduzisse os atuais conflitos descritivos no campo” para promover seu desenvolvimento por meio “das práticas abertas que se relacionam com a terminologia da Ciência Aberta”. No modelo proposto, a Ciência Aberta é ilustrada como uma fonte central que desencadeia uma série de ramificações e efeitos em diversas áreas. Como ilustrado na Imagem 1, essa abordagem coloca a Ciência Aberta como um catalisador para o avanço do conhecimento, incentivando a transparência, a colaboração e o acesso aos resultados de pesquisa para diversas finalidades. Na Figura 2, pode-se visualizar que, à medida que a Ciência Aberta se expande, suas ramificações se

<sup>5</sup> A autoria da Figura 1 não foi localizada. Por essa razão, citou-se a Fiocruz, de onde foi extraída.

estendem além dos limites disciplinares e institucionais, promovendo uma cultura de compartilhamento dos métodos e das descobertas científicas.

Figura 2 – Taxonomia da Ciência Aberta

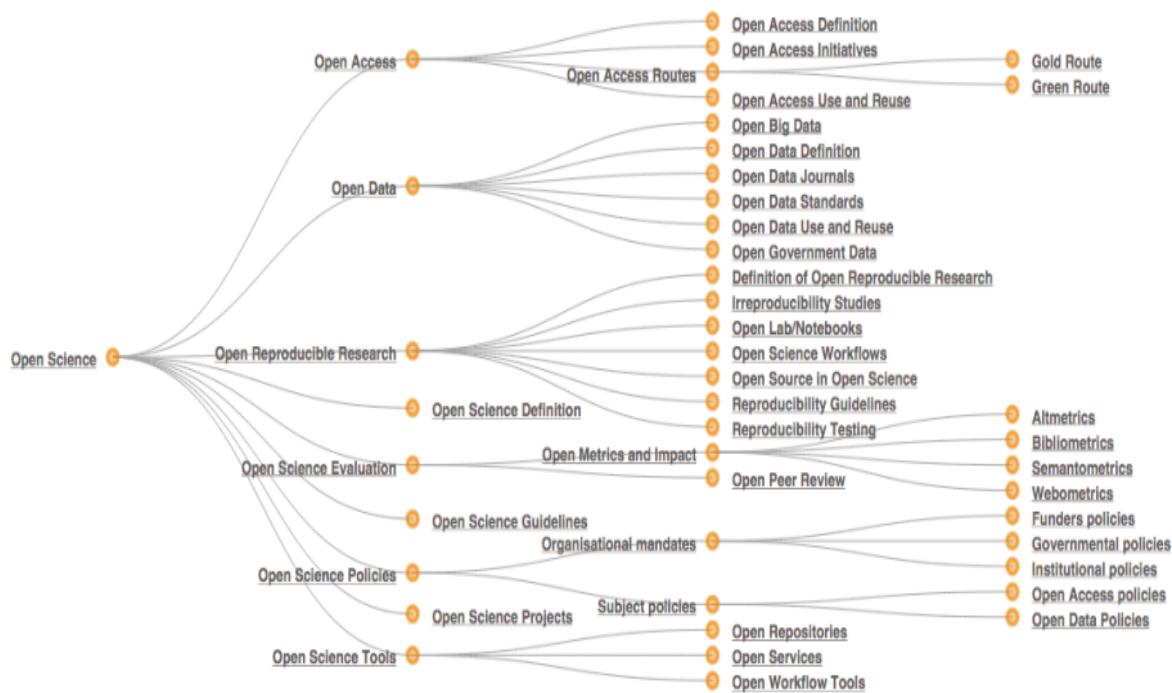
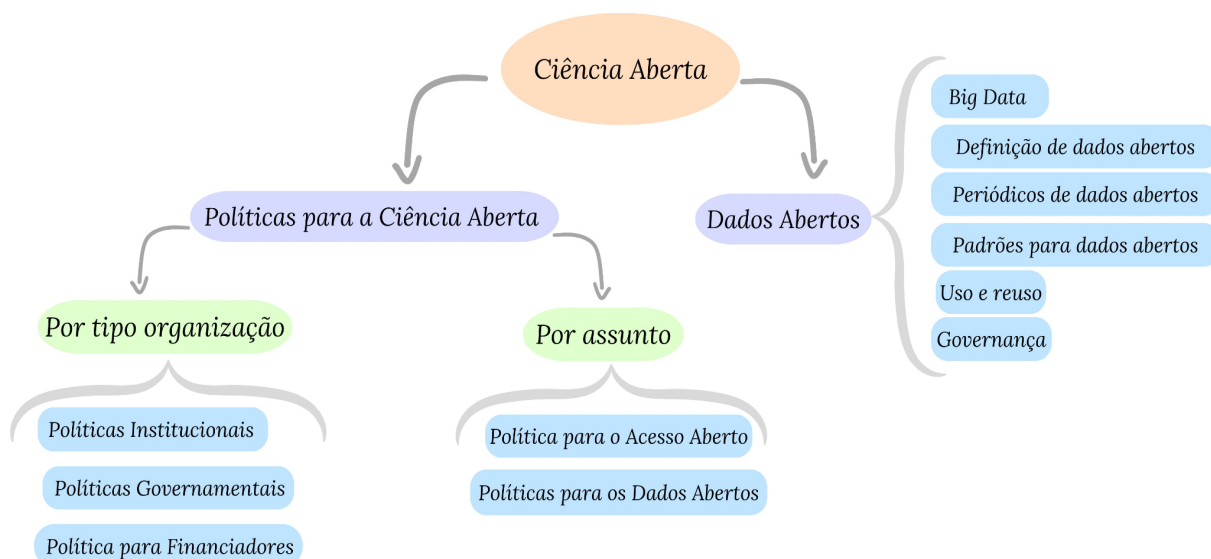


Figure 1. Open Science Taxonomy

Fonte: PONTIKA *et al.*, 2015, p. 3.

O modelo de taxonomia apresentado demonstra uma estrutura de Ciência Aberta com seus eixos e ramificações, abrangendo uma ampla variedade de atividades que englobam a organização do conhecimento, exemplificando a extensão e a diversidade de possibilidades existentes nessa área. Nesse contexto, os documentos regulatórios, como as políticas, desempenham um papel crucial, pois os desdobramentos possuem características distintas que podem ser interoperáveis, mas devem ser tratados conforme seu eixo dentro da Ciência Aberta. A documentação, tanto a de governança quanto a específica do elemento, torna essencial para garantir o funcionamento das engrenagens do sistema proposto.

Figura 3 – Os dados de pesquisa em evidência



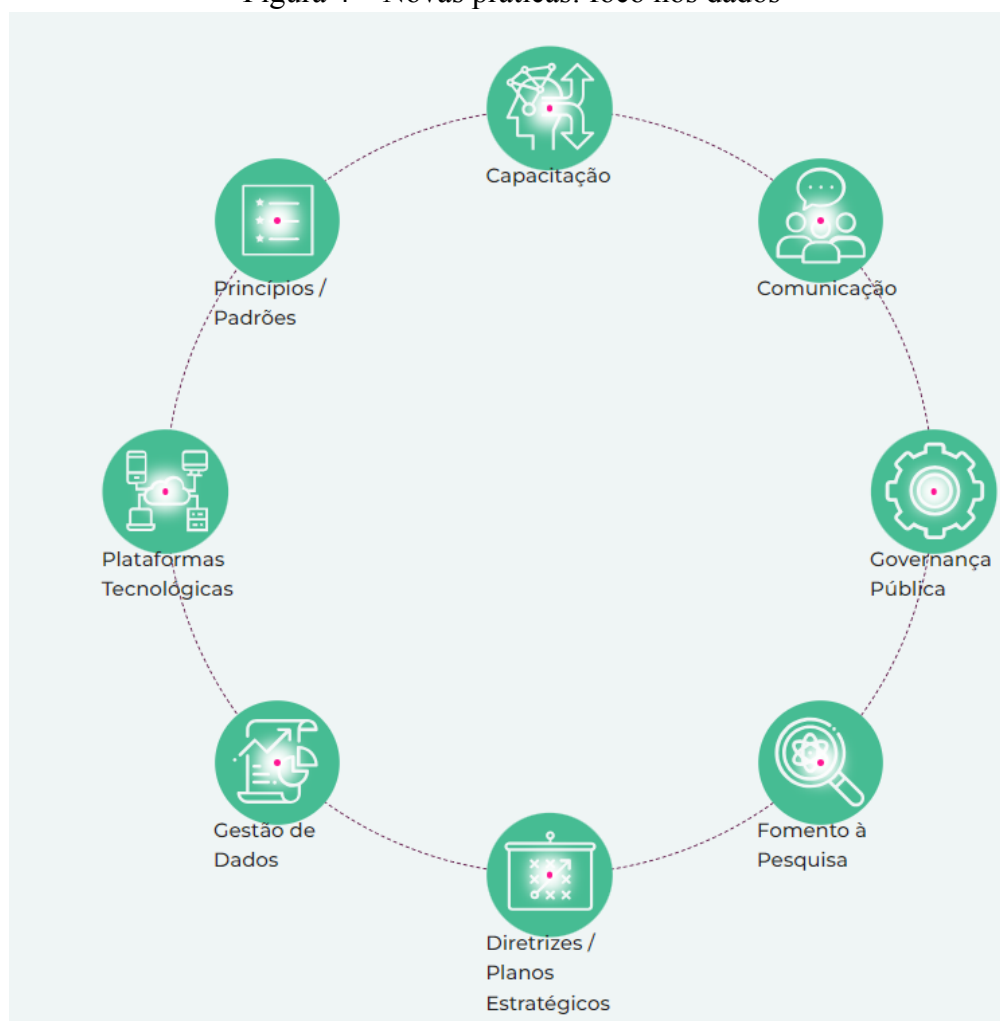
Fonte: Elaborada pela autora, 2023.

Pontika *et al.* (2015, p. 3) evidencia a interconexão da diversidade de serviços comuns na Ciência Aberta, mesmo que as linhas não se entrelacem diretamente. A incidência dos dados de pesquisa e políticas é destacada pela quantidade de desdobramentos dentro das ramificações deste modelo, como evidenciado na Figura 3.

Ao afirmarem que dados são a “pedra angular da ciência”, Wallis, Rolando e Borgman (2013) indicam que isso se deve à sua capacidade de viabilizar amplas oportunidades por meio de compartilhamento, reuso e combinação. Todavia, para a possibilidade de reanálise de evidências, verificação de resultados, minimização de esforços duplicados e aceleração da inovação, o GDP exige um esforço coletivo por parte de pesquisadores, instituições, organizações e até mesmo cooperação entre nações (OCDE, 2007, 2020; Borgman, Bourne, 2021; Sayão; Sales, 2015; Singh, Monu, Dhingra, 2018; Pasquetto; Randles; Borgman, 2017).

Na Figura 4, a Fiocruz (2018) apresenta um ecossistema que considera que “as estratégias adotadas pelos países para a implantação da Ciência Aberta convergem para a criação de um novo ecossistema de pesquisa adequado às novas práticas”.

Figura 4 – Novas práticas: foco nos dados



Fonte: FIOCRUZ, 2018.

Esta dinâmica apresentada representa os fluxos de GDP em uma instituição de pesquisa. A Figura 5, elaborada por Costa (2017) apresenta uma gestão de dados científicos como uma engrenagem, onde o movimento de uma peça impulsiona e depende da outra. As representações das Figuras 4 e 5 sugerem uma dinâmica para a Ciência Aberta, onde os componentes dependem da ação de *stakeholders*. É possível que a elaboração de uma taxonomia para as partes envolvidas seja uma abordagem útil para alinhar os movimentos síncronos.

Figura 5 – Reflexões sobre a gestão de dados científicos



Fonte: COSTA, 2017, p. 61.

Na perspectiva do DCC (2013), citado por Costa e Cunha (2019) o planejamento da gestão de dados é composto pelos itens: coleção de dados (quais e como dados serão coletados); Documentação e Metadados; Ética e *Compliance* Legal; Armazenamento e Backup; Retenção e Preservação; Compartilhamento de dados; Responsabilidades e recursos. Esses itens também são abordados por Sayão e Sales (2015), e acrescentam o ciclo de vida e a citação dos dados.

### 2.3 O ACESSO ABERTO, CIÊNCIA ABERTO E O BRASIL

No Brasil, a maior parte das pesquisas científicas são realizadas nas instituições de ensino superior. Logo, podemos inferir que estas também geram e coletam a maior quantidade de dados de pesquisa. Quando se discutiam problemas e soluções para viabilizar o Acesso Aberto aos documentos textuais “escritos e publicados”, Tomaél e Silva (2007) apontaram que algumas informações produzidas não estão disponíveis para os acadêmicos da própria

instituição e, tampouco, para a comunidade científica. Os documentos que não foram publicados ou que não estão acessíveis formam a grande massa de informação e podem ser perdidos, pois não há formas eficientes para o acesso (Tomaél; Silva, 2007; Sayão; Sales, 2014). Tomaél e Silva (2007, p. 2) afirmam que:

Isto ocorre em função de vários aspectos que perpassam o complexo ambiente da comunicação científica, dentre eles, a morosidade dos processos de publicação, os custos envolvidos na publicação de documentos, a grande quantidade de documentos impressos – e mesmo de documentos eletrônicos – dispersos nas coleções.

Os problemas relatados são similares aos enfrentados hoje quanto à viabilização de Acesso Aberto aos dados de pesquisa.

Com a implementação dos Repositórios Institucionais (RI), os problemas relacionados ao acesso a documentos escritos e publicados foram reduzidos, embora existam ainda, infelizmente, IES que não adotaram esta ferramenta. Ainda em analogia aos RI, Gibbons (2004), citado por Tomaél e Silva (2007, p. 9), levantou questões que também precisam ser respondidas para a GDP:

- Quem pode depositar trabalhos no repositório institucional? Somente os membros da organização ou também membros externos?
- Tipos de materiais e objetos digitais que poderão ser incluídos?
- O repositório institucional poderá ser utilizado para propósitos comerciais?
- A coleção será formada por materiais depositados pelos próprios autores ou por terceiros que se responsabilizarão pelo depósito dos materiais dos autores?
- Haverá aprovação prévia para a disponibilização do material?
- Os documentos depositados no repositório institucional poderão ser atualizados?
- Quem será o responsável legal do repositório institucional?

Na situação atual do Brasil, as questões mencionadas ainda estão sem respostas (conforme documentos já consultados). A adoção das práticas para GDP ainda enfrenta obstáculos significativos, como um menor envolvimento das IES, agências de fomento, periódicos nacionais e institutos de pesquisa em relação ao cenário internacional. Assim como ocorreu com o movimento em prol do Acesso Aberto aos documentos textuais, em que foi necessário o desenvolvimento de softwares, ferramentas, fluxos de trabalho e, sobretudo, de políticas institucionais para viabilizar os serviços, as Instituições de Ensino Superior agora se deparam com uma nova demanda: a elaboração de **políticas institucionais para a gestão dos dados de pesquisa**.

O conceito de Política para Gestão e Dados de Pesquisa é apresentado por Sayão e Sales (2018) como sendo um conjunto de instruções fundamentais para orientar/instituir normas, planos, condutas e procedimentos a serem adotados pelos membros de uma instituição, visando a gestão e o compartilhamento de dados de pesquisa. O Portal Foster ([20--]) conceitua como “um conjunto de princípios, normalmente exigidos por vários órgãos, que devem ser seguidos ao gerenciar dados de pesquisa”. Sayão e Sales (2018, p. 99) complementam:

[...] os compromissos que a instituição se obriga em relação aos seus principais stakeholders – pesquisadores, curadores, consumidores, financiadores, coletores de metadados entre muitos outros – e em relação às diversas etapas do ciclo de vida das coleções de dados que estão sendo gerenciadas. Esse posicionamento idealmente deve estar manifestado em um documento publicado na página web da plataforma explicitando a política da instituição em relação ao serviço que está sendo disponibilizado. Pode incluir: política de conteúdo, de submissão e de depósito, de direitos autorais, de acesso e reuso da informação, de preservação digital e questões éticas entre outras. A política deve permear todos os processos da gestão de dados, além do mais, ela deve ser um rebatimento harmônico das políticas institucionais e nacionais e das diretrizes internacionais.

#### 2.4 A GDP NO BRASIL: OS PRIMEIROS PASSOS

Algumas IES já estão oferecendo algum tipo de serviço relacionado ao planejamento e gerenciamento dos dados, majoritariamente, no estado de São Paulo. A Fapesp, foi a primeira agência de fomento a solicitar a apresentação de um PGD na submissão de projetos à agência, a partir de 2017 (Almeida, 2019, p. 19, Costa E Cunha, 2019, p. 302). No texto, a Fundação “[...] reconhece a importância da gestão adequada dos dados de pesquisa como parte essencial das boas práticas” e considera necessário que os dados sejam gerenciados e compartilhados (Fapesp, [201-b]) [2019]. A fundação segue a tendência internacional, como financiadora de pesquisa, mobilizar as instituições que produzem ciência, a planejar a gestão dos dados de pesquisa, como indicam Higman e Pinfield (2015).

Sayão e Sales (2013) enfatizam que a infraestrutura organizacional, o desenvolvimento de coleções de dados, a pesquisa, a formação de recursos humanos, a sustentabilidade econômica e as implicações sociais, legais e éticas não devem ser negligenciados pelas instituições. Esses componentes não são uma novidade para as academias, portanto, espera-se que algumas dessas demandas já tenham sido atendidas. A implementação da GDP

necessitaria de recursos adicionais, como infraestrutura mais robusta e pessoal especializado. Os recursos disponíveis podem ser identificados por meio de diagnósticos institucionais e de um planejamento estratégico para mapear as possibilidades e lacunas e focar nas necessidades de implantação (Jones; Pryor; Whyte, 2013).

Não é correto e nem seguro se restringir apenas às demandas de agentes externos, pois deve-se garantir que as IES não sejam meramente subordinadas às normas das editoras e financiadores (Lewis, 2010). Com os devidos cuidados, as diretrizes para a GDP podem ser construídas com base em coerência e em conformidade com a realidade da instituição. Costa e Cunha (2019, p. 293), questionam:

Em alguns casos, a diretriz institucional chega a ser insuficiente dada a complexidade da questão, como por exemplo – se o pesquisador receber fomento de agências internacionais e elas exigirem o compartilhamento de dados como contrapartida do apoio, qual a orientação política do Brasil quanto ao caso?

É fato que o pesquisador precisa se submeter aos termos do edital, considerando a necessidade de financiamento da pesquisa. A política institucional para GDP pode orientar ou instruir as agências de fomento a construírem editais que não prejudiquem os direitos da IES e dos pesquisadores, uma vez que essas agências também têm interesse nos resultados das pesquisas. As diretrizes estabelecidas por ambas as partes podem ser tornadas interoperáveis ou, pelo menos, negociáveis, garantindo que todos tenham ganhos juntos.

Costa e Cunha (2019, p. 293) informam que “parece coerente afirmar que um conjunto de respostas às indagações precisa ser elaborado por meio de uma política pública para a GDP”. Como uma possível iniciativa nacional, o IBICT, dentre suas metas apresentadas no Programa Brasileiro de Acesso Aberto a Dados de Pesquisa, informou que até 2019, pretendiam estruturar um padrão mínimo de metadados de acordo com as diretrizes internacionais; incentivar e apoiar a implantação de dez repositório de dados para as instituições de pesquisas nacionais; e promover três novos estudos em grupos de pesquisa sobre acesso aberto a dados de pesquisa (IBICT, 2018). No Termo de Compromisso e Gestão 2022 (TCG), publicado em junho de 2023<sup>6</sup>, consta como uma das Metas de Desempenho do Plano Diretor a promoção de três novos estudos em grupos de pesquisa sobre Acesso Aberto a

---

<sup>6</sup> Todos os TCG do IBICT podem ser consultados no link:  
<https://www.gov.br/mcti/pt-br/acesso-a-informacao/unidades-de-pesquisa/termos-de-compromisso-de-gestao/ibict-instituto-brasileiro-de-informacao-em-ciencia-e-tecnologia>.

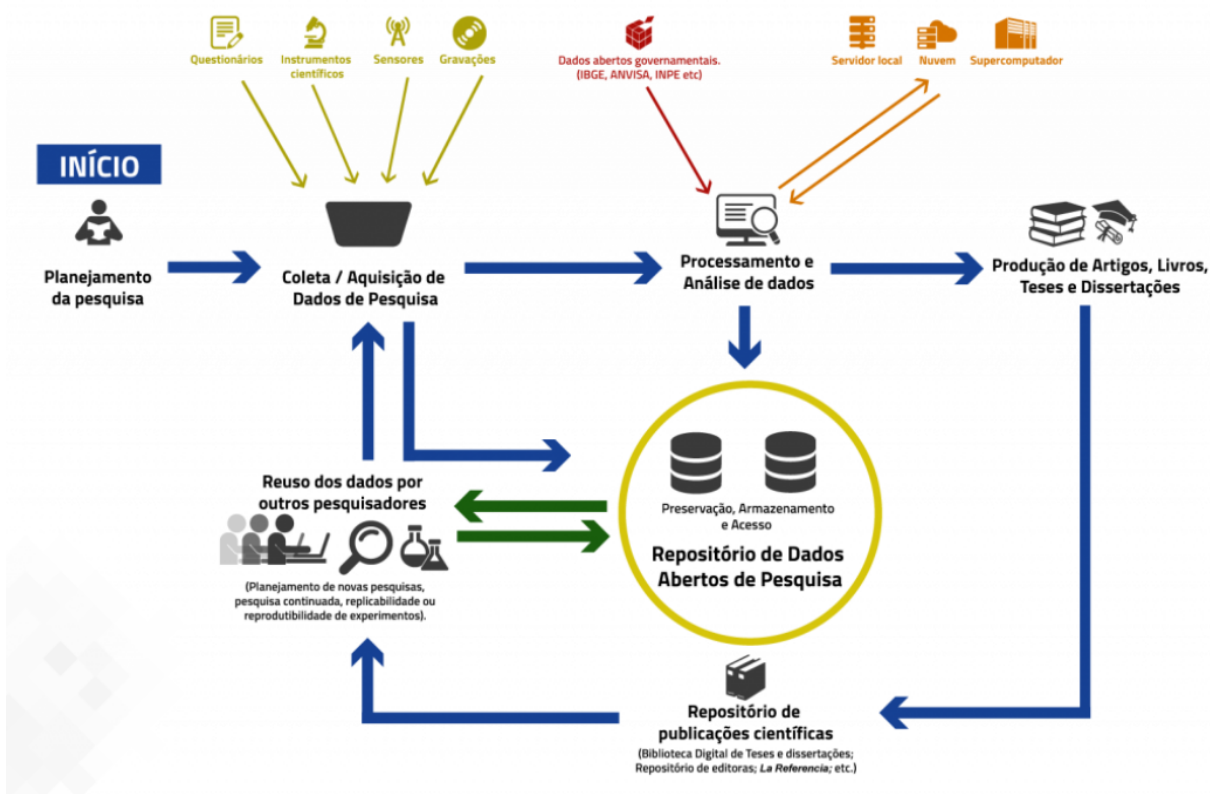


Dados de Pesquisa, além da criação de uma rede de instituições e pesquisadores voltada ao mesmo tema no Brasil. Outra iniciativa foi o lançamento em 2021, do edital da Rede Nacional de Ensino e Pesquisa (RNP), CNPq e IBICT para selecionar projetos para a criação e incubação de RD no país. Também por meio da RNP e IBICT, o Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira (BRASIL, 2018a) foi composto em conjunto com a Universidade Federal do Rio Grande do Sul (UFRGS) e a Universidade Federal do Rio Grande (FURG) com o objetivo de:

[...] executar atividades que contribuam para a identificação de práticas, mapeamento de requisitos e prototipação de sistemas que facilitem a disseminação de informações científicas. O projeto visa colaborar com a criação de um Programa Nacional de Acesso Aberto a Dados de Pesquisa, uma iniciativa com o objetivo de promover e incentivar o compartilhamento de dados entre pesquisadores, conferindo maior eficiência na produção de conhecimento científico no Brasil (BRASIL, 2018a).

As metas da RNP e do IBICT são a inclusão de procedimentos para gerenciar os dados de pesquisa nos processos de disponibilização e disseminação da informação. Isso é uma promessa para o desenvolvimento dos serviços científicos no país. No entanto, para que esta rede de procedimentos funcione, é necessário um sistema de governança que organize e estabilize práticas institucionais com a participação de *stakeholders*, e aporte documental.

Figura 6 – Dados de pesquisa: promessas da RNP e do IBICT



Fonte: BRASIL, 2018a.

Borgman e Bourne (2021) identificam como *stakeholders* internos de instituições de ensino superior: os pesquisadores individuais e os líderes de pesquisa, responsáveis pela geração/coleta dos dados e organização dos trabalhos; as instituições acadêmicas e suas lideranças, responsáveis por estabelecer políticas (por meio da criação de comissões) e fornecer suporte para GDP; as equipes de tecnologia de informação (TI), no gerenciamento da infraestrutura tecnológica; as bibliotecas acadêmicas, com serviços e recursos relacionados; e os departamentos e escolas, que apoiam a gestão dos dados em suas áreas específicas de pesquisa, seja com setor de apoio seja como disciplinas ofertadas, geralmente nos cursos de pós-graduação.

Nesta complexa rede, o estabelecimento de política formal e mandatória é fundamental para compor o processo da governança dos dados de pesquisa. A Governança, no conceito de Dillo e Doorn (2014), concentra-se em políticas e procedimentos, compondo o processo de gerenciamento da disponibilidade, usabilidade, integridade e segurança dos dados em sistemas corporativos, com base em padrões e políticas internas de dados que também

controlam o uso de dados. Nas considerações de Borgman (2018), a Governança de dados é um processo que indica “quais entidades têm quais responsabilidades, com base em quais critérios e políticas” (Borgman, 2018, p. 384). A Fiocruz (2020) indica que a Governança “Envolve políticas, processos e pessoas, além de uma estrutura organizacional para apoiar o gerenciamento de dados institucionais”. É evidente o papel que as políticas desempenham nesse cenário.

Nas leituras das políticas das IES brasileiras localizadas, verificou-se que, além da necessidade de maior abrangência, falta a inclusão de componentes de Governança para efetivar as práticas de gestão. Conforme Ribeiro e Oliveira (2019, p. 3):

A realidade brasileira voltada à Ciência Aberta e aos dados de pesquisa abertos nas universidades públicas federais encontra-se em um estágio incipiente. Tal afirmação baseia-se no número de produções acadêmico-científicas acerca do assunto, nas iniciativas em andamento e na ausência de políticas e diretrizes que guiam essas ações.

Costa e Cunha (2019, p. 299) indicam que, em termos de política explícita (formalmente estabelecidas e anunciadas), observa-se que algumas instituições, “ainda que de forma embrionária, usam sua autonomia para desenvolver políticas locais que atendam aos editais de fomento internacionais”. É nesse momento que boa parte dos pesquisadores tomam conhecimento sobre o planejamento. Sobre a gestão colaborativa, Costa e Cunha (2019, p. 303, grifo nosso), advertem que:

O papel das universidades não pode ser subestimado. A exemplo do processo de organização e preservação das teses e dissertações produzidas no Brasil, sugere-se que as **universidades** e os **institutos de pesquisa brasileiros** dividam a responsabilidade sobre a organização desses dados **em parceria com o IBICT**. Entende-se que uma instituição sozinha não terá condições de abraçar todo o processo de gestão de dados científicos no Brasil em função da sua complexidade, dimensão territorial, das dificuldades orçamentárias e da escassez de recursos humanos, dentre tantos outros fatores. Nesse cenário, a formação de uma espécie de consórcio para a gestão de dados científicos parece profícua no sentido de aproveitar-se do que cada instituição possui de melhor.

O cenário descrito por Costa e Cunha (2019) foi um caso de sucesso na Holanda em que, conforme Dillo e Doorn (2014), foram desenvolvidos uma infraestrutura e um modelo federal para armazenamento, *backup* e serviços *de back office* (traduzido como retaguarda), que é o processo de gestão interna que fornece recursos e suporte para arquivamento e acessibilidade a longo prazo; e *front office* (linha de frente) para apoiar, aconselhar e oferecer

treinamento aos pesquisadores e estudantes no gerenciamento de dados (Dillo, Doorn, 2014; Dijk; Doorn, 2014). Os autores relatam que essa infraestrutura e esse modelo de referência para universidades e institutos de pesquisa foram desenvolvidos pelo *Data Archiving and Networked Services* (DANS), que é um instituto nacional de dados e serviços para as ciências humanas e sociais que tem como objetivo promover a gestão, o arquivamento e a acessibilidade de dados de pesquisa na Holanda. O DANS é uma organização independente, financiada pelo Ministério da Educação, Cultura e Ciência da Holanda que, além dos serviços citados, possui um Repositório de Dados desde 2005, sendo considerado um dos repositórios mais importantes da Europa (DANS, 2023).

Os progressos das instituições de pesquisa e ensino da Holanda ganharam notoriedade do DDC (2019) que evidenciam que os PGD e os Princípios FAIR estão se tornando comuns na Holanda. Os dois principais financiadores, o ZonMw (sigla criada por meio de abreviações, sem tradução exata) e *Netherlands Organisation for Scientific Research* (NWO), exigem que a gestão de dados seja realizada conforme os Princípios FAIR nas solicitações de financiamentos.

Novamente citando as universidades holandesas, também têm a seu favor a disponibilidade de códigos de conduta que abordam a integridade da pesquisa, como o *Netherlands Code of Conduct for Research Integrity* (2018) elaborado pela *Association Of Universities In The Netherlands* (VSNU), pela *Royal Netherlands Academy Of Arts And Sciences* (KNAW); *Netherlands Federation Of University Medical Centres* (NFU); *Associated Applied Research Institutes* (TO2 FEDERATION); *Netherlands Association Of Universities Of Applied Sciences* e a NWO. A colaboração entre diferentes organizações, com documentos reguladores em comum, promove práticas de conduta mais fluidas e coerentes. Além do Código citado, as IES holandesas também se beneficiam das instruções fornecidas pelo *The European Code of Conduct for Research Integrity*, publicado pela *All European Academies* (ALLEA, 2017), documento norteador para as academias de países que integram a União Europeia e o *General Data Protection Regulation* (GDPR). Outros documentos importantes são a Lei do Ensino Superior e da Pesquisa Científica (HOLANDA, 2002) que, embora não trate especificamente da gestão de dados de pesquisa, estabelece requisitos e padrões para IES holandesas e a Lei do Governo Aberto (HOLANDA, 2023) que estabelece a transparência e a divulgação de informações e ainda que não tenha como foco a gestão de dados, impactam na disponibilidade de Acesso Aberto a dados para a pesquisa.

Embora os objetivos da gestão de dados possam variar entre diferentes categorias institucionais, o desenvolvimento de boas práticas para o acesso e reuso de dados é de interesse coletivo. Movidos pela necessidade de melhorar a infraestrutura para reutilização de dados, um grupo formado por representantes de academias de pesquisa, indústria, agências de financiamento e editores acadêmicos, especialistas em dados abertos se reuniram em uma conferência chamada *Jointly Designing a Data Fairport*, na cidade de Leiden, na Holanda, em 2014, e apresentaram um “rascunho de formulação de um conjunto de princípios fundamentais” (Wilkinson *et al.*, 2016, p. 3). Posteriormente, em 2015, um grupo de trabalho *FAIR Principles Working Group* foi formado para ajustar e aprimorar com mais detalhes sobre esses princípios (Wilkinson *et al.*, 2016, p. 3, European Commission, 2020). Assim, após consultas públicas, revisões e discussões, foram elaborados documentos e diretrizes que oferecem orientações mais claras, denominados como Princípios FAIR (*Findable, Accessible, Interoperable e Reusable*). O grupo apontou que esses princípios “colocam ênfase específica no aprimoramento da capacidade das máquinas de encontrar e usar automaticamente os dados, além de apoiar sua reutilização por indivíduos” (Wilkinson *et al.*, 2016, p. 1). Tais princípios trazem uma série de outras diretrizes e recomendações para cada um dos itens que compõem a sigla, que hoje são uma referência mundial na área de gestão de dados gerados para todos os tipos de organizações que pretendem gerenciar seus dados. Outras organizações foram surgindo desde então, como: o *Research Data Alliance*<sup>7</sup> (RDA), a rede *GO FAIR*, o *Internet of FAIR Data & Services* (IFDS)<sup>8</sup> o FAIRsharing<sup>9</sup>, e o *European Open Science Cloud* (EOSC).

Borgman (2018); Sales e Sayão (2019) relataram que quando os conjuntos de dados eram pequenos, raramente surgiam problemas na gestão, mas com os conjuntos de dados cada vez maiores e o aumento das pesquisas colaborativas, tornaram-se mais sofisticadas as ferramentas necessárias para minera-los e combiná-los. Borgman (2018, p. 368) alerta que é uma responsabilidade das lideranças institucionais iniciar o processo da construção da política e chama a atenção para o seguinte problema:

Os dados assumem muitas formas, têm muitas origens e têm muitos usos. A propriedade dos dados raramente é clara, especialmente para dados de pesquisa, e os custos e mecanismos de administração são mal compreendidos. Embora os sejam

---

<sup>7</sup> O RDA é uma organização internacional sem fins lucrativos que promove a abertura, o compartilhamento e o uso de dados de pesquisa.

<sup>8</sup> O IFDS é uma visão de uma rede global de dados e serviços que sejam FAIR

<sup>9</sup> O FAIRsharing é um portal que fornece acesso a recursos para ajudar os pesquisadores a encontrar, acessar e reutilizar dados de pesquisa. É mantido pelo RDA.

difíceis de gerenciar e governar em qualquer instituição, as universidades enfrentam um conjunto particularmente complexo de responsabilidades e riscos.

Sob a perspectiva de um professor e pesquisador, o astrofísico João Evangelista Steiner<sup>10</sup> (2014) relatou que um telescópio gera/coleta trinta milhões de medidas em apenas vinte minutos. Isso, conforme o professor, é algo que não conseguem mais tratar com os métodos clássicos, pois os olhos e o cérebro tornaram-se inúteis para extrair todas as informações. Logo, foi necessário o desenvolvimento de novas ferramentas e softwares para analisar a grande quantidade de dados. Entretanto, mesmo assim, extraem a informação de que precisam e descartam em torno de 95% dos dados. Por fim, Steiner afirma que muito esforço e dinheiro foram empregados e que esses dados poderiam ser reaproveitados por outros (Steiner, 2014). Em 2014, ano em que o pesquisador foi entrevistado, já haviam instituições promovendo a GDP, mas as IES do Brasil estavam longe de qualquer iniciativa.

O relato e as considerações do pesquisador descrevem muito bem os desafios que a *e-science* provoca com o dilúvio de dados. Como alerta Borgman (2018), as preocupações com a administração desses dados se tornam mais urgentes. O reaproveitamento, citado por Steiner (2014) pode ser viabilizado com as ferramentas disponíveis, hoje, pelo emprego dos Princípios FAIR e documentação formal.

Costa (2017, p. 2.019), no contexto de sua pesquisa, descreve que “Nas palavras de um pesquisador entrevistado “as universidades só vão se mexer quando uma instância superior disser, faça”. Conclui que as experiências no exterior revelam que o sucesso da implementação de política para a GDP dependerá das agências de fomento.

Borgman (2018, p. 368) aponta que:

A instituição deve tomar sua parte na questão e a principal forma de fazê-lo, é por meio de uma política eficiente. As universidades são guardiãs de grandes quantidades de dados. Esses dados oferecem muitas novas oportunidades para pesquisa, ensino, administração, parcerias e planejamento estratégico. Os dados assumem muitas formas, têm muitas origens e têm muitos usos. A propriedade dos dados raramente é clara, especialmente para dados de pesquisa, e os custos e mecanismos de administração são mal compreendidos. Embora os dados sejam difíceis de gerenciar e governar em qualquer instituição, as universidades enfrentam um conjunto particularmente complexo de responsabilidades e riscos.

---

<sup>10</sup> João Evangelista Steiner (1950-2020) foi um importante astrofísico brasileiro, professor titular do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da USP.

De acordo com Shearer (2015, p. 4), uma função importante da política é que “obriga os pesquisadores a pensarem sobre como eles administram seus dados antes do projeto, um requisito essencial para boas práticas de gestão de dados”. Entretanto, a política não pode ser encarada apenas como um meio para obrigar a GDP, mas também para garantir que esses insumos colaborem com a evolução do conhecimento e com o desenvolvimento socioeconômico (Singh; Monu; Dhingra, 2018). O levantamento dos elementos essenciais para compor o modelo de política universitária precisa ser ponderado de modo que possa proteger seus interesses, mas também dialogar com as políticas de outras instituições com as quais podem se relacionar (Shearer, 2015; Borgman; Bourne, 2021).

Em resumo, as políticas para gestão e compartilhamento de dados de pesquisa são imperativas para promover o Acesso Aberto, submissão de todos os membros às normas (tornando mandatória), padronizar os procedimentos, promover a transparência, estabelecer responsabilidades, garantir a preservação em todo o ciclo de vida dos dados, garantir a integridade e qualidade dos dados (Dijk; Doorn, 2014).

Em perspectivas diferentes quanto ao reuso de dados, Borgman e Bourne (2021, p. 4) relatam um polêmico incidente em editorial da revista *New England Journal of Medicine*: o texto “*The Data Parasite*” - o parasita de dados - referindo-se a pesquisadores que reusam os dados sem contribuir de volta e ainda sugeriram que “a pesquisa secundária deveria ser limitada àquela envolvendo colaboração direta com pesquisadores primários” (The New England Journal Of Medicine, 2016). Em resposta, Borgman e Bourne (2021) relataram que um grupo de pesquisadores criaram o prêmio *The Research Parasite Awards: Celebrating Rigorous Secondary Data Analysis*, para reconhecer o trabalho de pesquisadores que geraram novas hipóteses a partir de dados existentes, importantes para o processo da ciência.

Os primeiros passos do Brasil em direção à GDP já foram dados, mas ainda estão longe do ideal. Enfatiza-se também, a questão ética em acompanhar as tendências mundiais e atuar de forma recíproca. O acesso e reuso de dados produzidos em instituições estrangeiras, sem contribuir ou mesmo abrir os dados, mas sem os mesmos padrões de qualidade e colaboração, é problemática. Nesse cenário, o Brasil poderia ser visto como um "parasita de dados", literalmente. Sem uma estrutura equitativa e práticas diplomáticas, poderíamos nos isolar na busca por colaboração global em benefício da pesquisa e da sociedade. Logo, deve-se assumir a responsabilidade e contribuir ativamente para a colaboração e benefício mútuo na comunidade científica (Sayão; Sales, 2015, p. 53). Assim, corroborando com a frase

de Costa e Cunha (2019, p. 298) “fica evidente que o Brasil carece de um posicionamento uniformizado para todas as suas instituições de pesquisa e para todos os seus pesquisadores”.

## 2.5 DOCUMENTOS LEGAIS PARA ORIENTAR A GESTÃO DE DADOS DE PESQUISA

Não é possível organizar e padronizar a GDP em uma instituição sem a regulação formal que proteja os interesses da instituição e de seus *stakeholders* internos (Jones; Pryor; Whyte, 2013; Singh; Mony; Dhingra, 2018). Essa afirmativa é compartilhada por Sayão e Sales (2013) que atentam para a necessidade de um compromisso institucional para a condução da pesquisa de forma sustentável.

O dever no cumprimento da legislação governamental que abarca, direta ou indiretamente, a gestão de dados deveria ser suficiente para que as IES promovessem as práticas pertinentes, embora simplesmente obedecer às leis não seja suficiente. É preciso que a IES esteja preparada com uma documentação construída para atender os objetivos da GDP, e os seus interesses. Isso evitará que a gestão de dados chegue ao caos, transformando-se em um problema, ao invés de uma solução. A Fiocruz, por exemplo, para elaborar sua política, promoveu um diálogo com a comunidade científica e com todos os profissionais da instituição por meio de visitas técnicas às unidades, Câmaras Técnicas, Fórum das Unidades Regionais, seminários, grupos focais e consulta pública ao Termo de Referência que sistematizou uma série de Princípios e Diretrizes que dão orientações referentes aos dados da pesquisa científica (FIOCRUZ, 2020).

O Brasil possui legislação que estabelece normativas consistentes para orientar a gestão dos dados para o acesso e proteção. O Decreto 8.777, publicado em 2016, institui a Política Nacional de Dados Abertos do Poder Executivo Federal (BRASIL, 2016). Sendo as IES subordinadas ao Poder Executivo, obviamente se submetem, também, às disposições do documento. Nas Disposições Gerais, Art. 2º e 3º apresentam definições, princípios e diretrizes claras e objetivas para todos os setores abrangidos pelos decretos, sendo:

Art. 2º Para os fins deste Decreto, entende-se por:

I - dado - sequência de símbolos ou valores, representados em qualquer meio, produzidos como resultado de um processo natural ou artificial;



II - dado acessível ao público - qualquer dado gerado ou acumulado pelo Governo que não esteja sob sigilo ou sob restrição de acesso nos termos da Lei nº 12.527, de 18 de novembro de 2011 ;

III - dados abertos - dados acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento, limitando-se a creditar a autoria ou a fonte;

IV - formato aberto - formato de arquivo não proprietário, cuja especificação esteja documentada publicamente e seja de livre conhecimento e implementação, livre de patentes ou qualquer outra restrição legal quanto à sua utilização; e

V - Plano de Dados Abertos - documento orientador para as ações de implementação e promoção de abertura de dados de cada órgão ou entidade da administração pública federal, obedecendo os padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações.

Art. 3º A Política de Dados Abertos do Poder Executivo federal será regida pelos seguintes princípios e diretrizes:

I - observância da publicidade das bases de dados como preceito geral e do sigilo como exceção;

II - garantia de acesso irrestrito às bases de dados, as quais devem ser legíveis por máquina e estar disponíveis em formato aberto;

III - descrição das bases de dados, com informação suficiente para a compreensão de eventuais ressalvas quanto à sua qualidade e integridade;

IV - permissão irrestrita de reuso das bases de dados publicadas em formato aberto;

V - completude e interoperabilidade das bases de dados, as quais devem ser disponibilizadas em sua forma primária, com o maior grau de granularidade possível, ou referenciar as bases primárias, quando disponibilizadas de forma agregada;

VI - atualização periódica, de forma a garantir a perenidade dos dados, a padronização de estruturas de informação e o valor dos dados à sociedade e atender às necessidades de seus usuários; e

VII - designação clara de responsável pela publicação, atualização, evolução e manutenção de cada base de dados aberta, incluída a prestação de assistência quanto ao uso de dados (BRASIL, 2016).

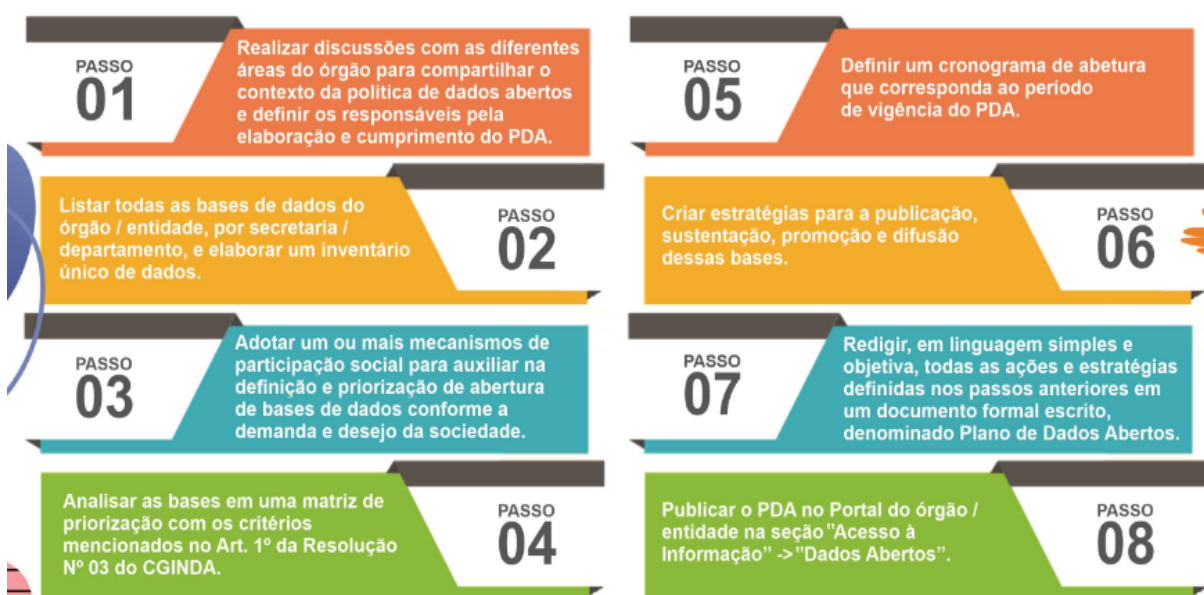
Os conceitos de 'Dado' apresentados no Art. 2º e de 'Dados de Pesquisa' da OCDE (2007) permitem o reconhecimento e a validação pela similaridade, embora, independentemente do conceito da OCDE, os dados de pesquisa estejam de acordo com a definição do Decreto.

O Decreto 8.777 também prevê que cada órgão deve apresentar Planos de Dados Abertos (PDA), no qual devem-se especificar os papéis e as responsabilidades das unidades envolvidas e designar um responsável, conforme o Art. 40, da LAI (BRASIL, 2011, grifo nosso):

[...] o órgão/entidade realize discussões com as diferentes áreas do órgão/entidade para compartilhar o contexto da Política de Dados Abertos e definir responsáveis pela elaboração e cumprimento do **PDA**. Um Grupo de Trabalho – GT pode ser criado, se o órgão desejar, para melhor conduzir as atividades.

O PDA, conforme descrito no Decreto, é um documento orientador das ações voltadas para promover a abertura de dados em cada órgão ou entidade da administração pública federal, seguindo alguns padrões para tornar as informações compreensíveis e reutilizáveis. A Figura 7 ilustra a dinâmica para iniciar os processos para a criação do PDA, conforme o Decreto 8.777:

Figura 7 – Etapas para a construção de um PDA



Fonte: BRASIL, [2017].

O PDA é um instrumento que destaca as estratégias e ações positivas para gerenciar e abrir os dados gerados/coletados pelo poder público. No entanto, é importante ressaltar que embora tenham alguns propósitos em comum, a estrutura do PDA não possui os mesmos elementos de um PGD para pesquisa científica, por esta razão não serve de referência, em sua completude. O PDA é obrigatório para todos os órgãos e entidades da administração pública federal direta, autárquica e fundacional, enquanto os PGD, voltados exclusivamente para dados gerados/coletados em IES públicas, vem sendo praticamente adotado de forma “voluntária”, por solicitação de outra organização com a qual possui algum envolvimento temporário.

Outro documento que deve ser considerado quanto ao tratamento dos dados é a Lei nº 12.527, de 18 de novembro de 2011 (BRASIL, 2011), a Lei de Acesso à Informação (LAI). O documento regula os procedimentos dos órgãos públicos para garantir o acesso à informação,

incluindo aquelas relacionadas à proteção de dados. A LAI (BRASIL, 2011, *passim*, grifo nosso) estabelece critérios para divulgação e sigilo, destacando-se:

- a) Art. 7, Parágrafo 1º: O acesso à informação previsto no caput não **compreende as informações referentes a projetos de pesquisa e desenvolvimento científicos ou tecnológicos cujo sigilo seja imprescindível à segurança da sociedade e do Estado**;
- b) Art. 7, Parágrafo 2º: Quando não for autorizado **acesso integral à informação** por ser ela parcialmente sigilosa, é assegurado o **acesso à parte não sigilosa** por meio de certidão, extrato ou cópia com **ocultação da parte sob sigilo**;
- c) Art. 23, Alínea VI: “prejudicar ou causar risco a **projetos de pesquisa e desenvolvimento científico ou tecnológico**, assim como a sistemas, bens, instalações ou áreas de interesse estratégico nacional”;
- d) Art. 31, Alínea II: “à realização de estatísticas e **pesquisas científicas** de evidente interesse público ou geral, previstos em lei, sendo **vedada a identificação da pessoa** a que as informações se referirem”.

No processo de gerenciamento e compartilhamento dos dados de pesquisa, as disposições da Lei nº 13.709, de 14 de agosto de 2018, Lei Geral de Proteção de Dados Pessoais (LGPD) (BRASIL, 2018) são imprescindíveis para regular o tratamento de dados pessoais e sensíveis de pessoa natural ou jurídica de direito público ou privado. Quando a pesquisa for para fins científicos, diversos requisitos devem ser atendidos pelas entidades que são autorizadas a gerar/coletar dados pessoais. Citando os pontos mais relevantes para esta pesquisa, o uso de dados pessoais deverá seguir as seguintes determinações (BRASIL, 2011, *passim*, grifo nosso):

Capítulo I: Disposições Preliminares:

XVIII - **órgão de pesquisa**: órgão ou entidade da administração pública direta ou indireta ou pessoa jurídica de direito privado sem fins lucrativos legalmente constituída sob as leis brasileiras, com sede e foro no País, que inclua em sua missão institucional ou em seu **objetivo social ou estatutário a pesquisa básica ou aplicada de caráter histórico, científico, tecnológico ou estatístico**.

Capítulo II: Do Tratamento De Dados Pessoais, Seção I: Dos Requisitos para o Tratamento de Dados Pessoais:

Art. 7º O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses:

IV - para a **realização de estudos por órgão de pesquisa**, garantida, sempre que possível, a **anonimização** dos dados pessoais;

Seção II: Do Tratamento de Dados Pessoais Sensíveis

Art. 11. O tratamento de dados pessoais sensíveis somente poderá ocorrer nas seguintes hipóteses:

c) realização de **estudos por órgão de pesquisa**, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis;

Art. 13. Na realização de estudos em saúde pública, **os órgãos de pesquisa** poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do

órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico e que incluam, sempre que possível, a **anonimização** ou **pseudonimização** dos dados, bem como considerem os devidos padrões éticos relacionados a estudos e pesquisas.

§ 1º A divulgação dos resultados ou de qualquer excerto do **estudo ou da pesquisa** de que trata o caput deste artigo em nenhuma hipótese poderá revelar dados pessoais.

§ 2º O **órgão de pesquisa será o responsável pela segurança da informação** prevista no caput deste artigo, não permitida, em circunstância alguma, a transferência dos dados a terceiro.

#### Seção IV: Do Término do Tratamento de Dados

Art. 16. Os **dados pessoais serão eliminados** após o término de seu tratamento, no âmbito e nos limites técnicos das atividades, **autorizada a conservação** para as seguintes finalidades:

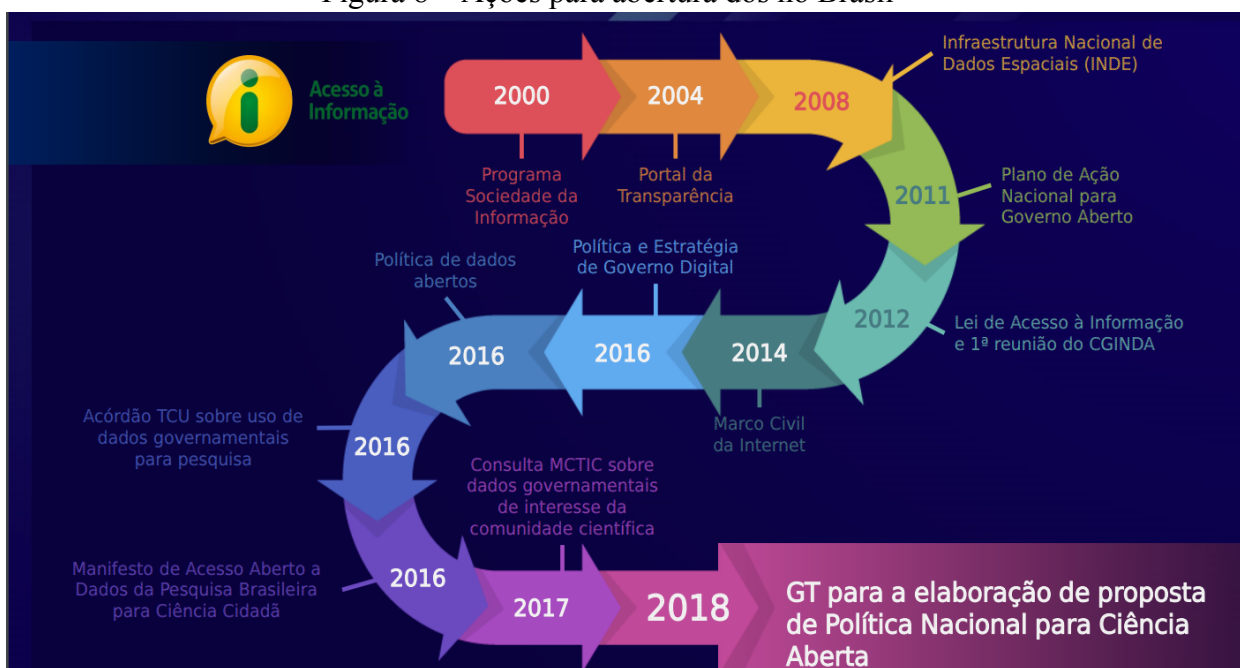
II - estudo **por órgão de pesquisa**, garantida, sempre que possível, a anonimização dos dados pessoais;

#### Capítulo X: Disposições Finais e Transitórias:

Art. 62. A autoridade nacional e o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), no âmbito de suas competências, **editarão regulamentos específicos para o acesso a dados** tratados pela União para o cumprimento do disposto no § 2º do art. 9º da Lei nº 9.394, de 20 de dezembro de 1996 (Lei de Diretrizes e Bases da Educação Nacional), e aos referentes ao Sistema Nacional de Avaliação da Educação Superior (Sinaes), de que trata a Lei nº 10.861, de 14 de abril de 2004 .

Ressalta-se que o Art. 12, parágrafo 4º da LGPD indica que os dados devem ser “mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico e que incluam, sempre que possível, a anonimização ou pseudonimização”. Outro ponto importante é a informação sobre a edição de “regulamentos específicos para o acesso a dados tratados pela União”, sendo atribuída à autoridade nacional e ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A Figura 8 ilustra os caminhos percorridos pela administração pública brasileira para promover o acesso aos dados gerados/coletados ou sob tutela federada.

Figura 8 – Ações para abertura dos no Brasil



Fonte: COSTA, 2018 *apud* COSTA; CUNHA, 2019, p. 290.

A Lei nº 13.243, de 2016, e a Lei da Inovação (BRASIL, 2016), dispõe sobre estímulos ao desenvolvimento científico, à pesquisa, à capacitação científica e tecnológica e à inovação. No Art. 9º existe a previsão de que as Instituições Científica, Tecnológica e de Inovação (ICTs) podem “celebrar acordos de parceria com instituições públicas e privadas para realização de atividades conjuntas de **pesquisa científica e tecnológica** e de desenvolvimento de tecnologia, produto, serviço ou processo” (BRASIL, 2016, grifo nosso). No Parágrafo 3º, é informado que a **propriedade intelectual** e a participação nos resultados **serão asseguradas às partes contratantes**, nos termos do contrato, podendo a ICT ceder ao parceiro privado a totalidade dos direitos de propriedade intelectual mediante compensação financeira ou não financeira, desde que economicamente mensurável. No Art. 5º, Parágrafo 1º “A propriedade intelectual sobre os resultados obtidos pertencerá à empresa, na forma da legislação vigente e de seus atos constitutivos” (BRASIL, 2016, grifo nosso). Não foram encontrados, durante a realização desta pesquisa, estudos sobre relatos de experiências para obter mais esclarecimentos sobre a prática no contexto da lei.

Criada pela Lei nº 14.129, de 29 de março de 2021, a Estratégia Nacional de Governo Digital (E-gov), ou Lei do Governo Digital, dispõe sobre princípios, regras e instrumentos para o aumento da eficiência da administração pública, especialmente por meio da

desburocratização, da inovação, da transformação digital e da participação do cidadão. A implantação dos serviços se encaixam nas demandas de abertura dos dados de pesquisa, principalmente quanto ao planejamento. Conforme especificado no Cap. VIII, “§ 6º Compete a cada ente federado disponibilizar as informações dos serviços prestados [...]”. O Art. 29 destaca que “Os dados disponibilizados pelos prestadores de serviços públicos, bem como qualquer informação de transparência ativa, são de livre utilização pela sociedade [...]”. A estratégia prevê que os servidores públicos devem adotar novas tecnologias e processos, com o planejamento traçando os objetivos, metas e prazos; capacitação para o uso das tecnologias. A exemplo das leis mais recentes citadas nesta subseção, os dados devem ser “[...] acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou tratamento por qualquer pessoa, física ou jurídica” (BRASIL, 2021).

Além dos suportes legais mencionados, um documento complementar que serve como base e regulamentação para ações no âmbito do Poder Executivo Federal é o Decreto nº 1.171, de 22 de junho de 1994, que institui o Código de Ética Profissional do Servidor Público Civil do Poder Executivo Federal (BRASIL, 1994). Esse decreto abrange, também, aos pesquisadores e demais servidores públicos das IES.

Por fim o texto “A Recomendação da UNESCO sobre Ciência Aberta”, lançado em 2021 apresentou uma série de prioridades que os estados-membros devem promover, destacando-se: a cultura da ciência aberta a compreensão de seus benefícios e desafios; ambiente político favorável; infraestrutura e serviços; treinamento, educação, alfabetização digital e capacitação; abordagens inovadoras em diferentes fases do processo científico; cooperação internacional e multissetorial; e outras recomendações. O documento apresenta um tópico específico para os dados de pesquisa publicados de forma aberta, a sua gestão e os princípios FAIR. O Brasil, como estado-membro da Unesco, assinou o documento e deve honrar seu compromisso.

Embora toda a legislação elencada ofereça recomendações, instruções e obrigações para as instituições públicas, é necessário mais do que isso. Precisa-se de documentos de abrangência nacional que estabeleçam boas práticas de pesquisa e métodos de gestão. É preciso que órgãos públicos especializados em ciência, tecnologia e inovação, como o IBICT, assumam sua tradicional liderança na viabilização de softwares, suporte, ferramentas de gestão, grupos de discussão, comitês nacionais e outras iniciativas para promover a

interoperabilidade e o intercâmbio de práticas em nível nacional para a promoção do reuso dos dados de pesquisa gerados/coletados nas IES do Brasil.

Por parte do governo federal, não basta apenas criar e impor o cumprimento das leis. É essencial fornecer contrapartidas investindo em infraestrutura física e tecnológica, capacitação e outras necessidades. É necessário fiscalizar.

Ao contrário dos países-membros da União Europeia, o Brasil não dispõe de programas de financiamento, e ainda não colocou em prática suas estratégias e metas a nível nacional para promover o Acesso Aberto aos dados de pesquisa. Mesmo diante dessas adversidades, as IES precisam encará-las e enfrentá-las como sempre fizeram e traçar estratégias para organizar as ações internamente. Destarte, é preciso que as lideranças institucionais sejam conscientizadas, por provocação interna ou pressão externa. Embora haja desafios, as IES têm a capacidade de superá-los, adaptando-se ao contexto brasileiro.

Elaborar e apresentar uma política institucional para a gestão de dados de pesquisa é um passo decisivo para implantar os serviços nas IES brasileiras como garantias de seu compromisso para promover a transparência, a integridade e o uso responsável dos dados, fortalecendo a confiança da comunidade acadêmica, dos órgãos governamentais e demais envolvidos. Costa e Cunha (2019, p. 304) alertam para o fato de que o Brasil precisa recuperar o atraso na gestão dos dados gerados por pesquisas financiadas com recursos públicos. Portanto, é preciso começar!

### 3 METODOLOGIA

Nesta seção, é descrito o método empregado para atingir os objetivos propostos. Maria Helena Michel (2015, p. 62) caracteriza métodos com uma “orientação necessária à realização da pesquisa social, sobretudo no que se refere à obtenção, processamento, análise e validação dos dados pertinentes à problemática que está sendo investigada”. A definição de Michel retrata a inclinação desta pesquisa como de natureza social, pois pretendeu obter dados pertinentes a uma contribuição compartilhada, por meio de coleta, análise e processamento, e validar os resultados alinhados com o problema investigado.

Para a abordagem do problema, utilizou-se o método **qualitativo**, pois foi fundamentada “a partir de análise feita de forma detalhada, abrangente, consistente e coerente, e na argumentação lógica das ideias” (Michel, 2015, p. 39) e **pesquisa aplicada**, pois teve sua finalidade voltada para a “aquisição de conhecimentos com vistas à aplicação numa situação específica” (GIL, 2022, p. 41), em que foi utilizada como característica, a exploração técnica, sistemática e exata. Para Michel (2015, p. 53), a pesquisa aplicada promove “a utilização dos conhecimentos e resultados adquiridos na pesquisa básica; volta-se mais para o aspecto utilitário da pesquisa, com vistas a torná-lo mais explícito”.

Quanto aos objetivos, é **exploratória**, pois utilizou-se dos meios **bibliográfico** e **documental**, com propósitos de identificar informações e subsídios para definição dos objetivos. Michel (2015, p. 4) aponta que esses procedimentos são realizados para levantamento de informações sobre o assunto, “recorrendo a documentos, de uma resposta a uma dúvida, uma lacuna de conhecimento” para “explicar um problema a partir de referências teóricas publicadas em documentos, dispensando a elaboração de hipóteses”. Gil (2022, p. 41) define que a pesquisa exploratória “tem como propósito proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito” e complementa informando que a coleta de dados realizada por meio de levantamento bibliográfico e análise de exemplos é necessária para a compreensão e viabilização de familiaridade com o problema, tornando-o mais explícito (Selltiz *et al.* *apud* GIL, 2022, p. 41).

Outra técnica aplicada foi a **análise de conteúdo**. Conforme Bardin (1977), a análise de conteúdo é um conjunto de técnicas de análise das comunicações que visa obter por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens indicadores



que permitam a inferência de conhecimentos relativos às condições de produção/recepção dessas mensagens. A análise de conteúdo foi considerada mais adequada por ser flexível para ser aplicada em diferentes fontes de informação, como textos de sites com recomendações e/ou oferta de serviços, relatórios de organizações afins, legislação, declarações e códigos que se envolvem com o tema, entrevistas, artigos científicos, entre outros. Este tipo de análise permitiu explorar e interpretar as informações, buscando identificar elementos e abordagens padrões.

Para que fosse possível a pesquisa em instituições e organizações, foi necessário realizar um recorte, pois, como apontado por Gil (2022, p. 115) “De modo geral, os levantamentos abrangem um universo de elementos tão grande que se torna impossível considerá-los em sua totalidade”.

### 3.1 ETAPAS E DESCRIÇÃO DA PESQUISA

**a) Levantamento bibliográfico:** as fontes bibliográficas foram buscadas nas bases de dados: *Scientific Electronic Library Online* (SciELO); Portal de Periódicos CAPES; Google Acadêmico, OasisBr e biblioteca de livros digitais Minha Biblioteca. Para a recuperação de documentos, foram realizadas buscas combinadas utilizando os termos: *research data; open research data; research data management; sharing research data; management and sharing research data; research data reuse; data management plan; institutional policy(ies); Brazil; Netherlands*. As buscas foram realizadas nos campos: assunto, título, resumo e palavras-chave. O período de busca compreende os meses de março de 2022 e julho de 2023.

Para filtrar a recuperação de documentos buscados, foram aplicados os seguintes critérios para refinamento:

- a) relação e pertinência do assunto, conforme os objetivos desta pesquisa;
- b) revisão por pares;
- c) acesso aberto;
- d) tempo majoritariamente inferior a dez anos.

Essa abordagem permitiu refinar os resultados e obter documentos relevantes para a análise. Além disso, as citações e referências apresentadas nos documentos consultados constituíram-se como uma fonte para novas leituras, o que possibilita que novos horizontes se

abram durante o processo de investigação (Bryman, 2012). Como mencionado por Gil (2022, p. 64), em virtude da “ampla disseminação de materiais em formato eletrônico, a localização de fontes bibliográficas vem sendo feita, principalmente, por intermédio das bases de dados”.

A seleção da bibliografia preliminar contou com as referências apresentadas em aula no PPGCI IBICT/UFRJ e estudos do Grupo de Pesquisa: Biblioteconomia, Representação, Interoperabilidade, E-science e Tecnologia (BRIET).

**b) Análise de conteúdo:** devido ao grande número de documentos, foi necessário examinar e interpretar textos para compreensão das nuances do tema, como seus elementos padrões e elementos em processos constantes de desenvolvimento, analisar e adaptar processos. A seleção dos documentos envolveu a escolha dos documentos conforme a relevância e adequação aos objetivos desta pesquisa.

Foram excluídos os documentos que não abordaram os tópicos necessários, como a gestão de dados que não eram relativos aos dados de pesquisa; documentos com informações que já haviam sido tratadas em documentos anteriores e não apresentavam novas perspectivas; informações desatualizadas e textos com acesso restrito. Foram realizadas leituras de título, palavras-chave e resumo para verificar se o documento se alinha aos propósitos desta dissertação. A credibilidade das fontes foi avaliada considerando o número de citações verificadas no Google Acadêmico, levando em conta as autoridades reconhecidas internacionalmente em pesquisa na área. Alguns documentos recentes, com um baixo número de citações, por exemplo, abaixo de dez, foram consultados e incluídos na lista de referências devido ao fato de que o assunto está em evolução, pois foram elaborados por autores proeminentes na área. Além disso, também se levou em consideração a frequência de citação desses documentos por outros autores conhecidos como grandes pesquisadores na área.

Após a seleção dos documentos, foi realizada a leitura do texto completo e foram selecionados os documentos que apresentavam informações sobre o histórico e a evolução da área; os que abordam os problemas e desafios enfrentados; perspectivas futuras; descrição de experiências; práticas de gestão de dados de pesquisa já estabilizadas e em desenvolvimento. Foram priorizados os documentos que abordaram a temática de políticas, conceitos, estruturas tecnológicas, ferramentas, serviços e envolvimento de *stakeholders* internos e externos, considerando que esses tópicos compõem o processo da GDP.

### 3.1.1 Mapeamento conceitual da área de Gestão de Dados de Pesquisa

Existe uma variedade de aspectos na geração/coleta de dados, metadados, armazenamento, arquivamento, compartilhamento, reuso e diversas outras questões. Com a proposição de uma política para GDP, é fundamental compreender os conceitos do campo.

Os conceitos são uma “reunião e compilação de enunciados verdadeiros a respeito de determinado objeto” (Dahlberg, 1978, p. 102).

Com o objetivo de apresentar conceitos relacionados à construção de uma política para GDP, foi realizada a **pesquisa bibliográfica** em artigos científicos, teses e livros; **pesquisa documental** em guias; normas, regulamentos, políticas para GDP, legislação, PGD amplamente adotados, como DMPtool e DMPonline; e **pesquisa exploratória** em instituições internacionais, como OCDE, DCC, Codata, *Fairsharing* e RDA, com diretrizes e princípios de abrangência globais; e nacionais, como Fiocruz, Fapesp, instituições de ensino superior e sites governamentais; o Descritores em Ciências da Saúde (DeCS) também contribuiu de forma significativa para a realização desta análise.

Foram realizadas pesquisas exploratórias em políticas institucionais para GDP em universidades estrangeiras, onde foram extraídos e analisados os elementos e abordagens para compreender as funcionalidades e finalidades, identificar as semelhanças e diferenças; as abordagens e padrões recorrentes. Foram realizadas comparações e abstrações para alinhar os elementos com as disposições nacionais sobre a GDP, verificando se é consistente, abrangente e se representa de forma precisa o objeto que está sendo modelado. Os modelos são “ideias estruturadas, conectando argumentos, que apresentam algum poder explanatório” (Skilling, 1964 *apud* Sayão, 2001, p. 85). Nesta mesma linha, sobre a construção de modelos, Cunha, Amaral e Dantas (2015 p. 88) apontam o seguinte:

- a) delimitar, definir e relacionar os distintos aspectos do processo de busca da informação;
- b) estabelecer as diferenças entre os distintos estudos centrados nos usuários do ponto de vista dos seus objetivos e da aplicação dos seus resultados;
- c) sistematizar o nível de abstração e propósito da pesquisa;
- d) usar uma base teórica como marco de referência para interpretar a conduta que surge em consequência de uma necessidade de informação;
- e) estabelecer uma relação mútua entre teoria e pesquisa empírica, de tal maneira que a teoria guia a pesquisa e que esta alimenta a teoria.

Com esses critérios, foi realizado o levantamento de informações consistentes e foram embasados os conceitos necessários para a compreensão e construção de uma política de gestão e compartilhamento de dados de pesquisa. Foram relacionadas as amostragens, suas definições e perspectivas nas fontes consultadas. Dahlberg (1978, p. 104), sobre as relações entre os conceitos, informa que:

Se dois conceitos diferentes possuem características idênticas e um deles possui uma característica a mais do que o outro, então entre eles se estabelece a relação hierárquica ou relação de gênero e espécie. Pode-se então falar de conceitos mais amplos ou mais restritos. Pode-se também falar de conceito superior e inferior. O conceito superior é o mais genérico e o inferior é o mais específico [...] Quando a comparação entre as características dos conceitos mostra que dois conceitos diferentes possuem uma ou duas características em comum, então há que se falar de relações entre tais conceitos.

As considerações de Dahlberg forneceram uma elucidativa orientação quanto ao reconhecimento de conceitos já estabelecidos e investigação daqueles que podem ser interpretados equivocadamente. Nesse sentido, foi realizada uma análise comparativa e qualitativa dos elementos, considerando os seguintes critérios: a presença de definições idênticas ou de maior similaridade; conceitos distintos, mas com o mesmo sentido, sendo aplicado neste texto, o que apresentou maior incidência; e conceitos iguais, mas aplicados a situações distintas, sendo selecionados os mais recorrentes nas fontes consultadas.

Para esta análise, foi elaborado um quadro contendo as colunas: conceito, definição, procedimentos e função e fontes. Em alguns elementos, a definição e a função são indissociáveis; há também, os casos que apresentam definições e procedimentos extensos. As informações foram estruturadas conforme o Quadro 1, que pode ser consultado na íntegra no Apêndice C.

Quadro 1 – Estrutura para análise de conceitos

TERMOS	CONCEITOS	FONTES
Acesso aberto	Acordo em que autores e detentores de direitos autorais garantem a t	Decs - Descritores da Área
Administrador de dados	Pessoa responsável por manter a qualidade, integridade e acordos de	LCRDM (201-)
Administrador de dados	Alguém responsável pelo cuidado profissional e cuidadoso dos c	Universidade de Utrecht (2
Administrador de dados	Os administradores de dados ajudam os pesquisadores em seus institu	Universidade de Amsterda
Administrador de dados	Funcionário, geralmente um pesquisador, responsável pelo gerenciam	Universidade de Tilburg (2
Administrador de dados (	Alguém que lida com um ou mais aspectos do gerenciamento de dad	Universidade de Utrecht (2
Administrador de dados (	Uma pessoa responsável pelo gerenciamento de objetos de dados, inc	LCRDM (201-)
Anonimização	Desvinculação irreversível de informações de identificação de dados	LCRDM (201-)
Anonimização	Processo pelo qual passam dados que contêm informações pessoais p	Sayão e Sales (2015)
Anonimização	Utilização de meios técnicos razoáveis e disponíveis no momento do	BRASIL. LGPD (2018)
Armazenamento	Ação de dispor ou guardar documentos nos lugares predeterminados	Cunha e Cavalcanti (2008)
Armazenamento	A colocação e guarda de coisas em um lugar especial para uso no futu	Dicionário Cambridge
Armazenamento	Serviços e funções usados para o armazenamento e recuperação de pe	Data Curation Centre

Fonte: Elaborado pela autora, 2023.

### 3.2 PREVISÕES LEGAIS NACIONAIS PARA A GESTÃO DE DADOS

É fato que há previsão nas leis federais brasileiras que torna obrigatório o compartilhamento e a proteção das informações e dos dados gerados/coletados pelos setores públicos. Essa prática tem sido adotada tanto pelos órgãos governamentais quanto pelas instituições de pesquisa. No entanto, as IES públicas brasileiras não compreenderam, por desconhecimento ou omissão, que os dados de pesquisa também estão sujeitos à força da legislação.

A seção de Discussão e Resultados apresentará pontos específicos sobre os dados de pesquisa. Também discutirá as implicações legais para as IES que tornam obrigatória a responsabilidade e a gestão dos dados; no contexto desta dissertação, os dados de pesquisa.

### 3.3 PANORAMA DA GESTÃO DE DADOS DE PESQUISA NO BRASIL

Para o segundo objetivo específico, que tencionava extrair os elementos de normativas brasileiras sobre GDP, foram realizadas pesquisas exploratórias nos sites de agências de fomento, institutos de pesquisa e instituições de ensino superior. Foram buscadas as

normativas relacionadas à gestão e ao compartilhamento de dados de pesquisa, como resoluções, portarias e diretrizes de órgãos reguladores e financiadores de pesquisa no Brasil. Por meio de análise de conteúdo, foram identificados os elementos e, nesta etapa, também foram extraídos os conceitos com suas definições, procedimentos e função.

A busca foi realizada nos sites das instituições e em ferramentas digitais, empregando o nome da instituição, combinando com as expressões: “plano de gestão de dados”, “dados de pesquisa”, “compartilhamento e reuso de dados”, “gestão de dados de pesquisa”, “políticas para gestão de dados de pesquisa” “normas”, “diretrizes” “acesso aberto”. Essa ação permitiu visualizar o panorama da GDP no Brasil sob diferentes perspectivas. Esse levantamento é “uma fase da pesquisa, cujo objetivo é auxiliar na definição de objetivos e levantar informações sobre o assunto/objeto de estudo” (Michel, 2015, p. 47). A busca por políticas institucionais para IES estrangeiras foi realizada da mesma forma, com as palavras no idioma inglês.

### **3.3.1 Agências de fomento/Financiadores**

Sendo as instituições de fomento e financiadores de pesquisa os grandes propulsores da GDP, buscou-se compreender sua relação com as instituições que realizam pesquisas científicas no Brasil.

Não foi localizada uma fonte de informação que apresentasse critérios para a seleção de financiadores de pesquisas no âmbito nacional. Por meio de uma pesquisa exploratória, foi localizado no site da USP uma investigação conduzida por Dudziak (2020) que, como base em dados obtidos na Plataforma *InCites*, identificou-se as agências que mais financiavam pesquisa no Brasil, sendo: a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Financiadora de Estudos e Projetos (FINEP), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Banco Nacional de Desenvolvimento (BNDES). Com base nesse levantamento, os sites

dessas organizações foram consultados para a busca de documentação ou informações sobre assuntos relacionados a PGD e GDP.

### **3.3.2 Institutos de pesquisa nacionais**

No Brasil, há diversos institutos públicos e privados reconhecidos nacional e internacionalmente pela qualidade de suas pesquisas científicas. Entretanto, não foi localizada uma fonte específica para orientar na seleção desta dissertação. Assim, optou-se por consultar os sites de institutos citados na literatura nacional, como Fiocruz, o Instituto Butantan, Instituto de Pesquisas Jardim Botânico (JBRJ), Instituto Nacional de Pesquisas Espaciais (INPE), Instituto Nacional de Pesquisas da Amazônia (INPA), Programa de Pesquisa em Biodiversidade no Brasil (PPBio), Embrapa e Fundação Espírito-Santense de Tecnologia (FEST). Essa abordagem buscou garantir a abrangência e a representatividade das informações coletadas em institutos que não têm vínculos administrativos com as IES.

Considerando as pesquisas colaborativas entre pesquisadores de IES e esses institutos, o compartilhamento de infraestruturas e financiamentos, para encontrar normativas, foram aplicadas buscas nos sites das referidas instituições e também por meio de buscas em ferramentas digitais, combinando os termos: nome da instituição e os termos indicados na subseção 3.3.

### **3.3.3 Instituições de Ensino Superior Nacionais**

Tendo em vista o grande número de IES em território nacional, como método de seleção, tomou-se como referência o Índice Geral de Cursos (IGC) do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), referente ao ano de 2019 e atualizado em 2021. Foram selecionadas as IES identificadas como universidades, centros universitários, faculdades e institutos federais, que obtiveram notas 6 e 7 nos cursos de pós-graduação *stricto sensu*, conforme os resultados da avaliação quadrienal da Capes (2017-2020). Buscou-se nos sites de todas as IES recuperadas por este método, a existência de

políticas, além de documentos relativos à GDP. O resultado das consultas e análises foi apresentado em quadros descrevendo o conteúdo de cada documento recuperado.

Essa análise possibilitou uma visão do panorama e a identificação de lacunas quanto à promoção dos dados de pesquisa e à gestão. Além disso, reafirmou de maneira enfática a necessidade de implementação de uma política institucional para que essas questões sejam apresentadas e adotadas de forma estruturada e abrangente.

### 3.4 BUSCANDO ELEMENTOS EM IES ESTRANGEIRAS

Para enriquecer o campo por meio da consulta em modelos internacionais, foram pesquisados em documentos estrangeiros em países que estão em estágio mais avançado na adoção dos princípios da Ciência Aberta. Considerando o grande número de nações, optou-se por realizar um recorte, e consultar apenas as políticas para GDP das universidades Holanda, levando em consideração, entre outros fatores, a existência de documentos de abrangência nacional quanto às práticas de ensino e pesquisa. A escolha se justifica também pelo destaque que recebeu no livro Livro Verde – Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional”, publicado pela Fiocruz, como um dos países proeminentes na GDP. A publicação cita que a Rede GO FAIR começou na Holanda, com estratégias internacionais voltadas para o desenvolvimento de serviços e dados FAIR na Europa e no mundo (Santos, P.; Almeida, B.; Henning, P. (org.), 2017, p. 23). Clínio *et al.* (2017, p. 12) também reconheceu que a Holanda se destaca pelo seu papel na promoção dos princípios FAIR, bem como pela sua atuação com a *European Open Science Cloud* (EOSC).

A Holanda possui uma infraestrutura nacional de dados de pesquisa integrada, onde diversas organizações atuam em conjunto para promover a Ciência Aberta. Em um primeiro momento, destacam-se a *Royal Netherlands Academy of Arts and Sciences* (KNAW) “Academia Real Holandesa de Artes e Ciências” e o *Dutch Research Council* (NWO) “Organização Holandesa para Pesquisa Científica”, que são grandes financiadores da pesquisa nacional. Em 2004, a KNAW e o NWO organizaram o primeiro *Netherlands Code of Conduct for Research Integrity* (NCCRI) “Código de Conduta para Integridade em Pesquisa da Holanda”. Na versão de 2018, foi incluído no texto a necessidade da gestão dos dados em



todas as instituições de pesquisa. Também em 2004, foi lançado o portal *National Academic Research and Collaboration Information System* (NARCIS), “Sistema Nacional de Informações sobre Pesquisa Acadêmica e Colaborações”, para arquivar as publicações de resultados de pesquisa e seus dados. Em julho de 2023, o NARCIS foi substituído pelo portal *Netherlands Search Results Portal* (PURE) “Portal de Resultados de Pesquisa da Holanda” (HOLANDA, 2023). O PURE foi construído para oferecer mais recursos e funcionalidades integrativas, reunindo os resultados de pesquisa de diferentes RD e RI. O Portal foi desenvolvido em parceria entre o *Dutch University Libraries* “Bibliotecas das Universidades da Holanda” e a *Royal Library of the Netherlands* (UKB) “Biblioteca Real da Holanda”, com o financiamento do governo federal.

Todo o sistema de informação e comunicação científica do país recebe suporte do *University Computing Centers Foundation* (SURF) “Fundação de centros de computação universitária”, organização colaborativa para TI em educação e pesquisa, e o *Data Archiving and Networked Services* (DANS) “Arquivamento de Dados e Serviços em Rede”, que desde 2005 fornece suporte e infraestrutura aos pesquisadores para o arquivamento, preservação, compartilhamento e acesso aos dados de pesquisa. Por fim, enfatiza-se que vários programas, ferramentas e softwares utilizados no país receberam o financiamento do *Open Access Infrastructure for Research in Europe* (OpenAIRE) “Infraestrutura de Acesso Aberto para Pesquisa na Europa”, e do Programa Horizon 2020, da União Europeia.

Existem ainda outras iniciativas importantes, como *National Coordination Point Research Data Management* (LCRDM) “Ponto Nacional para Gestão de Dados de Pesquisa”, criado em 2016, que é mantido e gerenciado por especialistas em GDP para facilitar os serviços e apoiar políticas e solução. Também em 2016, foi lançado o *Code of Conduct for Social and Behavioral Sciences* (CCBS) “Código de Conduta para Ciências Sociais e Comportamentais”, atualizado em 2018 com diretrizes para a proteção, privacidade e reutilização de dados. Como membro da União Europeia, seguem as normativas do bloco, como o *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020* e do *European Code of Conduct for Research Integrity*, publicado pela ALLEA (2017, 2023). O Código de Conduta da Holanda está em conformidade com os termos do Código Europeu.

Em 2022, o *National Programme Open Science* (NPOS) “Programa Nacional de Ciência Aberta da Holanda” apresentou o plano nacional para tornar a ciência mais aberta, o

“*Open Science 2030 in The Netherlands*”, que visa aumentar a colaboração entre instituições de pesquisa, órgãos públicos, setor industrial e cidadão, com objetivos, metas e ações para alcançar cada tópico da agenda. Por fim, cita-se o fato de que a Holanda foi sede da conferência *Jointly Designing a Data Fairport*, em 2014, que discutiu práticas para a gestão de dados, e lançou as ideias dos Princípios FAIR.

As iniciativas citadas e inseridas na seção de Metodologia evidenciam a qualidade e o comprometimento das instituições holandesas na promoção da Ciência Aberta e no gerenciamento e compartilhamento dos dados de pesquisa, justificando assim sua escolha para compor o *corpus* da pesquisa. É importante enfatizar que todas as universidades selecionadas para análise são públicas e esses recursos desempenham um papel fundamental na criação de políticas institucionais efetivas para a gestão de dados de pesquisa. Vale ressaltar que não é possível listar todos os serviços, sistemas colaborativos e ferramentas disponíveis no país, dada à sua vasta gama de recursos na área.

### **3.4.1 Critérios para selecionar as IES holandesas a serem analisadas**

As instituições foram selecionadas com base nos resultados do Centro de Estudos de Ciência e Tecnologia (CWTS) da Universidade de Leiden, na Holanda. O CWTS é internacionalmente reconhecido pelos seus indicadores de desempenho e qualidade de IES e está incluído no Catálogo de Bases de Dados do Portal de Periódicos da Capes, o que evidencia sua relevância e qualidade.

A seleção das Instituições foi realizada com base no *ranking* de Impacto Científico, tendo sido aplicados alguns filtros:

- a) Período: 2017-2022;
- b) Área: todas as áreas do conhecimento;
- c) País: Holanda (*Netherlands*);
- d) Indicadores: "Impacto Científico";
- e) Classificação por: P (número de publicações no período de 2017-2022; publicações colaborativas).

Após a verificação de disponibilidade e acessibilidade das políticas institucionais para GDP de cada instituição, os arquivos foram baixados e organizados em uma pasta específica utilizando o serviço de nuvem do Google Drive.

### 3.5 ORGANIZAÇÃO DAS INFORMAÇÕES COLETADAS NAS INSTITUIÇÕES CONSULTADAS

Os dados coletados foram organizados em planilhas no editor Excel, com uma planilha dedicada exclusivamente a cada categoria de instituição:

- a) agências de fomento,
- b) institutos de pesquisa,
- c) IES nacionais,
- d) IES holandesas.

Com os resultados, foi realizada a **análise comparativa** dos elementos e abordagens entre cada categoria. De acordo com Marconi e Lakatos (2010, p. 89) “o método comparativo permite analisar o dado concreto, deduzindo do mesmo os elementos constantes, abstratos e gerais”.

Na primeira linha da planilha, foi inserido o nome da categoria institucional; na linha inferior foram apresentadas em Escala Likert, as opções: Sim, Não e Não informa. Em seguida, na primeira coluna, foram listados os nomes das instituições, e nas colunas seguintes, os elementos abordados nas políticas.

Nesta fase da pesquisa, a análise e o relacionamento entre políticas exigiram um esforço significativo. Apesar do número relativamente pequeno de instituições, onze no total, foi necessário realizar leituras minuciosas e exaustivas para examinar cada documento e identificar os elementos comuns e distintos. Para garantir maior precisão na coleta de dados, todas as informações foram registradas na planilha. Este processo envolveu a verificação cuidadosa da menção de cada elemento nas diversas políticas, exigindo idas e vindas nos documentos para confirmar os resultados e evitar possíveis omissões, pois os mesmos tipos de informações ora estavam explícitas nos textos, ora em pequenas letras em notas de rodapé, ou mesmo, truncados em um parágrafo. Logo, para que a análise do conteúdo fosse mais precisa,

o processo de leitura dos documentos foi repetido diversas vezes. Ao todo, foram criadas 61 colunas representando cada elemento extraído das políticas.

Quadro 2 – Modelo da planilha para organizar os elementos

<b>UNIVERSIDADES HOLANDESAS</b>				
<b>UNIVERSIDADE</b>	<b>Título</b>	<b>Ano</b>	<b>primeira Vers</b>	<b>PGD próprio</b>
Universidade de Utrecht	Política de gerenciamento	2019	2016	Sim
Universidade de Amsterdar	Gerenciamento de Dados	2019	2015	Sim
Universidade de Groningen	Política de dados de pes	2021	2015	Sim
Universidade de Leiden	Regulamento de gerenc	2021	2016	Sim
Universidade de Radboud	Política da universidade	2021	2013	Sim
Universidade de Erasmus d	Política de Gerenciame	2020		Não informa
Universidade Livre de Ams	Política de gerenciamento	2020		Não informa
Universidade Técnica de D	Política de estrutura d	2020	2018	Não informa
Universidade de Wageninge	Política de dados na WI			Não informa
Universidade de Twente	Política de gerenciamento	2018	2015	Não informa
Universidade de Tilburg	Regulamento de Gerenc	2020	2017	Sim

Fonte: Elaborado pela autora, 2023.

Após a análise, verificou-se que, como assunto novo e multifacetado, os textos não apresentavam somente normativas institucionais, mas também, orientações, sugestões e flexibilidade para realização de acordos. Posteriormente, as informações coletadas foram organizadas e categorizadas de maneira lógica, o que facilitou a realização desta etapa do estudo.

### 3.6 APRESENTAÇÃO DAS DIRETRIZES PARA UMA POLÍTICA PARA GDP

A apresentação de diretrizes para a elaboração de uma política institucional para as IES nacionais foi estruturada de forma detalhada, com alternativas de abordagens e fluxos de trabalho de cada elemento. A pesquisa revelou que há itens gerenciados de forma padronizada, como os prazos mínimos para que os dados permaneçam arquivados e questões que ainda carecem de normativas mais claras para organizar os serviços. Em alguns casos, as

informações foram explícitas, mas em outros foi necessário realizar abstrações, como no caso em que se indicava a escolha de um “repositório confiável” para o arquivamento dos dados.

O modelo proposto, conforme o objetivo geral desta pesquisa, foi estruturado com as etapas para as ações, responsáveis pelas decisões e alternativas para iniciar a construção do documento. Para lidar com situações em que as informações são explícitas, ambíguas ou mesmo omissas, a modelagem recorreu a métodos comparativos e de abstração, que se revelaram essenciais para visualizar os modelos como recursos metodológicos, conforme elucidações de Sayão (2001, p. 83): Os modelos são “representações simplificadas e inteligíveis do mundo, que permitem vislumbrar características essenciais de um domínio ou campo de estudo”. O autor descreve de forma reflexiva, que “um modelo é uma criação cultural, um “mentefato”, destinada a representar uma realidade, ou alguns dos seus aspectos, a fim de torná-los descritivo qualitativa e quantitativamente e, algumas vezes, observáveis” (Sayão, 2001, p. 83).

Para aderir ao processo da GDP, apresentamos os PGD desta dissertação nos Apêndices A e B, utilizando os modelos FioDMP e DMPtool conforme proposto no projeto de qualificação. O DMPtool não foi atualizado durante a pesquisa para permitir a identificação de erros caso o PGD não seja revisado.

Para sintetizar o percurso metodológico e seu alinhamento com os objetivos específicos da pesquisa, os resultados produzidos podem ser visualizados no Quadro 3.

Quadro 3 – Correlação: caminhos para os resultados

Objetivo	Método	Resultado
Conhecer os conceitos que se envolvem com dos dados de pesquisa e sua gestão	Pesquisa exploratória, bibliográfica e documental, método qualitativo e comparativo. Foram extraídos os termos e conceitos citados nas fontes e organizados sistematicamente em planilha para comparação e análise criteriosa das definições, indicado a fonte. Por meio de método qualitativo e análise de conteúdo, foram selecionados os conceitos conforme o contexto, coerência em glossários especializados e incidência nas fontes consultadas.	Obteve-se um conjunto de conceitos e termos padronizados que permitiu uma compreensão mais profunda dos conceitos-chave, sua aplicação e inter-relações, fornecendo uma base sólida para a análise e discussão no contexto da pesquisa, evitando ambiguidades e equívocos no processo da GDP.
Identificar normativas em organizações nacionais	Pesquisa exploratória e documental, método qualitativo e comparativo. Normativas de diferentes categorias de instituições nacionais foram buscadas para extrair seus elementos e conhecer as práticas de gestão no país. Os dados foram organizados em planilhas individualmente e, em seguida, realizou-se uma análise comparativa dos elementos de acordo com o tipo de organização.	Os elementos foram identificados e permitiu uma visão abrangente das normativas em organizações nacionais, fornecendo subsídios para a compreensão do cenário regulatório e sua aplicação nas instituições pesquisadas.
Verificar normativas em IES estrangeira	Pesquisa exploratória, bibliográfica e documental e método qualitativo e comparativo. Visando maior coerência e objetividade, optou-se por selecionar as políticas de um único país, pois possuem normativas nacionais comuns. Os resultados foram organizados em uma planilha sistemática para facilitar a extração e visualização de cada elemento, num primeiro momento, isoladamente e posteriormente, comparativamente entre as IES.	Os resultados foram obtidos, permitindo a compreensão das práticas e abordagens utilizadas para a GDP no país selecionado. Foram obtidos resultados relevantes que contribuíram para a construção de ideias para o modelo nacional, conforme proposto nesta pesquisa.
Propor elementos para um modelo nacional	Com resultados alcançados nos objetivos específicos, realizou-se uma análise de conteúdo e identificou-se os elementos necessários para a construção de uma política institucional para GDP para as IES brasileiras. Os elementos foram organizados em blocos temáticos, que permitiram dissertar sobre aspectos relevantes. Os resultados obtidos proporcionaram uma base para a apresentação dos resultados e conclusões.	Durante o processo da pesquisa, adquiriu-se maior conhecimento e competência sobre o tema, o que permitiu que o conjunto de informações fosse aproveitado de forma qualificada. Os resultados foram apresentados de forma a facilitar o entendimento daqueles que consultarão o documento para a construção de uma política que aborde os itens indispensáveis e os que apresentam flexibilidade, que poderão ser abordados conforme interesses institucionais. Dessa forma, a pesquisa aplicada foi concretizada, atingindo o objetivo geral.

Fonte: Elaborado pela autora., 2023.

Na próxima seção, serão apresentados e discutidos os resultados obtidos a partir da aplicação dos procedimentos metodológicos descritos neste trabalho, buscando analisar de forma abrangente e detalhada as informações coletadas.

## 4 ANÁLISE DOS RESULTADOS

Nesta seção, apresentam-se os resultados obtidos na pesquisa sobre políticas para gestão e compartilhamento de dados de pesquisa. Ao longo das etapas, analisaram-se os conceitos envolvidos, identificaram-se as normativas e abordagens brasileiras e internacionais sobre o tema desta pesquisa e foram coletados elementos para proposição de elementos para um modelo nacional de política.

### 4.1 HORIZONTES PARA O COMPARTILHAMENTO DE DADOS NO BRASIL

A revisão de literatura desta pesquisa baseou-se nos documentos jurídicos e regulatórios que determinam diretrizes para a abertura e a proteção dos dados. Suas disposições abrangem tanto os dados de pesquisa quanto a responsabilidade das instituições para viabilizar a abertura e o reuso dos dados, quando aplicável, e garantir sua proteção, quando necessário.

Até mesmo a Lei mais importante do Brasil, a Constituição Federal (CF) (BRASIL, 1988), justifica o gerenciamento de dados. No cenário atual, as normativas da CF também abrangem as instituições de ensino superior, que geram dados, incluindo os de pesquisa. Como partes integrantes da administração pública, as IES são obrigadas a cumprir essas normativas. O Art. 37 aponta como princípios da administração pública federal: a Legalidade, a Impessoalidade, a Moralidade, a Publicidade e a Eficiência, conhecidos pelo acrônimo LIMPE. Relacionando esses princípios com a GDP, as IES devem:

- a) garantir que todos os dados sejam coletados, armazenados e processados de forma legal e ética, armazenados com segurança e processados conforme as leis de proteção de dados;
- b) tratar os dados de forma impessoal, garantindo que não serão usados para discriminar ou prejudicar nenhuma pessoa e grupo natural ou jurídico;
- c) responsabilizar-se quanto ao tratamento, de acordo com a moralidade, não os utilizando para fins ilegais, prejudiciais ou inescrupulosos;

- d) tornar os dados públicos, disponibilizando-os de forma aberta, a menos que exista previsão legal e ética para privados.
- e) tratar os dados de forma eficiente, em que a agilidade, a qualidade e a responsabilidade sejam garantidas durante todo o seu ciclo de vida.

Mesmo que na ocasião da promulgação da Constituição Federal (CF) os dados de pesquisa não fossem notáveis como atualmente, o LIMPE manifesta normativas similares aos princípios FAIR, como o armazenamento, a segurança, as restrições da publicidade e o ciclo de vida. O LIMPE, como princípios para orientar a administração pública a nível nacional, e o FAIR, como princípios para nortear o gerenciamento de dados de pesquisa a nível global, apresentam similaridades, apesar de serem diferentes em abrangência, criados para objetivos distintos. Além da CF (BRASIL,1988), outros documentos legais que justificam e fornecem bases para a Gestão de Dados de Pesquisa no Brasil são apresentados no Quadro 4. Ressalta-se que a legislação é volátil e podem sofrer alterações no decorrer do tempo. Além do mais, outras leis podem ser criadas, logo, é importante estar sempre atualizado.

Quadro 4 – Embasamento legal para a GDP no Brasil

Legislação	Disposições
Lei nº 12.527, de 18 de novembro de 2011. Lei de Acesso à Informação (LAI)	Regula o acesso a informações previsto em artigos da Constituição Federal e dá outras providências. Dispõe sobre os direitos do cidadão para acessar informações públicas produzidas ou custodiadas pelos órgãos públicos, incluindo dados de pesquisa.
Lei nº 12.965, de 23 de abril de 2014. Marco Civil da Internet	Estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil e determina as diretrizes para atuação da União, dos Estados, do Distrito Federal e dos Municípios em relação à matéria. Inclui a proteção e outras informações sobre os dados pessoais.
Lei nº 10.973, de 2 de dezembro de 2004 - Lei da Inovação	Estabelece medidas de incentivo à inovação e à pesquisa científica e tecnológica no ambiente produtivo, com vistas à capacitação tecnológica, ao alcance da autonomia tecnológica e ao desenvolvimento do sistema produtivo nacional e regional do País.
Decreto nº 8.777, de 11 de maio de 2016. Decreto de Dados Abertos	Estabelece diretrizes para a gestão e publicação de dados abertos no Brasil de órgãos e entidades da administração pública federal direta, autárquica e fundacional.
Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD)	Dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.
Lei nº 14.129, de 29 de março de 2021. Lei do Governo Digital (E-gov)	Dispõe sobre princípios, regras e instrumentos para o Governo Digital para o aumento da eficiência da administração pública,





## 4.2 EXPLORANDO OS CONCEITOS

Durante as leituras, verificou-se que haviam termos distintos utilizados para indicar as mesmas atividades em outros casos, eram imprecisos. Foi necessário padronizá-los com a máxima precisão para evitar o comprometimento da qualidade desta pesquisa. Uma análise do contexto foi realizada para compreender o sentido de componentes do texto para apresentar nomenclaturas e conceitos. Para esse fim, nesta subseção, utilizou-se como fonte de referência a incidência dos termos e conceitos na literatura, nas políticas, nas organizações que oferecem serviços de na área de GDP e serviço de vocabulário controlado. Em alguns casos, os elementos estavam tão entrelaçados que elaborar um conceito tornou-se inviável, dada a complexidade da atividade, que transcende um conceito concreto ou padrão. Portanto, optou-se por apresentar definições que relacionam a atividade às suas funções e objetivos.

Um caso mais recorrente, foi quanto o arquivamento e o armazenamento, quais são atividades realizadas em etapas distintas e com propósitos diferentes. As universidades de Utrecht e Radboud enfatizaram em suas respectivas políticas que os termos, embora possam parecer sinônimos, possuem objetivos diferentes.

O DeCS, sob a estrutura hierárquica da Ciência da Informação > Curadoria de Dados, apresenta o conceito de armazenamento como “Atividades organizadas relacionadas com a estocagem, localização, busca e recuperação de informação”. O conceito apresentado a seguir foi sintetizado pela autora desta pesquisa com base nas definições apresentadas nas fontes: Cunha e Cavalcanti (2008); Dicionário Cambridge; DCC; nas políticas para GDP das Universidades de Leiden; de Radboud; de Utrecht; Sayão e Sales (2015) e DeCS:

**Armazenamento:** conjunto de atividades realizadas para viabilizar a guarda e proteção dos dados brutos e processados assim que são gerados/coletados, ainda durante o processo de pesquisa. Nesta fase, o acesso é limitado aos membros da equipe, salvo nos casos em que a política prevê acordos entre pesquisadores que não fazem parte da pesquisa. Tem como objetivo manter os dados acessíveis para análise e consulta, preservar e proteger os dados evitando perdas, corrupção ou destruição dos arquivos e possibilitar a verificação da integridade durante a pesquisa.

Nas políticas institucionais consultadas, para realizar o arquivamento, a infraestrutura era oferecida pela IES, em servidores locais das Faculdades/Institutos, laboratório e/ou sistemas de gerenciamento em nuvem. Algumas políticas incluíram o gerenciamento dos

dados físicos. Nesses casos, o acondicionamento deveria ser realizado em um espaço seguro e exclusivo.

Por sua vez, sob a mesma estrutura hierárquica apresentada no DeCS, o arquivamento é definido como “Atividades de gestão necessárias à manutenção de dados de pesquisa de forma a assegurar que eles sirvam para uso contemporâneo e que estejam disponíveis para descoberta e reuso”. Nesta fase da GDP, são colocados em prática os critérios do gerenciamento apresentados no PGD para viabilizar o compartilhamento.

O conceito de Arquivamento foi sintetizado pela autora, por meio da consulta e síntese dos conceitos extraídos nas respectivas fontes: CODATA; LCRDM; Sayão e Sales (2015); Cunha e Cavalcanti (2008); DeCS/MeSH.

**Arquivamento:** conjunto de atividades realizadas para promover a organização, classificação e guarda de dados brutos e processados, de forma metodológica e sistemática para garantir a preservação a longo prazo, dos dados gerados/coletados após a realização da pesquisa ou da publicação dos resultados. Seu objetivo é assegurar a integridade e a acessibilidade, permitindo que outros pesquisadores possam validar, reproduzir ou reutilizar os dados.

Os locais comumente utilizados para o arquivamento são RD ou RI da IES (os RI são, geralmente, multidisciplinares), Repositório de Dados externos especializados ou plataformas governamentais.

No período de arquivamento, deve-se atribuir elementos adicionais para identificar e comunicar normativas aos que acessarem os dados. No Quadro 5 são especificados alguns itens que devem ser atribuídos durante o armazenamento e arquivamento. A organização do quadro tomou como base as informações disponíveis nas fontes: Sayão e Sales (2015); LCRDM; CODATA; Scielo; IBGE (2020); Borgam e Bourne (2021); JBRJ (2021); Embrapa (2019); PPBio (2009).

Quadro 5 – Alguns conceitos e as fases na gestão

Fase da GDP	Elemento	Conceito
Armazenamento e Arquivamento	Metadados	São dados sobre os dados. Servem para descrever o contexto, o conteúdo, além de facilitar sua encontrabilidade e acesso. Incluem nome do arquivo, do autor, formato, a data de criação e outros.
Armazenamento e Arquivamento	Documentação	Conjunto de informações sobre uma coleção de dados, necessária para que a integridade e qualidade possam ser verificadas e contextualizadas. São considerados como documentação os cadernos de laboratório, planilhas, formulários, prontuários e outros. Embora pouco citado na literatura, a documentação deve incluir o PGD e a política para GDP vigente quando os dados foram gerados/coletados.
Arquivamento	Licenças	Códigos, geralmente atribuídos pelos detentores dos direitos autorais, que indicam os níveis de acesso e uso dos dados a outros usuários. As licenças são importantes para garantir que os dados serão usados de forma ética e responsável.
Arquivamento	Identificadores	Códigos atribuídos aos dados para identificá-los de forma única e persistente que facilitam o acesso aos dados

Fonte: Elaborado pela autora, 2023.

Conforme Costa (2017, p. 210) os termos curadoria de dados e gestão de dados foram utilizados como sinônimos e isso “reflete a literatura internacional sobre o tema”. A autora também argumenta que, no contexto dos dados científicos da *e-Science*, esta “terminologia ainda é incipiente e carece de estudos para padronizar o uso de termos de forma a não se sobreponem em significado, bem como respeitar a origem (etimologia) dos termos oriundos de outras áreas, a exemplo do termo gestão” (Costa, 2017, p. 211).

O conceito de curadoria mas abrangente encontrado, foi do CODATA [20--a]:

O processo de gerenciar, ao longo do ciclo de vida dos dados, pelo qual os dados/coleções de dados são limpos, documentados, padronizados, formatados e inter-relacionados. Isso inclui dados de versão ou formação de uma nova coleção de várias fontes de dados, anotação com metadados, adição de códigos a dados brutos (por exemplo, classificação de uma imagem de galáxia com um tipo de galáxia como “espiral”). Níveis mais altos de curadoria envolvem a manutenção de links com anotações e outros materiais publicados. Assim, um conjunto de dados pode incluir um link de citação para a publicação cuja análise foi baseada nos dados. O objetivo da curadoria é gerenciar e promover o uso de dados desde seu ponto de criação para garantir que sejam adequados para propósitos contemporâneos e disponíveis para descoberta e reutilização. Para conjuntos de dados dinâmicos, isso pode significar enriquecimento ou atualização contínua para mantê-los adequados à finalidade. Formas especiais de curadoria podem estar disponíveis em repositórios

de dados. O próprio processo de curadoria de dados deve ser documentado como parte da curadoria. Assim, curadoria e proveniência estão altamente relacionadas.

O conceito do CODATA [20--a] apresenta diversas e complexas atribuições. Encontrar um curador com estas qualificações parece algo inatingível. Soma-se ao fato de que as pessoas podem se desvincular da IES, mudar de setor ou assumir outras funções.

O DeCS define os serviços de curadoria como “Atividades de gestão necessárias à manutenção de dados de pesquisa de forma a assegurar que sirvam para uso contemporâneo e que estejam disponíveis para descoberta e reuso”. A Fiocruz descreve como “Ações voltadas para o gerenciamento de dados de pesquisa durante o seu ciclo de vida; envolve manter, preservar e adicionar valor aos dados”.

Sintetizando as definições encontradas nas fontes: Scielo; Dudziak (2019); LCRDM; DeCS/MeSH Descritores em Ciências da Saúde; National Research Council (2015); Sayão e Sales (2015) e Fiocruz ([20--]), apresenta-se o seguinte conceito:

**Curadoria de dados:** serviços para gerenciar os dados de pesquisa durante todo o seu ciclo de vida, de forma a garantir a preservação e o acesso em longo prazo, a integridade, a encontrabilidade, a acessibilidade para uso e reuso. As atividades compreendem a revisão dos dados e metadados, disponibilidade da documentação e verificação quanto à conformidade legal, ética e das normativas institucionais, incluindo a política para GDP.

Por sua gama de atividades, esse serviço demanda de profissionais com conhecimento holístico de todo o processo de gestão. A atualização de metadados, formatos de arquivos, softwares e outros, levando em consideração o volume de arquivos, parece inexecutável para um profissional. Mesmo que a IES indique mais de uma pessoa para ser o curador, provavelmente, seria também professor e pesquisador, os desafios são imensos.

Ao descrever quais competências devem ser desenvolvidas por profissionais para o trabalho com curadoria digital, Sayão e Sales (2016, p. 70) indicam que:

[...] os profissionais de informação e as infraestruturas informacionais e tecnológicas subjacentes às bibliotecas de pesquisa devem estar imbricadas nas atividades dos laboratórios e das outras atividades acadêmicas e de pesquisas da instituição. Isto demanda uma reformulação nos perfis profissionais dos bibliotecários e arquivistas que agora se tornam profissionais de dados. Além disso, a interlocução necessária à curadoria exige ainda novos profissionais provenientes dos domínios específicos – como pesquisadores - e profissionais da área de computação.

Essas considerações ficam apenas para reflexão, pois nos limites desta pesquisa, a complexidade da nomenclatura de curadoria e gestão de dados e do curador e gestor de dados carece de estudos de profissionais com maior conhecimento na área.

Na literatura, as atividades de um curador correspondem às de um administrador de dados, que é o termo recorrente nas políticas para GDP das IES holandesas. Constatou-se, também, que as atividades do administrador de dados variam entre as instituições. Mesmo após uma análise exaustiva, não foi possível apresentar um conceito que possa ser mais assertivo.

Como proposta para indicar as funções do Administrador de Dados, tendo como referência as considerações das políticas para GDP das universidades de Utrecht, Amsterdam, Tilburg e LCRDM, seguir como conceito:

**Administrador de dados/Curador de dados:** o profissional responsável por gerenciar os dados durante todo o seu ciclo de vida, que inclui atuar na construção de políticas e procedimentos de gerenciamento de dados; projetar e implementar sistemas de gerenciamento de dados; verificar a qualidade dos dados; protegê-los contra acesso não autorizado; arquivar e preservar; e compartilhar os dados com outros usuários.

Sendo os tipos de dados categorizados conforme a sua estrutura, conteúdo ou propósito, existem diferentes formas de acesso. A seguir, apresenta-se nos Quadros 6, alguns tipos de dados; no Quadro 7, a classificação dos dados e no Quadro 8, as formas para acessos.

O Quadro 6 foi construído com base nas informações extraídas das fontes: LCRDM, Universidade de Utrecht (2019); Embrapa (2019); Universidade de Rotterdam (2020); Universidade de Leiden (2021); Santos, P.; Almeida, B.; Henning, P. (org.), (2017); Scielo; JBRJ (2021); CODATA; Sayão e Sales (2015) e IBGE (2020).

Quadro 6 – Conceituando alguns tipos de dados

Termo	Conceito
Dados abertos	São dados publicados em um formato legível por máquina e sem restrição de licenças, patentes ou mecanismos de controle, de modo a estarem livremente disponíveis para serem utilizados e redistribuídos para qualquer pessoa.
Dados sigilosos	São dados que não podem ser acessados devido a diversas razões: em processo de patente, comprometimento da soberania e segurança pública, dados que contenham informações sensíveis e pessoais e outros.
Dados brutos	Dados que vêm diretamente dos instrumentos científicos ou coletados diretamente da fonte sem sofrer nenhuma manipulação ou processamento.
Dados processados	Dados que foram processados, minerados, selecionados, recodificados, visualizados, categorizados, agregados, mesmo sendo de resultados negativos e inconclusivos.
Dados Sensíveis	Dados que contêm informações sobre raça, origem étnica, grupo político, religião ou crenças, filiação sindical, doença física e/ou mental, vida sexual e outros.
Dados Pessoais	Dados de pessoa natural identificada ou ‘identificável’, que pode ser reconhecida, direta ou indiretamente pelo nome, número dos documentos, endereço residencial, contatos telefônicos, e-mail e outros dados.
Dados Anonimizados	Dados que passaram por um processo pelo qual não seja possível identificar o indivíduo ao qual se referem.
Dados Pseudonimizados	Dados originais que foram substituídos por um identificador por um código com o objetivo de evitar a identificação direta do participante/pessoa.

Fonte: Elaborado pela autora, 2023.

A classificação, o impacto e as características dos dados são apresentadas no Quadro 7, com base nas fontes: Fiocruz [20--], Embrapa (2019); Scielo; IBGE (2020); JBRJ (2021) e CODATA.

Quadro 7 – Classificação: impactos e riscos

Classificação	Impacto	Características	Adequado para
Sigiloso/ Fechado	Alto risco	É o nível de risco mais alto quando ao acesso. Apenas um pequeno grupo de pessoas podem acessá-los, por meio de autorização expressa, seguindo protocolos de segurança	Dados pessoais, sensíveis, segurança nacional e confidenciais
Acesso restrito	Risco Moderado	Nível intermediário, permite que um grupo maior de pessoas acesse os dados, mas de forma controlada	Dados que apresentam menor grau de risco, como dados embargados
Acesso aberto	Baixo risco	O acesso aberto é o nível de acesso mais baixo e permite que qualquer pessoa acesse os dados.	Dados que não são sensíveis ou confidenciais

Fonte: Elaborado pela autora, 2023.

As políticas holandesas mencionam tipos de dados, classificação e níveis de acesso ao longo do texto, mas algumas não incluem um glossário. Isso pode ocorrer porque os pesquisadores podem já ter adquirido conhecimento suficiente para compreensão ou que devem buscar informações complementares em outras fontes. No cenário brasileiro, seria altamente recomendado apresentar os conceitos na política para evitar equívocos e omissões, pois é um serviço novo e ainda pode ser desconhecido por muitos pesquisadores. É fundamental compreender o significado das palavras e seus propósitos no contexto dos dados de pesquisa para assimilar as normas e sugestões apresentadas nas políticas para GDP. Essa assimilação de conceitos também pode aumentar a confiança e reduzir os receios dos pesquisadores, permitindo o acesso e o reuso de seus dados e evitando restrições desnecessárias.

Tomar conhecimento das formas de acesso ajuda a compreender, também, o motivo pelo qual um conjunto de dados está fechado ou indisponível, para uma transparência e compreensão do contexto da pesquisa.

O Quadro 8 foi construído com informações extraídas das fontes: LCRDM, Universidade de Utrecht (2019); Embrapa (2019); Universidade de Leiden (2021); Santos, Almeida e Henning (org.) (2017); Universidade de Rotterdam (2020); Scielo; JBRJ (2021); CODATA; Sayão e Sales (2015) e IBGE (2020):



Quadro 8 – Tipos de acessos: relacionando conceitos

Tipo de acesso	Característica
Acesso aberto	Quando o acesso e o reuso são permitidos livremente por qualquer pessoa sem restrições técnicas ou legais.
Acesso restrito	O acesso apresenta restrições e controles para limitar quem pode acessar os dados.
Acesso fechado	Constitui-se em um nível ainda mais restrito de acesso, autorizado apenas a um grupo muito limitado de indivíduos ou uma única entidade.
Acesso embargado	Dados em que o acesso e o reuso serão permitidos após um tempo predeterminado.
Acesso Compartilhado	Dados em que o acesso e o reuso são limitados a determinadas pessoas ou grupos.

Fonte: Elaborado pela autora, 2023.

Por fim, apresentar um conceito para “Política para Gestão de Dados de Pesquisa” é imprescindível para que se possa reconhecer uma “política”, mesmo quando o documento apresenta nomenclatura diversa e para excluir os que não atendam aos critérios necessários.

Destarte, sintetizando as definições relacionadas nas fontes: LCRDM; CODATA e Portal Foster:

**Política para Gestão de Dados de Pesquisa** “é um documento que estabelece os princípios, diretrizes e procedimentos que orientam o gerenciamento de dados de pesquisa. Isso engloba o planejamento da geração/coleta de dados, armazenamento, processamento, compartilhamento, preservação e descarte. Além disso, a política define as responsabilidades dos profissionais envolvidos, os recursos institucionais disponíveis e aborda questões éticas e legais relacionadas à gestão de dados”. No contexto das IES, o documento deve ser um instrumento formal e obrigatório, assinado pelo representante máximo da instituição.

Os conceitos que foram apresentados nesta subseção são uma síntese das leituras realizadas durante esta pesquisa. Alguns estão em conformidade com as previsões legais e éticas e outros foram apresentados por IES e organizações afins. Os resultados não são absolutos e podem variar conforme as circunstâncias.

### 4.3 A GESTÃO DE DADOS DE PESQUISA: BUSCANDO ABORDAGENS NACIONAIS

Como as potencialidades dos dados de pesquisa foram reconhecidas em diversos setores de natureza pública e privada, incluindo empresas, organizações governamentais, institutos de pesquisa e outros, tomar conhecimento das normativas para GDP dessas organizações é importante, considerando o relacionamento contínuo destas com as IES. Nesse cenário, emerge também a presença e a força das agências de fomento na promoção da GDP. O Quadro 9 apresenta instituições onde foram localizadas normativas ou abordagens sobre o tema.

Quadro 9 – A busca por normativas no território nacional

Instituição	Título	Ano
Programa de Pesquisa em Biodiversidade no Brasil (PPBio)	Política de Dados do PPBio	2009
Empresa Brasileira de Pesquisa Agropecuária (Embrapa)	Política de Governança de Dados, Informação e Conhecimento da Embrapa	2019
Fundação Espírito-Santense de Tecnologia (FEST)	Plano de Gestão de Dados	2019
Fundação Oswaldo Cruz (Fiocruz)	Política de gestão, compartilhamento e abertura de dados para pesquisa: Princípios e Diretrizes	2020
Jardim Botânico do Rio de Janeiro (JBRJ)	Política de Gestão, Acesso, Publicação e Compartilhamento de Dados e Recursos de Informação do JBRJ	2021

Fonte: Elaborado pela autora, 2023.

A política do PPBio se destaca por apresentar um corpo significativo de elementos e pela data de sua publicação, quando ainda eram raras as instituições nacionais e internacionais que apresentavam normas para gerenciar seus dados. Embora haja algumas lacunas em relação às práticas atuais, reconhece-se que muitos recursos foram desenvolvidos anos depois.

Apesar do título, a Política de gestão, compartilhamento e abertura de dados para pesquisa: Princípios e Diretrizes da Fiocruz, enfatiza os princípios e diretrizes da fundação, conforme indicado no seu subtítulo, com características de um documento de governança institucional. Entretanto, foi possível extrair alguns elementos para fins práticos, necessários para uma efetiva GDP.

Devido às características peculiares das normativas, não foi possível apresentar um quadro comparativo, pois possuem particularidades que não são passíveis de comparações.

No entanto, no Quadro 10, foram destacadas as características relevantes para uma melhor visualização das informações obtidas.

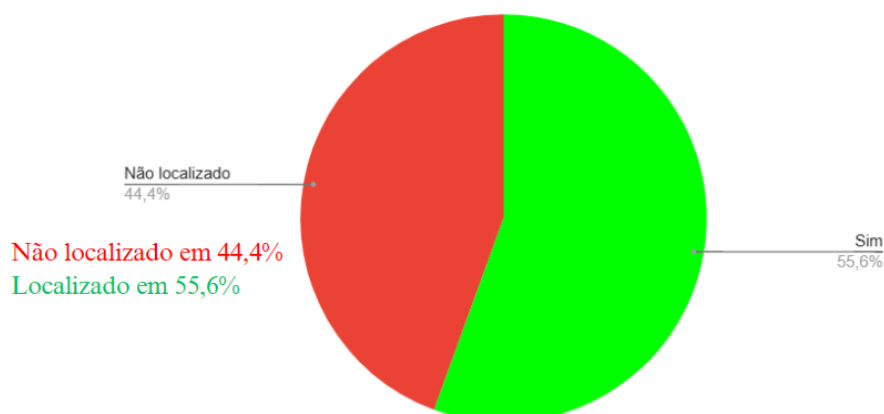
Quadro 10 – Normativas em institutos de pesquisa nacionais

Instituição	Principais elementos e abordagens
PPBio	Contextualização dos objetivos da GDP; glossário, a classificação e tipos de dados a serem gerenciados, inclui os dados físicos; critérios para acessos e metadados públicos para todos os tipos de dados; propriedade intelectual, stakeholders e responsabilidades, subordinação à política, segurança e proteção, citação e referência. Embora haja a previsão para revisão e atualização, não foram encontrados documentos mais recentes.
Embrapa	RD institucional, aborda o armazenamento e arquivamento, formatos abertos, reuso, citação e referência, funções e responsabilidade de stakeholders, administrador de dados, protocolos específicos nas unidades da empresa e glossário.
FEST	Armazenamento em formatos abertos, o reuso, propriedade intelectual, compartilhamento de dados durante a pesquisa, responsabilidade pela GDP, a citação e referência dos dados, identificadores persistentes, stakeholders e suas responsabilidades, administrador de dados e suas funções.
Fiocruz	Política de gestão, compartilhamento e abertura de dados para pesquisa: Princípios e Diretrizes.
JBRJ	Repositório de dados, classificação dos tipos de dados a serem gerenciados, metadados, armazenamento e arquivamento em formatos abertos, responsável pela GDP, níveis de acesso, licenças de uso, lista de stakeholders, previsão de atualização da política e glossário.

Fonte: Elaborado pela autora, 2023.

Outras instituições listadas na Metodologia desta pesquisa foram o Instituto Butantan, o Instituto Nacional de Pesquisas Espaciais (INPE), o Instituto de Matemática Pura e Aplicada (IMPA) e o Instituto de Pesquisas Tecnológicas (IPT). No entanto, entre agosto de 2022 e janeiro de 2023 não foram localizados documentos ou qualquer outra informação explícita sobre a gestão de dados de pesquisa. O Gráfico 1 aponta as porcentagem de documentos localizados nas instituições de pesquisa nacionais.

Gráfico 1 - O panorama nos institutos de pesquisa nacionais



Fonte: Elaborado pela autora, 2023.

Os modelos de políticas encontrados abordam diversos elementos e aspectos importantes da gestão de dados. Apesar dos avanços, existem lacunas em suas abordagens, como a ausência de informações sobre o período mínimo para o arquivamento, prazos para embargo, previsão dos custos da gestão, serviços para capacitação, pesquisas realizadas em colaboração com pesquisadores externos, dentre outros elementos. O aprimoramento, incorporando elementos para uma gestão mais completa, demanda revisão e atualização da política, que é outro componente presente nas políticas mais abrangentes.

#### 4.3.1 Iniciativas das agências de fomento e apoio à pesquisa

A experiência internacional mostra que o envolvimento das agências de fomento foi essencial para o sucesso da GDP, como na Holanda, EUA, Canadá, China, Austrália e outros. Como resultado, esses países vêm desenvolvendo um forte ecossistema de gestão que está apresentando resultados e benefícios.

Para certificar se as agências de fomento e apoio às pesquisas nacionais estão estimulando as práticas para a GDP, foram consultados os principais financiadores, apontados por Dudziak (2018), como sendo: o BNDES, o CNPq, a Capes, a Finep, a Fapemig, a Fapesp, a Faperj, a Fapergs e o Fundo Nacional de Desenvolvimento Científico e Tecnológico. A pesquisa foi realizada nos sites de cada instituição, em agosto de 2022.

Dentre os financiadores de pesquisa consultados, apenas na Fapesp foi encontrado um documento que incentiva a GDP nas IES, a declaração intitulada “Diretrizes para Planos de Gestão de Dados para Propostas de Centros”, publicada em 2017. No Quadro 11, apresenta-se de forma sintética, algumas perguntas que devem ser respondidas no PGD. Sugere-se consultar o documento completo para informações detalhadas.

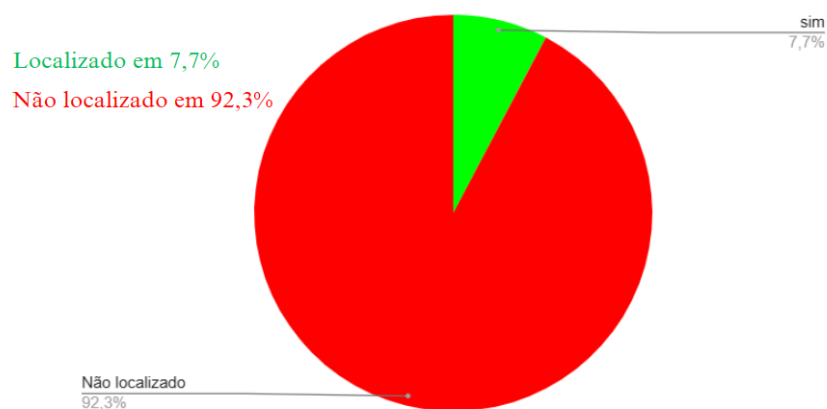
Quadro 11 – Padrão Fapesp para planos de gestão de dados

Tópico	Conteúdo
Sumário descritivo	Tipos e formatos de dados a serem gerados/coletados; Estimativa do volume; Armazenamento e compartilhamento de dados brutos; Quais os outros objetos a serem gerados (software, workflows, modelos etc.).
Gerenciamento	Padrões de metadados para documentação e dos objetos digitais; Formatos de dados e metadados para viabilizar o compartilhamento; Providências para permitir o reuso; Atribuição de identificadores.
Preservação e compartilhamento	Abertura e disponibilização livre dos dados; Tipos de dados com restrições de acesso e justificativa; Atribuição de licenças de uso; Mecanismos para compartilhamento; Políticas diferenciadas dentro do Centro para compartilhamento de dados, níveis de acesso.
Gerenciamento, responsabilidades e Segurança	Responsável pela GDP; Comitê gestor de dados; Processo de decisão sobre a preservação e descarte dos dados; Mecanismos de preservação durante a existência do Centro; Planejamento para preservação a longo prazo; Medidas de segurança para evitar acesso indevido ou destruição; Monitoramento e atualização do PGD ao longo da pesquisa.
Conformidade legal e ética	Quais são as restrições éticas e legais que podem impactar o compartilhamento dos dados; Medidas para conformidade com a legislação; Formas para viabilizar reuso.
Repositórios	Criação de um repositório próprio de dados; Descrição do planejamento para criação e gestão do repositório; Possibilidade de uso de repositórios institucionais para armazenamento dos dados.

Fonte: Elaborado pela autora, 2023.

O Gráfico 2 revela o baixo envolvimento das agências de fomento na promoção da GDP no Brasil.

Gráfico 2 – O panorama nas agências de fomento nacionais



Fonte: Elaborado pela autora, 2023.

A limitada adesão ao movimento das agências de fomento impacta na velocidade com que os serviços são implementados. Não se deve esperar que apenas as agências tenham iniciativas, mas é inegável sua força e influência na promoção e divulgação de novas práticas científicas.

Na subseção 4.3.2, destaca-se a promoção da gestão de dados de pesquisa como parte dos serviços oferecidos pelas IES.

#### 4.3.2 Os primeiros acordos das IES nacionais

A fim de levantar informações sobre iniciativas para a GDP nas IES nacionais, foram explorados os sites de 53 instituições, selecionadas conforme o índice IGC, de acordo com critérios apontados na Metodologia.

O Quadro 12 foi organizado de acordo com a Coluna B, na qual é apresentada a sigla da unidade da federação.

Quadro 12 – As IES nacionais: uma delimitação do campo

Nº	UF	Sigla	Instituição
1	AM	UFAM	Universidade Federal do Amazonas
2	BA	UFBA	Universidade Federal da Bahia
3	CE	UNIFOR	Universidade de Fortaleza
4	CE	UECE	Universidade Estadual do Ceará
5	CE	UFC	Universidade Federal do Ceará
6	DF	UNB	Universidade de Brasília
7	DF	UCB	Universidade Católica de Brasília
8	GO	UFG	Universidade Federal de Goiás
9	MA	UFMA	Universidade Federal do Maranhão
10	MG	FAJE	Faculdade Jesuíta de Filosofia e Teologia
11	MG	PUC/MG	Pontifícia Universidade Católica de Minas Gerais
12	MG	UNIMONTES	Universidade Estadual de Montes Claros
13	MG	UFLA	Universidade Federal de Lavras
14	MG	UFU	Universidade Federal de Uberlândia
15	MG	UFV	Universidade Federal de Viçosa
16	MS	UFMS	Universidade Federal de Mato Grosso do Sul
17	PA	UFPA	Universidade Federal do Pará
18	PB	UFPB	Universidade Federal da Paraíba
19	PB	UFCG	Universidade Federal de Campina Grande
20	PE	UFPE	Universidade Federal de Pernambuco
21	PE	UFRPE	Universidade Federal Rural de Pernambuco
22	PR	UFPR	Universidade Federal do Paraná
23	PR	PUC/PR	Pontifícia Universidade Católica do Paraná
24	PR	UEM	Universidade Estadual de Maringá
25	RJ	UFF	Universidade Federal Fluminense
26	RJ	PUC-RIO	Pontifícia Universidade Católica do Rio de Janeiro
27	RJ	UERJ	Universidade do Estado do Rio de Janeiro
28	RJ	UENF	Universidade Estadual do Norte Fluminense Darcy Ribeiro
29	RJ	UFRJ	Universidade Federal do Rio de Janeiro
30	RJ	IME	Instituto Militar de Engenharia
31	RJ	FGV	Fundação Getúlio Vargas
32	RN	UFERSA	Universidade Federal Rural do Semiárido
33	RS	UFRN	Universidade Federal do Rio Grande do Norte
34	RS	PUC/RS	Pontifícia Universidade Católica do Rio Grande do Sul

35	RS	UNISINOS	Universidade do Vale do Rio dos Sinos
36	RS	UFPEL	Universidade Federal de Pelotas
37	RS	UFSM	Universidade Federal de Santa Maria
38	RS	FURG	Universidade Federal do Rio Grande
39	RS	UFRGS	Universidade Federal do Rio Grande do Sul
40	SC	UNESC	Universidade do Extremo Sul Catarinense
41	SC	UNIVALI	Universidade do Vale do Itajaí
42	SC	UFSC	Universidade Federal de Santa Catarina
43	SP	USP	Universidade de São Paulo
44	SP	UNESP	Universidade Estadual Paulista
45	SP	ITA	Instituto Tecnológico de Aeronáutica
46	SP	PUC/SP	Pontifícia Universidade Católica de São Paulo
47	SP	UNICSUL	Universidade Cruzeiro do Sul
48	SP	UNICAMP	Universidade Estadual de Campinas
49	SP	UNIFESP	Universidade Federal de São Paulo
50	SP	UFSCAR	Universidade Federal de São Carlos
51	SP	UPM	Universidade Presbiteriana Mackenzie
52	SP	USF	Universidade São Francisco
53	SP	UNIVERITAS	Universidade Universitas Veritas

Fonte: Elaborado pela autora, 2023.

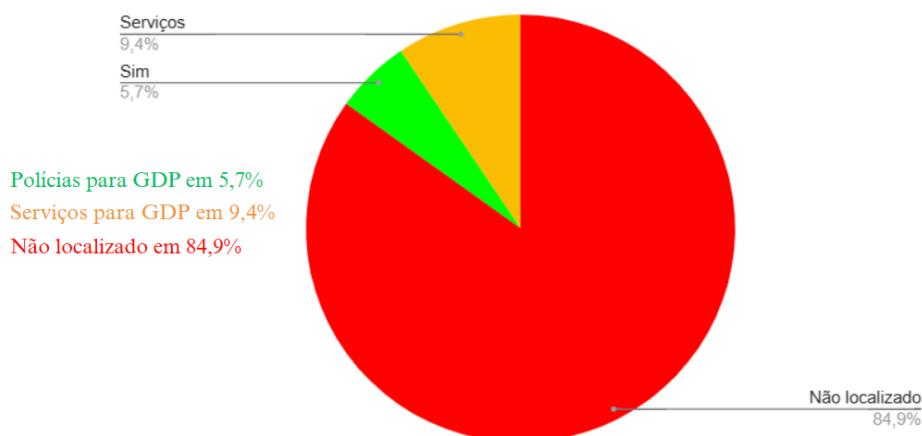
Apenas na UFPR, na USP e na UNIFESP foram localizadas políticas para GDP, conforme o conceito de política apresentado nesta pesquisa. Foram encontrados registros em outras IES quando a promoção de serviços para o gerenciamento de dados, sendo a UFSC, a UNESP, a UFSCar, a Unicamp e a UFABC. No entanto, nestas IES, não foram localizados documentos que preenchiam os critérios necessários para serem considerados como políticas. Os serviços encontrados incluem diretrizes para o uso do repositório de dados institucional, orientações para elaboração de planos de gestão de dados e links para esclarecimentos quanto a outros serviços relacionados.

Em 45 IES, não há ou não foi localizado qualquer tipo de normativa, serviço ou orientação relacionada ao gerenciamento dos dados de pesquisa. Este número significativo pode ser atribuído à falta de conhecimento de sua comunidade acadêmica, e/ou desinteresse em implementar os serviços ou ainda, as limitações financeiras, na infraestrutura física e



tecnológica e na de pessoal capacitado. No Gráfico 3, apresenta-se o panorama das IES nacionais.

Gráfico 3 – O panorama nas IES nacionais



Fonte: Elaborado pela autora, 2023.

As políticas encontradas apresentaram variações no conteúdo e abrangência nos elementos. Isso é esperado, pois as IES são diferentes em termos de tamanho, estrutura, recursos e prioridades. É importante reconhecer que estas universidades estão dando os primeiros passos na GDP. Portanto, não é correto exigir a completude que se encontra em IES estrangeiras, pois realizam serviços conforme suas possibilidades.

No Quadro 13, foram relacionados, resumidamente, os elementos da UFPR, UNIFESP e USP. Devido à diversidade de informações, alguns itens não foram mencionados e descritos com precisão.

Quadro 13 – Elementos de políticas para GDP em IES nacionais

UFPR	
<i>Título</i>	<i>Diretrizes da Base de Dados Científicos da Universidade Federal do Paraná, de 2018</i>
<i>Normas, elementos e serviços</i>	<ul style="list-style-type: none"> <li>- Possui Repositório de Dados Institucional;</li> <li>- Oferta cursos de capacitação;</li> <li>- Oferece serviços de orientação e suporte;</li> <li>- Possui uma Equipe de Gestão do RD;</li> <li>- Oferece serviços de apoio no sistema de bibliotecas;</li> <li>- Possui um PGD;</li> <li>- Aponta o responsável pela integridade dos dados;</li> <li>- Incentiva a publicação em Acesso Aberto visando o reuso;</li> <li>- Gerencia os dados das pesquisas de pós-graduação;</li> </ul>

	<ul style="list-style-type: none"> <li>- Aponta critérios para embargos e exclusão de dados;</li> <li>- Gerencia dados gerados em pesquisas de pós-graduação;</li> <li>- Lista quem pode arquivar dados no RD;</li> <li>- Arquia dados com a documentação;</li> <li>- Atribui metadados, licenças de uso e identificadores persistentes;</li> <li>- Mantém metadados visíveis para qualquer tipo de acesso;</li> <li>- Possui formatos de arquivos abertos;</li> <li>- Possui sistema de backup;</li> <li>- Informa o espaço máximo que pode ser ocupado por cada conjunto de dados.</li> <li>- Orienta aos pesquisadores que façam cópias dos arquivos;</li> <li>- Permite o arquivamento em RD externo;</li> <li>- Solicita que os dados reusados sejam citados e referenciados;</li> <li>- Possui glossário;</li> <li>- Tem previsão de revisão e atualização das normativas do documento.</li> </ul>
<b>UNIFESP</b>	
<i>Título</i>	<i>Política Institucional de Informação para o Repositório Institucional Unifesp, 2017</i>
<i>Normas, elementos e serviços</i>	<ul style="list-style-type: none"> <li>- Repositório de Dados Institucional;</li> <li>- Comitê Gestor do Repositório Institucional;</li> <li>- Listagem de quem pode arquivar dados no RD;</li> <li>- Publicação com Acesso Aberto, salvo em exceções previstas;</li> <li>- Publicação em formato padronizado e normatizado;</li> <li>- Metadados padronizados; licenças de uso e identificadores persistentes;</li> <li>- Informação de quando os dados devem ser arquivados;</li> <li>- Informação sobre prazos para embargos;</li> <li>- Indicação dos tipos de dados que devem ser arquivados;</li> <li>- Apresentação dos critérios que desobrigam os autores a publicarem os dados.</li> </ul>
<b>USP</b>	
<i>Título</i>	<i>Gestão de Dados Científicos na Universidade de São Paulo, 2019</i>
<i>Normas, elementos e serviços</i>	<ul style="list-style-type: none"> <li>- Possui RD Institucional;</li> <li>- Possui uma Equipe de Gestão do RD;</li> <li>- Solicita o PGD (não informa se possui um modelo);</li> <li>- Arquia o PGD junto aos dados;</li> <li>- Tem um responsável pela integridade dos dados;</li> <li>- Tem espaço máximo para armazenamento;</li> <li>- Delimita tempo para armazenamento interno;</li> <li>- Solicita os metadados para armazenar os dados;</li> <li>- Possui critérios para custear os dados armazenados.</li> </ul>

Fonte: Elaborado pela autora, 2023.

Sendo uma atividade em estágios iniciais no Brasil, suas normativas são especialmente importantes, pois marcam o surgimento de novas perspectivas para as instituições que promovem o ensino e a pesquisa, projetando o Brasil neste novo contexto para uma pesquisa mais colaborativa e transparente. Nas IES em que foi encontrado algum tipo de serviço ou informações, foram elencados alguns de seus elementos para ampliar o campo de visão sobre o envolvimento de outras universidades. A UFPR, a USP e a UNIFESP não foram incluídas na lista do Quadro 14 por já terem sido citadas anteriormente.

Quadro 14 – Os serviços de gestão nas IES nacionais

IES	Serviços
UFSC	<ul style="list-style-type: none"> <li>- Comissão de concepção do serviço de suporte à pesquisa e gerenciamento de dados;</li> <li>- Seminário de Suporte à Pesquisa e Gestão de Dados Científicos, em 2017.</li> </ul>
UNESP	<ul style="list-style-type: none"> <li>- Apresentação de definições e iniciativas da Ciência Aberta, do Acesso Aberto, dos Dados Abertos e sobre a Licença Creative Commons;</li> <li>- Lista de ferramentas para criar um PGD;</li> <li>- Repositório Institucional da UNESP;</li> <li>- Integra-se ao Metabuscador de dados da Fapesp;</li> <li>- Possui uma “Política de Acesso Aberto à Produção Intelectual da Unesp”.</li> </ul>
UNICAMP	<ul style="list-style-type: none"> <li>- Comissão de Gestão de Dados de Pesquisa;</li> <li>- Treinamentos para capacitação;</li> <li>- Seção de Perguntas e Respostas que instruem sobre GDP;</li> <li>- Repositório de Dados de Pesquisa da Unicamp (REDU);</li> <li>- Integra-se ao Metabuscador de dados da Fapesp.</li> </ul>
UFSCAR	<ul style="list-style-type: none"> <li>- Repositório Institucional (arquiva dados e documentos gerais);</li> <li>- Manual de autodepósito de dados de pesquisa.</li> </ul>
UFABC <sup>11</sup>	<ul style="list-style-type: none"> <li>- Repositório de Dados de Pesquisa da UFABC;</li> <li>- Informações sobre outros RDs;</li> <li>- Integra-se ao Metabuscador de dados da Fapesp.</li> </ul>

Fonte: Elaborado pela autora, 2023.

É possível que sejam oferecidos serviços e informações além do especificado. No entanto, foram listadas apenas as localizadas entre os meses de março e julho de 2023.

A Unicamp e a Unesp possuem política institucional de acesso aberto à produção intelectual e científica da universidade. Os documentos não atendem aos critérios para serem consideradas política para GDP, mas são importantes registros para a promoção da governança.

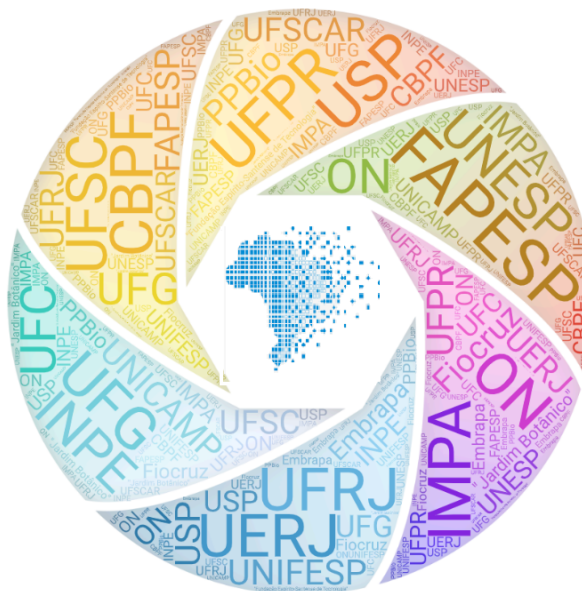
Nas IES do estado de São Paulo, conforme apresentado no Quadro 13, houve um consenso quanto à disponibilidade de serviços para o arquivamento dos dados. No entanto, de forma geral, apesar dos esforços empreendidos, algumas IES apresentam informações mínimas, limitando-se a disponibilizar links para documentos externos, sem objetividade e clareza em comparação com as IES estrangeiras. Porém, é importante destacar o mérito dessas IES por seu compromisso com a adoção das tendências globais, mesmo diante de limitações e desafios.

A nuvem de palavras da Figura 10 resume todas as instituições consultadas que, de

<sup>11</sup> A UFABC não foi selecionada conforme o IGC, mas como constava no Metabuscador da Fapesp, considerou-se relevante registrar suas iniciativas.

alguma forma, estão aderindo à GDP.

Figura 10 – Iniciativas de IES nacionais para a promoção da GDP



Fonte: Elaborada pela autora, 2023.

Na Figura 10, visualiza-se que, mesmo que sejam poucas, há um sinal positivo de que o Brasil está dando os primeiros passos na direção da gestão de dados de pesquisa. Com o tempo e com o investimento adequado, essas instituições poderão contribuir com serviços mais amplos, servindo de referência e incentivo a outras organizações que fazem pesquisa ou que de alguma forma, se relacionam com as IES.

Quanto à presença dos órgãos públicos, ainda é tímida a sua atuação para o desenvolvimento e disponibilização de ferramentas e suporte para a GDP no Brasil. Além disso, chama-se a atenção para o fato de que a administração pública federal deveria atuar na cobrança e fiscalização do cumprimento da legislação relacionada ao gerenciamento e abertura dos dados de pesquisa.

É possível que algumas IES públicas ainda não tenham plena consciência de sua natureza estatutária e de sua subordinação ao Poder Executivo Federal, que implica na obrigatoriedade de submissão à legislação. Por outro lado, é possível também, que a administração pública não tenha reconhecido a aplicabilidade da legislação às IES. Ambas as partes compartilham a responsabilidade pelo desperdício de recursos públicos e pelos prejuízos decorrentes. Isso traz implicações negativas para a sociedade brasileira, pois:

- 1) Dificulta o acesso aos dados de pesquisa à comunidade acadêmica, o que pode limitar a capacidade dos pesquisadores de replicar estudos e gerar novos conhecimentos.
- 2) Limita a projeção no cenário internacional ocultando a qualidade das pesquisas geradas no país.
- 3) Fere os direitos dos cidadãos quanto ao acesso aos dados financiados por meio de recursos públicos.
- 4) Opõe-se ao cumprimento da legislação nacional que dispõe sobre a abertura como dever e restrição como regra.
- 5) Prejudica a transferência de tecnologia para a indústria, o que pode prejudicar a competitividade da economia brasileira.

Existem ações promissoras da RNP e do BICT, como o RDP-Brasil<sup>12</sup> e o RDA Brasil<sup>13</sup>. O Repositório de Dados de Pesquisa do Brasil (RDP-Brasil) contém registros de trabalhos iniciados nos anos de 2018 e 2019, no entanto, não foram encontrados registros após esse período. É possível que o processo ainda esteja em andamento, mas, dado todo esse período, é importante aumentar a divulgação dos resultados já alcançados ou retomar as atividades, caso tenham sido interrompidas.

Outra iniciativa em andamento é a Incubação de Repositórios de Dados de Pesquisa, promovido pela RNP, CNPq e IBICT, com a “Chamada Aberta: Incubação de Repositórios de Dados de Pesquisa”. Os contemplados pelo edital de 2021 foram a UFC, UFG, a UFRA e CBPF (esta última não é uma IES). Conforme informações obtidas diretamente com estas instituições, os trabalhos estão em fase de tratativas quanto aos aspectos tecnológicos e de infraestrutura, além de documentos para a governança.

Na próxima subseção será apresentado o contexto sobre as universidades holandesas selecionadas com base nos critérios estabelecidos na metodologia desta pesquisa, bem como os serviços e a infraestrutura que possuem para a gestão de dados.

---

<sup>12</sup> A RDP-Brasil é uma iniciativa da RNP e IbiCT em colaboração com a UFRGS e FURG para formar um Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira para o desenvolvimento de um projeto em conjunto sobre o Acesso Aberto a Dados de Pesquisa (AADP)

<sup>13</sup> .O RDA Brasil é um grupo criado na plataforma do RDA para promover a discussão sobre a GDP e promover conexões entre pesquisadores, gestores de dados de pesquisa, formuladores de políticas públicas, profissionais da Ciência da Informação e atores em geral.

#### 4.4 ALÉM DAS FRONTEIRAS: AS UNIVERSIDADES HOLANDESAS COMO REFERÊNCIA

A escolha das universidades da Holanda como referência para a busca de elementos importantes para a gestão de dados de pesquisa e sua incorporação em uma política institucional para IES brasileiras foi fundamentada na seção de Metodologia. As universidades da Holanda são públicas, financiadas pelo governo, e com uma antiga e rica história no envolvimento da pesquisa e ensino, reconhecidas mundialmente. A mais antiga é a Universidade de Leiden, fundada em 1575, seguida pela Universidade de Groningen, fundada em 1614, e a Universidade de Utrecht, em 1636.

Com base nos critérios do *CWTS Leiden Ranking*, as treze instituições selecionadas foram:

- 1) *Universiteit Utrecht*: a Universidade de Utrecht foi fundada em 1636 por Guilherme II, Príncipe de Orange, na cidade de Utrecht. É membro da *European League of Research Universities* e classificada entre as melhores universidades do mundo.
- 2) *Universiteit van Amsterdam*: a Universidade de Amsterdam foi fundada em 1632, pelo *Gemeente Amsterdam*, pelo “Conselho da Cidade de Amsterdam” e é a maior universidade da Holanda, localizada na capital do país, Amsterdam. É membro da *European League of Research Universities*, classificada entre as melhores universidades do mundo.
- 3) *Rijksuniversiteit Groningen*: a Universidade de Groningen foi fundada em 1614, pelo príncipe Maurício de Nassau. É membro da *European League of Research Universities*, classificada entre as melhores universidades do mundo.
- 4) *Universiteit Leiden*: a Universidade de Leiden é a mais antiga universidade da Holanda e uma das mais antigas do mundo. Foi fundada em 1575, por Guilherme I, Príncipe de Orange, para promover a educação e a pesquisa protestantes. Possui sete faculdades e está localizada em Leiden e Haia. É membro da *European League of Research Universities*, classificada entre as melhores universidades do mundo.
- 5) *Radboud University*: a Universidade de Radboud foi fundada em em 1923, na cidade de Nijmegen. A universidade leva o nome do bispo Radboud, conhecido no século IX pelo apoio aos povos marginalizados.

- 6) *Erasmus Universiteit Rotterdam*: a Universidade de Erasmus de Rotterdam foi fundada em em 1913, na cidade de Rotterdam. Seu nome é uma homenagem a Desiderius Erasmus Roterodamus, um humanista e teólogo do século XVI, importante líder intelectual do Renascimento.
- 7) *Technische Universiteit Delft*: a Universidade Técnica de Delft é a maior e mais antiga universidade politécnica da Holanda. Foi fundada em 1842, na cidade de Delft. Formou a maior parte dos profissionais empenhados na realização do “Plano Delta”, um sistema de defesa contra a subida do nível do mar, na Holanda.
- 8) *Wageningen University & Research*: a Wageningen Universidade e Pesquisa foi fundada em 1918, na cidade de *Wageningen*, e se destaca pelas pesquisas realizadas na área da agricultura.
- 9) *Universiteit Twente*: a Universidade de Twente, localizada na cidade de Enschede, foi fundada em 1961.
- 10) *Tilburg University*: a Universidade de Tilburg foi fundada em 1927 como Colégio Comercial Católico Romano, na cidade de Tilburg
- 11) *Vrije Universiteit Amsterdam*: a Universidade Livre de Amsterdam foi fundada em 1880 por um grupo de cristãos protestantes, em Amsterdam.
- 12) *Eindhoven University of Technology*: a Universidade de Tecnologia de Eindhoven foi fundada em 1956, na cidade de Eindhoven, como uma escola de engenharia.
- 13) *Maastricht University*: a Universidade de Maastricht foi fundada em 1976, na cidade de Maastricht, sendo a mais jovem das universidades públicas da Holanda.

Figura 11 – As universidades holandesas selecionadas

## Universities of The Netherlands



Fonte: Holanda. Universiteiten van Netherland

Não foram localizadas políticas para a gestão de dados de pesquisa na Universidade de Tecnologia de Eindhoven e na Universidade de Maastricht. Portanto, foram excluídas do escopo desta pesquisa. É importante ressaltar que, embora as políticas não tenham sido encontradas, ambas as instituições oferecem diretrizes e apoio consistentes para sua comunidade.

Na UE, os pesquisadores que recebem financiamento do programa *Horizon Europe* vigência 2014-2020, tiveram que gerenciar os dados de pesquisa, disponibilizando-os conforme os princípios FAIR. O programa *Horizon Europe* 2021-2027 reitera estas mesmas normas. A Holanda recebe financiamento para realizar pesquisa e desenvolver soluções para promover a Ciência Aberta e, além do apoio da UE, o governo holandês investe em



infraestrutura e ações, desenvolvimento de repositórios de dados abertos e políticas de Acesso Aberto.

O Quadro 15 apresenta alguns documentos e organizações encontradas que pertencem à rede de incentivo à promoção e apoio à Ciência Aberta da Holanda.

Quadro 15 – O suporte para a Ciência Aberta, na Holanda

Documentos e Organizações	Características e Objetivos
<b>Organizações</b>	
<i>All European Academies (ALLEA)</i>	Rede de academias nacionais de toda a Europa para promover a pesquisa e a inovação na Europa e representar as academias europeias em nível internacional.
<i>Association of Universities in the Netherlands (VSNU)</i>	Representa as universidades no governo, parlamento, organizações governamentais e cívicas promovendo debates, desenvolvendo e divulgando posições comuns para melhorar a reputação do setor universitário e escolhas estratégicas para fortalecer a educação científica e a pesquisa na Holanda.
<i>Data Archiving and Networked Services (DANS)</i>	Instituição holandesa especializada em gerenciamento, preservação, compartilhamento e acesso a dados de pesquisa.
<i>Dutch Research Council (NWO)</i>	Financia pesquisas em universidades e institutos e orienta o curso da ciência holandesa por meio de subsídios e programas de pesquisa.
<i>European Open Science Cloud (EOSC)</i>	Organização da Comissão Europeia para desenvolver infraestrutura que forneça serviços para promover as práticas da Ciência Aberta.
<i>GO FAIR</i>	Rede holandesa com estratégias internacionais destinadas a impulsionar o desenvolvimento de serviços e dados FAIR.
<i>Royal Netherlands Academy of Arts and Sciences (KNAW)</i>	Organização dedicada ao avanço da ciência e da literatura na Holanda. Financia e realiza pesquisas; assessora o governo holandês e outras instituições sobre questões científicas e de política científica; promove atividades como conferências, publicações e prêmios, cursos e treinamentos.
<i>National Academic Research and Collaborations Information System (NARCIS)</i>	Portal holandês para buscar informações sobre pesquisadores e pesquisas em Acesso Aberto e agregar dados de cerca de 30 repositórios institucionais. <i>Foi desativado em julho de 2023 e substituído pelo PURE.</i>
<i>National Coordination Point Research Data Management (LCRDM)</i>	Rede nacional de especialistas em GDP criada para facilitar os serviços de gerenciamento de dados e apoiar políticas e soluções em nível nacional.
<i>OpenAIRE – Infraestrutura para o Acesso Aberto de Pesquisa na Europa</i>	Infraestrutura europeia para a promoção do acesso aberto a pesquisas, criado pela Comissão Europeia que visa aumentar a visibilidade e acessibilidade da pesquisa e facilitar a reutilização de dados de pesquisa.
<i>OpenAIRE Connect</i>	Portal que permite aos pesquisadores encontrar e reutilizar dados de pesquisa de toda a Europa.
<i>OpenAIRE Graph</i>	Recurso do OpenAIRE para criar uma infraestrutura de dados abertos interconectada, permitindo a descoberta e o link de

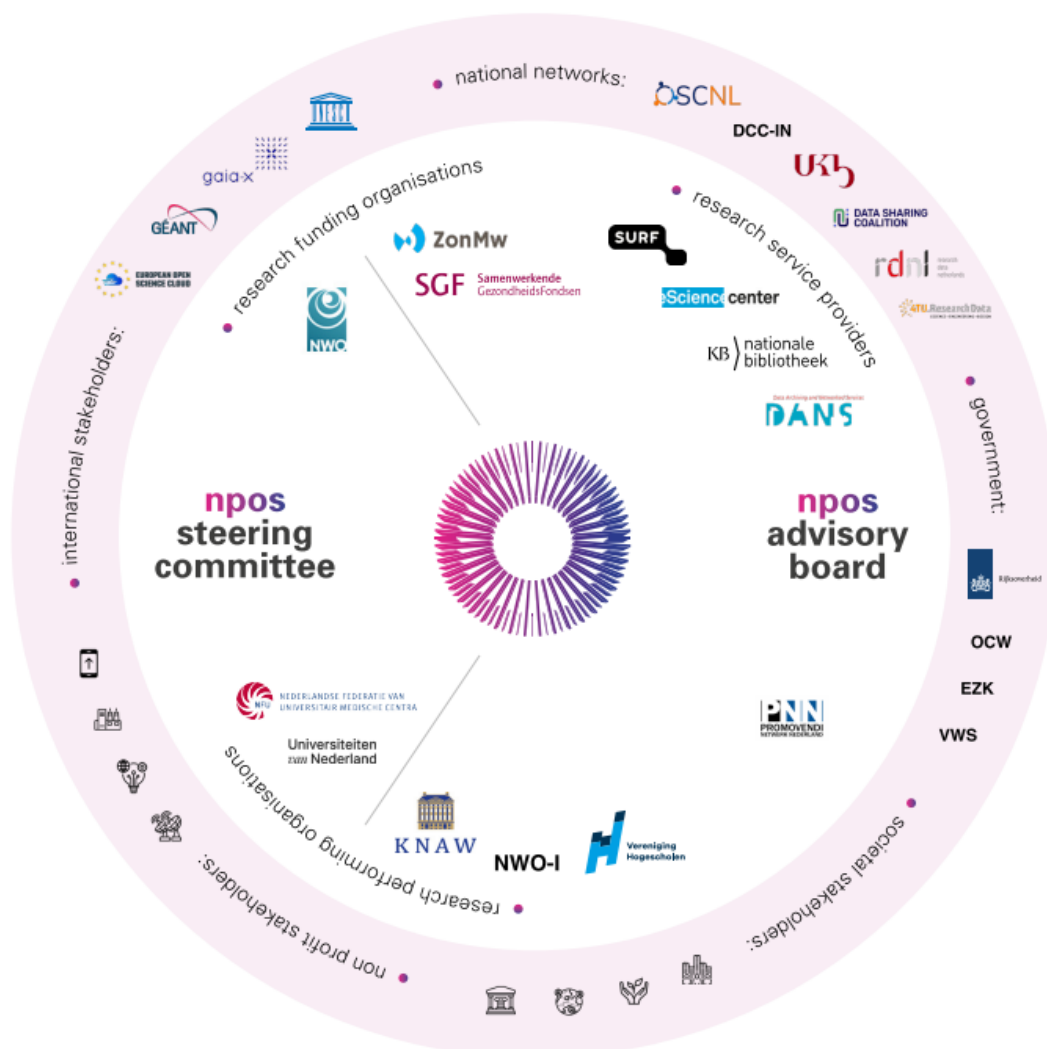
	informações relacionadas à pesquisa, como publicações, conjuntos de dados, projetos e autores, por meio de um modelo de dados interoperável e gráficos semânticos.
<i>National Open Science Program</i> (NPOS)	Rede de organizações holandesas que trabalham para promover, coordenar a transição e divulgar a importância da Ciência Aberta entre diversos <i>stakeholders</i> .
<i>Netherlands Research Portal</i> (PURE)	Cobre publicações, conjuntos de dados e projetos de universidades e institutos de pesquisa na Holanda com o objetivo de conectar a pesquisa holandesa a pesquisadores e <i>stakeholders</i> . Apresenta resultados e projetos de pesquisas holandesas. <i>Esse sistema substituiu o portal NARCIS.</i>
<i>Open Science 2023 in The Netherlands: NPOS2030 Ambition Document</i>	Diretrizes e metas para a promoção da Ciência Aberta na Holanda com princípios orientadores que fundamentam o programa, os objetivos e os requisitos; descreve as linhas de ação por período.
<i>Netherlands Code of Conduct for Research Integrity</i> (2018)	Conjunto de princípios e diretrizes que visam garantir a integridade da pesquisa na Holanda, guiados por cinco princípios: honestidade, escrupulosidade, transparência, independência e responsabilidade.
<i>Guidelines on Implementation of Open Access to Scientific Publications and Research Data</i>	Conjunto de orientações que visam promover o Acesso Aberto e o reuso dos dados de pesquisa da UE, em projetos apoiados pelo <i>European Research Council</i> no âmbito do Horizonte 2020.
<i>European Code of Conduct for Integrity of Research</i> (2023)	Estabelece princípios e diretrizes éticas para garantir a integridade e a qualidade da pesquisa científica na Europa.
<i>General Data Protection Regulation</i> (RGPD)	Estabelece a proteção e os dados pessoais dentro da União Europeia.
<i>Dutch Copyright Law – Auteurswet</i>	Protege os direitos autorais de obras de literatura, ciência e arte a partir do momento em que a obra é criada.
<i>Dutch Data Prize</i>	Premiação concedida ao pesquisador ou à equipe que fez uma contribuição especial para a ciência, tornando os dados de pesquisa acessíveis para pesquisas novas ou adicionais.

Fonte: Elaborado pela autora, 2023.

A menção de alguns dos programas que dão suporte no país foi considerada necessária para demonstrar o que leva as IES holandesas a serem tão comprometidas com a Ciência Aberta e, conseqüentemente, grandes promotoras de serviços que viabilizam a GDP.

Em tempo, destaca-se que todas as universidades holandesas selecionadas são públicas, e o maior financiador de pesquisas no país, a NWO, é vinculado ao governo federal. Além disso, o NPOS publicou o *Open Science 2030 in the Netherlands* que tem como objetivo facilitar a colaboração de todos os *stakeholders* nacionais que cooperam para uma transição “da ciência como é para a ciência como será” (NPOS, 2022). Na Figura 12, apresenta-se a estrutura proposta pelo NPOS.

Figura 12 – Aspirações do Open Science 2030 in the Netherlands



Fonte: NPOS, 2022.

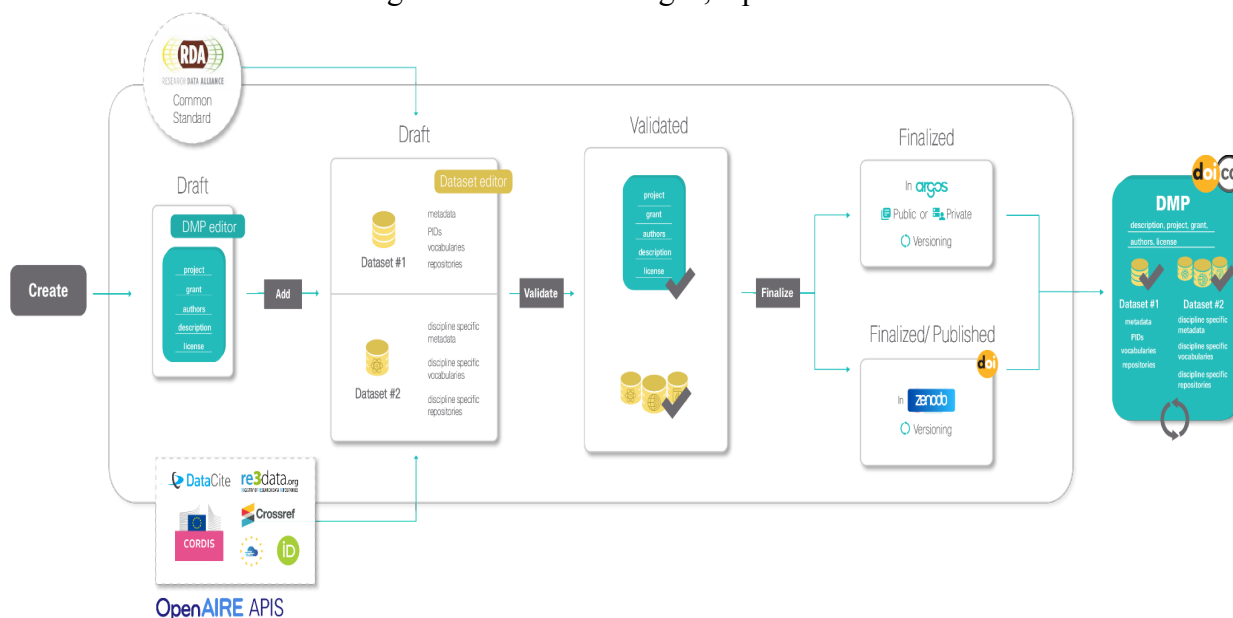
A proposta da NPOS, apresentada na Figura 12 descreve a pretensão da rede de trabalhar para “uma colaboração estreita entre instituições de conhecimento, governo, indústria e cidadãos para fortalecer a posição internacional da ciência holandesa e otimizar os processos de criação, compartilhamento e comunicação de conhecimento para o benefício da sociedade” (NPOS, 2023, p. 21). Além de aplicar recursos financeiros na educação básica e superior no país, o governo Holandês dispõe de recursos de tecnologia e documentos nacionais e do bloco europeu. Dentre as normativas estão obrigatoriedade de: a) Garantir a

gestão dos dados de pesquisa; b) Apresentar um Plano de Gestão de Dados; c) Permitir a publicação com Acesso Aberto; d) Apoiar infraestruturas para a gestão e a proteção de dados; e) Compartilhar os dados quando possível e protegê-los, quando necessário.

Além dos programas e incentivos já existentes, a OpenAIRE lançou, em 2023, uma nova ferramenta para elaboração de PGD: o **Argos**. Esta ferramenta apresenta uma série de vantagens em comparação com outros planos de gestão de dados atuais, que incluem a disponibilidade de modelos personalizáveis, integração com repositórios de dados, facilitando o armazenamento e a promoção da colaboração entre pesquisadores para criar e gerenciar PGD mais complexos.

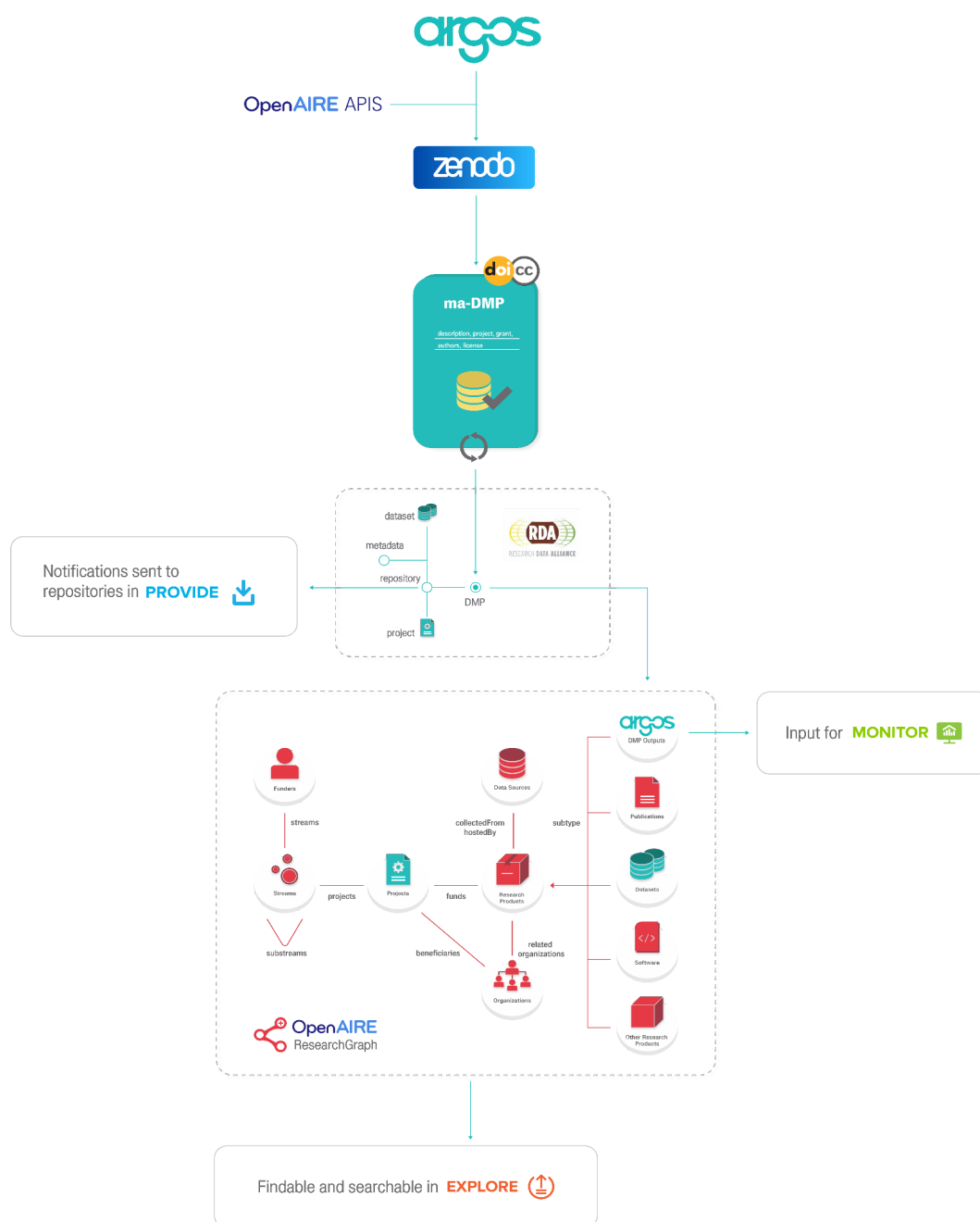
As funcionalidades do Argos possibilitam a integração com outros planos de gestão correlatos já arquivados. As Figuras 13 e 14 fornecem uma representação visual dos fluxos e conexões facilitados pelo Argos. É relevante observar que, como uma ferramenta nova, não foi mencionado nas políticas para GDP consultadas. Dado o conjunto de vantagens que oferece, supõe-se que, quando estas políticas forem atualizadas, o Argos possa ser adotado como padrão, uma vez que está alinhado com os objetivos do NPOS.

Figura 13 – Sistema Argos, OpenAIRE



Fonte: OpenAIRE [202-?].

Figura 14 – Ecossistema do Argos no OpenAIRE



Fonte: OpenAIRE [202-?].

Na subseção 4.5 serão apresentados os resultados da análise das políticas para GDP, conforme o processo metodológico desta pesquisa.

#### 4.5 AS POLÍTICAS DE GESTÃO DE DADOS DE PESQUISA NAS IES DA HOLANDA

Nesta fase, foram realizados levantamento, análise e relacionamento entre políticas para GDP das universidades holandesas. O resultado desse esforço foi um mapeamento completo das normativas das onze instituições, com identificação dos elementos comuns e distintos, bem como dos pontos de convergência e divergência. Esse mapeamento representou um passo importante para conhecer e compreender os elementos necessários para o gerenciamento dos dados, antes mesmo de gerá-los/coletá-los, possibilitando que os potenciais responsáveis saibam o que é necessário fazer para construir uma política.

A compilação dos títulos dos documentos foi evidenciada no Quadro 16 para exemplificar a diversidade de denominações que uma política pode apresentar.

Quadro 16 – As políticas das IES holandesas

Universidade	Título da política	Ano
Universidade de Tilburg	Regulamento de Gerenciamento de Dados de Pesquisa	2020
Universidade Livre de Amsterdam	Política de gerenciamento de dados de pesquisa	2020
Universidade de Wageningen	Política de dados na WUR	2020
Universidade de Utrecht	Política de gerenciamento de dados de pesquisa	2019
Universidade de Amsterdam	Gerenciamento de Dados de Pesquisa UvA	2019
Universidade de Groningen	Política de dados de pesquisa da UG	2021
Universidade de Leiden	Regulamento de gerenciamento de dados da Universidade de Leiden	2021
Universidade de Radboud	Política da universidade para armazenar e gerenciar dados de pesquisa	2021
Universidade de Erasmus de Rotterdam	Política de Gerenciamento de Dados de Pesquisa da Universidade de Erasmus de Rotterdam	2020
Universidade de Twente	Política de gerenciamento de dados de pesquisa da Universidade de Twente	2018
Universidade Técnica de Delft	Política de estrutura de dados de pesquisa da TU Delft	2018

Fonte: Elaborado pela autora, 2023.

Primeiramente, destaca-se que todas as políticas para GDP da Holanda reafirmam que os pesquisadores devem permitir o reuso dos dados, salvo nos casos de impedimentos previstos. Outras incidências significativas de elementos, com 90,9% foram em questões que se referem à estrutura geral do documento, como a contextualização da Ciência Aberta e da

abertura dos dados de pesquisa nacional e internacionalmente, a preocupação com a capacitação de todos os *stakeholders*.

No quadro 17, são apresentados os elementos encontrados nas políticas com as porcentagens indicando a incidência nos documentos. Entretanto, não se pode afirmar que as porcentagens apresentadas são uma representação exata da realidade, pois outras universidades podem ter adotado práticas iguais ou similares, mas que não estavam explícitas na política institucional.

Quadro 17 – Os elementos para GDP na Holanda: ocorrências e incidências

Incidência dos elementos nas políticas		
1	Introdução contextualizando a Ciência Aberta e a GDP	90,90%
2	Capacitação e orientação	90,90%
3	Monitoramento/fiscalização	90,90%
4	Princípios FAIR como meta	90,90%
5	Indica quem é considerado <i>stakeholders</i> da IES	90,90%
6	Comissão para elaborar a GDP	81,80%
7	Legislação e documentos relacionados à ética e integridade	81,80%
8	Protocolos específicos para Faculdades/Institutos	81,80%
9	Subordinação à política	81,80%
10	Função e responsabilidade de cada <i>stakeholder</i>	81,80%
11	Pós-graduação devem gerenciar os dados	63,60%
12	Biblioteca como parte da equipe serviços	54,50%
13	Gerenciamento de dados físicos	54,50%
14	Administrador de Dados	54,50%
15	Revisão e atualização periódica	45,50%
16	Glossário	45,50%
17	Obrigatoriedade de PGD de candidatos às vagas de pós-graduação	36,40%
18	Disciplina na grade curricular na IES	36,40%
19	Propriedade intelectual	27,30%
20	Representante da CI na Comissão	18,20%
21	Proprietário/dono dos dados	18,20%
22	Irregularidades na coleta/geração e/ou processamento	9,10%
23	Atribuições do Administrador de Dados	9,10%
A gestão dos dados		
24	Todos os pesquisadores devem permitir o reuso	100%
25	Responsável pela GDP	90,90%
26	Infraestrutura interna para armazenamento	90,90%

27	Prazo padronizado para armazenar e arquivar	72,70%
28	Proteção contra invasões, roubos e corrupção dos dados	72,70%
29	PGD é obrigatório para todos os projetos	63,60%
30	Responsável pela integridade dos dados	63,60%
31	Prazo mínimo para o armazenamento	63,60%
32	GDP quando o pesquisador perde o vínculo com a IES	63,60%
33	PGD institucional	54,50%
	Prazo mínimo para o arquivamento	54,50%
34	Prazo mínimo para o arquivamento do PGD	54,50%
35	Obrigatoriedade de formatos abertos	54,50%
36	Quando os dados devem ser arquivados	54,50%
37	Identificadores persistentes	54,50%
38	Atualizar o PGD durante a pesquisa	45,50%
39	Classificação dos dados no PGD	36,40%
40	Tipos de dados que devem ser gerenciados	36,40%
41	Responsabilidade pelos custos	36,40%
42	Backup	36,40%
43	GDP em pesquisas colaborativas	36,40%
44	Arquivamento em RD externos	36,40%
45	Licenças de uso	36,40%
46	Os metadados públicos para dados de acesso fechado/restrito	27,30%
47	O segurança reforçada para os dados confidenciais	27,30%
48	RD institucional	27,30%
49	Rastreabilidade dos dados para identificação de indivíduos	27,30%
50	Citação e referência	27,30%
51	Prazo e critérios para embargos	9,10%
52	Compartilhamento com outros pesquisadores antes da conclusão/publicação dos resultados	9,10%
53	Prazo para armazenar os brutos s	9,10%
54	Compartilhamento/permuta de dados entre pesquisadores internos e externos, durante a pesquisa	9,10%
55	Alteração de dados após a pesquisa	9,10%

Fonte: Elaborado pela autora, 2023.

A exposição dos elementos derivados das políticas holandesas, conforme o Quadro 17, não adere a uma sequência determinada para orientar as decisões, assim como não delinea os passos para a implementação de serviços, visto que a pesquisa não tem como finalidade servir de um manual ou guia. É pertinente considerar que a disposição das informações serve



meramente como uma representação dos elementos identificados, visando aprimorar a compreensão do conteúdo.

As porcentagens médias focaram em elementos diversos, tanto de estrutura do documento quanto no gerenciamento dos dados em si, tais como a revisão e atualização periódica da política, o glossário, decisões relacionadas aos dados e pesquisas realizadas em pós-graduação, dentre outras questões relevantes.

Os elementos que apresentam uma incidência de 9,10% revelam que o componente foi abordado por apenas uma IES. Embora abordados de maneira isolada, esses elementos ainda representam aspectos significativos que merecem atenção ao se elaborar políticas nas IES brasileiras. Por exemplo, podem-se mencionar as decisões e os procedimentos adotados diante de eventuais irregularidades identificadas nos processos de coleta, geração e/ou processamento dos dados.

Há ocorrências de pesquisas despublicadas e de desvinculação dos responsáveis pela pesquisa de suas respectivas instituições, quando os dados foram inadequadamente processados comprometeram a confiabilidade nos resultados das pesquisas. Assim, normativas para essas situações podem evitar erros, e o constrangimento público aos pesquisadores e à instituição. Outro item importante é informar quais são as atribuições do Administrador de Dados, que raramente são claras no contexto das políticas.

## 5 PROPOSTAS PARA A CONSTRUÇÃO DE UMA POLÍTICA INSTITUCIONAL

Nesta seção, são apresentados os elementos extraídos das políticas para Gestão de Dados de Pesquisa das instituições de ensino superior holandesas e alternativas para abordagens.

O trabalho para promover a gestão dos dados é desafiador, porém necessário e benéfico. A construção do documento requer uma organização e pesquisa minuciosa, assegurando que sua implementação seja possível, conforme as necessidades e possibilidades da IES. Assim como o planejamento para a gestão dos dados deve ser atualizado para acompanhar o andamento da pesquisa, a política para gestão de dados é um documento em evolução e deve ser revisado e atualizado continuamente.

Para melhor compreensão da proposta apresentada no Quadro 18, foi necessário identificar os possíveis responsáveis pelo processo da construção de uma política. Os "Responsáveis/Stakeholders" apresentados no quadro, referem-se à:

a) *Comissão*: um grupo de pessoas designadas para estudar um assunto e realizar uma tarefa específica, de interesse geral. No contexto desta pesquisa, recomenda-se que sejam nomeados um ou mais representantes de cada categoria envolvida no processo da GDP;

b) *Reitor*: cargo administrativo mais alto na IES, o dirigente máximo responsável por liderar e gerenciar todos os aspectos da instituição para seu bom funcionamento acadêmico, administrativo e financeiro. Pode ser identificado por outras denominações em algumas IES;

c) *Pró-reitores*: indivíduos que ocupam cargos de confiança do reitor, responsáveis pela direção de uma determinada área de atuação, como a pesquisa, pós-graduação, administração, extensão, graduação e outros;

d) *Diretores de faculdades/institutos*: são os responsáveis pela gestão das respectivas unidades acadêmicas de uma instituição de ensino superior, geralmente de uma área do conhecimento específica;

e) *Coordenadores de cursos*: são os responsáveis pela gestão de um curso específico de graduação ou pós-graduação e supervisionam as atividades acadêmicas, administrativas e financeiras do curso;

f) *chefes de departamentos*: são os responsáveis pela gestão de um departamento acadêmico e devem supervisionar as atividades acadêmicas, administrativas e financeiras do

departamento;

g) *chefes de laboratórios*: profissionais que possuem um cargo de liderança e supervisão das atividades do laboratório, como o planejamento, execução e análises da pesquisa;

h) *Pesquisadores individuais*: são profissionais que realizam pesquisas de forma independente, sem fazer parte de um grupo ou equipe de pesquisa;

i) *Setor de tecnologia de informação*: setor responsável pela gestão e suporte aos assuntos relacionados à infraestrutura de tecnologia e sistemas de informação, redes e comunicação da instituição;

j) *Biblioteca universitária*: setor responsável por organizar, proteger e fornecer acesso a recursos de informação e aprendizagem à comunidade acadêmica e sociedade em geral, além de prestar serviços de orientação e capacitação.

k) *Equipe de pesquisa*: grupo de pessoas que trabalham em conjunto para realizar a pesquisa. A equipe pode ser interna, composta por pessoas exclusivamente da IES ou colaborativa, com cientistas de outras instituições;

O Quadro 18 foi organizado na seguinte forma:

1) Na primeira coluna: “**Elementos**”, apresenta-se os elementos e atividades necessárias para construir a política.

2) Na segunda coluna, “**Responsáveis/Stakeholders**”, apresenta-se os possíveis agentes para executar as ações que, majoritariamente, são debatidas e acordadas em conjunto.

3) Na terceira coluna, “**Abordagens**”, propõe-se algumas possibilidades para abordar cada elemento ou ações para compor a política.

Os elementos podem ser abordados de diferentes formas, pois cada IES decide o mais viável para a sua realidade, necessidade e interesse. A intenção é apresentar alternativas, e não escolhas engessadas. Não é possível apresentar um modelo com representações absolutas, uma vez que o processo decisório deve ser flexível e adaptável.

Os passos sugeridos no Quadro 18 foram construídos a partir das informações obtidas nas políticas para GDP das universidades da Holanda. Contudo, é possível que outros itens importantes tenham sido omitidos, considerando a diversidade e complexidade no processo da gestão.

Quadro 18 – Políticas para GDP em IES brasileiras: diretrizes

<i>1 – Passos para iniciar os trabalhos - Atividade e/ou Processos</i>		
<i>Elementos</i>	<i>Responsáveis/Stakeholders</i>	<i>Abordagens/Planejamento</i>
Criar a Comissão	<input type="checkbox"/> Reitoria	<input type="checkbox"/> nomear os membros da Comissão; <input type="checkbox"/> nomear o líder da Comissão; <input type="checkbox"/> designar membros de setores especializados e professores de diferentes áreas de pesquisa; <input type="checkbox"/> submeter a proposta de política para análise dos conselhos pertinentes; <input type="checkbox"/> encaminhar as alterações propostas à Comissão; <input type="checkbox"/> aprovar e publicar o documento nos canais de comunicação institucional; <input type="checkbox"/> estabelecer uma estrutura de governança; <input type="checkbox"/> fazer valer as normativas da política;
Atividades da Comissão	<input type="checkbox"/> Comissão	<input type="checkbox"/> elaborar estratégias e cronogramas para execução das atividades; <input type="checkbox"/> levantar materiais de informação para adquirir maior conhecimento sobre o assunto; <input type="checkbox"/> elencar os serviços, termos e conceitos relacionados; <input type="checkbox"/> levantar a legislação e documentos afins (ética e integridade da pesquisa); <input type="checkbox"/> elencar os elementos que compõem o processo da GDP e respectivos procedimentos; <input type="checkbox"/> realizar um diagnóstico das necessidades e interesses da instituição para cada elemento; <input type="checkbox"/> definir as abordagens dos elementos e procedimentos na instituição; <input type="checkbox"/> promover consultas internas e ações colaborativas internamente para o envolvimento da comunidade; <input type="checkbox"/> permitir que aos stakeholders não membros da comissão possam contribuir na elaboração do documento; <input type="checkbox"/> propor cronograma para revisão e atualizar os termos da política, periodicamente;

Ações/decisões gerais	<input type="checkbox"/> Reitoria <input type="checkbox"/> Comissão <input type="checkbox"/> Pró-reitorias <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Prog. de pós-graduação	<input type="checkbox"/> decidir se cada Faculdades/Institutos poderá criar protocolos individuais; <input type="checkbox"/> definir formas de fiscalização e monitoramento da GDP de acordo com a política; <input type="checkbox"/> estabelecer procedimentos nos casos de suspeitas de violação no processo coleta/geração e na gestão dos dados; <input type="checkbox"/> desenvolver formas para, reconhecer, valorizar e premiar os pesquisadores por contribuições;
-----------------------	---	---

*2 – Informações para a Introdução do texto da política para gestão de dados de pesquisa*

<i>Elementos</i>	<i>Responsáveis/Stakeholder</i>	<i>Abordagens/Planejamento</i>
Escopo do texto: a Introdução	<input type="checkbox"/> Comissão	<input type="checkbox"/> contextualizar a Ciência Aberta e a GDP na instituição; <input type="checkbox"/> esclarecer sobre o reuso dos dados da instituição; <input type="checkbox"/> destacar a importância da gestão de dados de pesquisa para a instituição; <input type="checkbox"/> enfatizar sobre as normativas imperativas da legislação brasileira; <input type="checkbox"/> expor os métodos para construção da política para GDP; <input type="checkbox"/> regular as normativas política como imperativas em toda a instituição; <input type="checkbox"/> listar os stakeholders institucionais subordinados à política; <input type="checkbox"/> descrever as responsabilidades e as funções de cada stakeholder; <input type="checkbox"/> apresentar glossário; <input type="checkbox"/> adotar os Princípios FAIR; <input type="checkbox"/> salientar a escolha por arquivos em formatos abertos; <input type="checkbox"/> apresentar alternativas de formatos abertos; <input type="checkbox"/> enfatizar a necessidade da citação e referência; <input type="checkbox"/> apresentar padrões para citação e referência; <input type="checkbox"/> elencar a legislação e documentos correlatos (ética e integridade); <input type="checkbox"/> recomendar a realização de uma busca em RD para verificar a disponibilidade de dados que possam ser reutilizado;

		<input type="checkbox"/> informar sobre a obrigatoriedade de citar e referenciar os dados reusados; <input type="checkbox"/> determinar se todos os pesquisadores devem permitir o reuso; <input type="checkbox"/> instruir sobre a anonimização e pseudonimização de dados; <input type="checkbox"/> indicar o responsável por resolver os casos omissos na política;
Administração dos dados de pesquisa	<input type="checkbox"/> Pró-reitorias <input type="checkbox"/> Programa de pós-graduação <input type="checkbox"/> Faculdades/Institutos	<input type="checkbox"/> nomear um administrador dos dados geral; <input type="checkbox"/> nomear um administrador de dados para cada disciplina; <input type="checkbox"/> definir os critérios para nomear o administrador de dados; <input type="checkbox"/> definir as atribuições do administrador de dados;
Serviços de suporte e orientação	<input type="checkbox"/> Progr. de pós-graduação <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Departamentos <input type="checkbox"/> Laboratórios <input type="checkbox"/> Setor de TI <input type="checkbox"/> Biblioteca	<input type="checkbox"/> oferecer cursos, workshops e treinamentos; <input type="checkbox"/> indicar os responsáveis pelas ações de suporte; <input type="checkbox"/> criar apenas um setor centralizado; <input type="checkbox"/> criar canais especializados para cada disciplina; <input type="checkbox"/> oferecer uma disciplina na grade curricular; <input type="checkbox"/> manter website informativo (Na Holanda estes serviços são hospedados nos sites das bibliotecas);
Pós-graduação	<input type="checkbox"/> Pró-reitorias <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Progr. de pós-graduação	<input type="checkbox"/> os dados gerados/coletados devem ser gerenciados; <input type="checkbox"/> o PGD é pré-requisito para concorrer às vagas; <input type="checkbox"/> responsável pela gestão dos dados da pós-graduação <input type="checkbox"/> responsável pela qualidade e integridade dos dados; <input type="checkbox"/> obrigatoriedade de cursar a disciplina de GDP ou outras formas de capacitação ofertados pela instituição;
Tipos e formatos de dados a serem gerenciados	<input type="checkbox"/> Progr. de pós-graduação <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Departamentos	<input type="checkbox"/> todos os dados gerados/coletados devem ser gerenciados; <input type="checkbox"/> apenas alguns tipos de dados poderão ser gerenciados (nem todos os dados têm relevância para uma determinada área); <input type="checkbox"/> dados em formatos físicos;

		<input type="checkbox"/> dados digitais, apenas; <input type="checkbox"/> dados digitais e digitalizados; <input type="checkbox"/> dados digitais, digitalizados e físicos; <input type="checkbox"/> definido nos protocolos individuais das Faculdades/Institutos;
--	--	--

### 3 – O planejamento no processo da gestão dos dados de pesquisa - Atividade e/ou Processo

<i>Elementos</i>	<i>Responsáveis/Stakeholder</i>	<i>Abordagens/Planejamento</i>
Responsável pela gestão dos dados	<input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Departamentos <input type="checkbox"/> Laboratórios	<input type="checkbox"/> a instituição; <input type="checkbox"/> a gestão é compartilhada (pesquisadores, equipe de TI, e outros) <input type="checkbox"/> pesquisador principal/líder da pesquisa; <input type="checkbox"/> o pesquisador poderá atribuir a função a terceiros; <input type="checkbox"/> indicar o responsável quando o gestor dos dados perde o vínculo institucional;
Proprietário dos dados	<input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Pesquisador individual	<input type="checkbox"/> pertencem à instituição; <input type="checkbox"/> pertencem aos pesquisadores; <input type="checkbox"/> deve ser compartilhada entre os pesquisadores e a instituição; <input type="checkbox"/> é definida conforme o edital do financiador; <input type="checkbox"/> é definida por meio de acordos entre as partes (a instituição e/ou o financiador); <input type="checkbox"/> criar protocolos institucionais para firmar acordos; <input type="checkbox"/> responsável analisar e aprovar os acordos; <input type="checkbox"/> procedimentos quando houver conflitos de interesses;
Fiscalização	<input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento <input type="checkbox"/> Chefes de laboratórios <input type="checkbox"/> Administrador de dados	<input type="checkbox"/> verificar se os dados estão/foram gerenciados conforme o PGD; <input type="checkbox"/> verificar o cumprimento das normativas da política para GDP; <input type="checkbox"/> indicar um responsável pela fiscalização; <input type="checkbox"/> procedimentos em caso de irregularidades na GDP; <input type="checkbox"/> procedimentos quando constatado falhas na coleta/geração e/ou processamento dos dados;

Plano de gestão de dados	<input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento	<input type="checkbox"/> todos os projetos de pesquisa devem apresentar um PGD; <input type="checkbox"/> a instituição terá seu próprio modelo; <input type="checkbox"/> poderá ser utilizado qualquer modelo de PGD; <input type="checkbox"/> poderão ser utilizados apenas modelos específicos; <input type="checkbox"/> o periódico/editora terá primazia na escolha do PGD; <input type="checkbox"/> o financiador terá primazia quanto as escolha do PGD;
Responsável pelos custos da GDP	<input type="checkbox"/> Reitoria <input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento <input type="checkbox"/> Agências de fomento <input type="checkbox"/> Pesquisadores	<input type="checkbox"/> a instituição; <input type="checkbox"/> a instituição, parcialmente; <input type="checkbox"/> o financiador da pesquisa; <input type="checkbox"/> o pesquisador; <input type="checkbox"/> o pesquisador deverá buscar meios;
Armazenamento	<input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento	<input type="checkbox"/> a instituição irá disponibilizar sua infraestrutura tecnológica; <input type="checkbox"/> haverá limite de espaço em disco, por pesquisador; <input type="checkbox"/> sistema segurança reforçada para dados confidenciais (chaves, criptografia); <input type="checkbox"/> sistema de segurança contra roubos, invasões e corrupção para todos os dados; <input type="checkbox"/> permitir o armazenamento em sistemas externos; <input type="checkbox"/> definir o prazo para armazenar os dados brutos; <input type="checkbox"/> definir o prazo para armazenar os dados processados; <input type="checkbox"/> definir o tempo mínimo e máximo em que permanecerão retidos; <input type="checkbox"/> o descarte poderá ser realizado assim que forem arquivados; <input type="checkbox"/> todas as versões de dados reprocessados serão mantidas; <input type="checkbox"/> apenas a versão final dos dados processados será mantida; <input type="checkbox"/> como e quando os dados pessoais e sensíveis serão destruídos, <input type="checkbox"/> indicar local para armazenar os dados físicos; <input type="checkbox"/> permitir o acesso e/ou reuso aos dados por pesquisadores internos que não façam parte da equipe;



		<input type="checkbox"/> permitir o acesso e/ou reuso de dados com pesquisadores externos, que não fazem parte da equipe;
Responsável pelo Backup	<input type="checkbox"/> Reitoria <input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento	<input type="checkbox"/> a instituição; <input type="checkbox"/> a responsabilidade é do pesquisador/equipe; <input type="checkbox"/> indicar onde, como e quando será realizado; <input type="checkbox"/> apenas a instituição poderá manter o backup; <input type="checkbox"/> prazo para mantê-los sistema; <input type="checkbox"/> os pesquisadores podem manter cópia em dispositivos particulares;
Arquivamento	<input type="checkbox"/> Reitoria <input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento <input type="checkbox"/> Administrador de dados	<input type="checkbox"/> após a publicação da pesquisa, definir o período limite para que o pesquisador arquivar os dados; <input type="checkbox"/> estabelecer o prazo mínimo para a manter dos dados arquivados; <input type="checkbox"/> estabelecer o prazo máximo para a manter dos dados arquivados; <input type="checkbox"/> os dados podem ser arquivados antes da finalização da pesquisa; <input type="checkbox"/> os dados podem ser arquivados antes da publicação dos resultados da pesquisa; <input type="checkbox"/> poderá ser concedido prazo maior para arquivamento; <input type="checkbox"/> determinar o local para arquivar os dados físicos; <input type="checkbox"/> solicitação de acessos fechados e restritos serão analisadas previamente; <input type="checkbox"/> a justificativa para proibição de acesso será disponibilizada publicamente; <input type="checkbox"/> os metadados devem ser públicos para todos os tipos de acesso; <input type="checkbox"/> determinar prazo máximo para o embargo; <input type="checkbox"/> critérios para acessar dados confidenciais; <input type="checkbox"/> identificadores persistentes para todos os dados; <input type="checkbox"/> identificadores persistentes serão atribuídos por uma organização específica; <input type="checkbox"/> o nível de acesso deve ser informado por meio de licenças de uso; <input type="checkbox"/> as licenças serão atribuir por uma organização específica; <input type="checkbox"/> arquivar os dados com sua documentação; <input type="checkbox"/> arquivar o PGD com os dados;

		<input type="checkbox"/> prazo de arquivamento do PGD deve ser igual ou maior que os dados; <input type="checkbox"/> dados devem ser arquivados em formatos abertos; <input type="checkbox"/> o uso de formatos fechados deve ser justificado; <input type="checkbox"/> critérios para permitir acesso a dados confidenciais; <input type="checkbox"/> autorizar a rastreabilidade de dados para identificar indivíduos; <input type="checkbox"/> arquivar apenas no RD institucional; <input type="checkbox"/> arquivamento em RD externo; <input type="checkbox"/> arquivamento em RD institucional e externo; <input type="checkbox"/> RD externo deve possuir certificação de qualidade;
Metadados	<input type="checkbox"/> Comissão <input type="checkbox"/> Faculdades/Institutos <input type="checkbox"/> Coordenadores de curso <input type="checkbox"/> Chefes de departamento	<input type="checkbox"/> serão padronizados pela instituição; <input type="checkbox"/> seguirão padrões externos nacionais e/ou internacionais; <input type="checkbox"/> metadados adicionais podem ser atribuídos pelo responsável pelo arquivamento, conforme os padrões do RD; <input type="checkbox"/> deve-se consultar o pesquisador sobre adições antes de arquivar os dados;
Promoção do acesso	<input type="checkbox"/> Bibliotecários <input type="checkbox"/> Equipe de TI <input type="checkbox"/> Arquivistas <input type="checkbox"/> Administrador de dados	<input type="checkbox"/> responsável por indexar os dados no RD; <input type="checkbox"/> descartar os dados que não se enquadram na definição de dados de pesquisa; <input type="checkbox"/> indexar os dados e com sua documentação no RD; <input type="checkbox"/> verificar se os metadados atribuídos estão de acordo com os padrões institucionais; <input type="checkbox"/> orientar o pesquisador a realizar buscas por dados em RD antes de elaborar o PGD; <input type="checkbox"/> tornar públicos os dados embargados após o período de retenção; <input type="checkbox"/> assegurar que os dados sigilosos e confidenciais serão devidamente protegidos contra acessos não autorizados; <input type="checkbox"/> deve-se justificar a proibição do reuso quando os dados não forem confidenciais; <input type="checkbox"/> descartar os dados quando completarem seu ciclo de vida.

É altamente recomendável que a Comissão promova um diálogo com a comunidade acadêmica e ações participativas, assegurando que as deliberações sejam realizadas de maneira participativa, buscando, na medida do possível, atender aos interesses gerais.

Quando se trata da nomeação do “Responsável pelos dados”, trata-se de uma responsabilidade de grande importância no processo de gestão. É necessário designar um pesquisador líder da pesquisa ou outro profissional especializado para responder pela qualidade e integridade dos dados, para ser uma referência como orientador e supervisor dos processos de geração/coleta, processamento, e outros aspectos relacionados à pesquisa. Além disso, o responsável pelos dados deve garantir que as pesquisas sejam conduzidas de acordo com princípios legais e éticos.

Sintetizando o processo para a construção de uma política a nível institucional, deve-se seguir uma ordem de ações que, primordialmente, parte do cargo de máxima liderança da instituição. As tratativas seguintes são realizadas para definir os detalhes específicos da política, destacando-se:

- a) nomear responsáveis para iniciar os trabalhos;
- b) relacionar os elementos e as práticas necessárias para a GDP;
- c) levantar legislação e documentos pertinentes;
- d) realizar um diagnóstico institucional de infraestrutura de pessoal e tecnologia;
- e) levantar os custos para a GDP;
- f) verificar se a instituições poderá assumir os custos totalmente ou parcialmente;
- g) analisar a disponibilidade infraestrutura tecnológica e de pessoal;
- h) desenvolver a política para a GDP;
- i) promover capacitação dos responsáveis;
- j) conscientizar sobre a importância e obrigatoriedade da GDP;
- k) manter canais para apoio e suporte contínuo;
- l) revisar o cenário e atualizar a política.

Levantar os elementos essenciais para a elaboração de uma política para GDP para as IES brasileiras foi um objetivo parcialmente alcançado. Ao longo da pesquisa, ao analisar as políticas das universidades na Holanda, observou-se que o país possui uma extensa rede de recursos tecnológicos e documentais em âmbito nacional e regional

(na União Europeia). Essa rede oferece suporte para procedimentos e estabelece alguns padrões que asseguram a funcionalidade e integração a sistemas nacionais e internacionais.

Os objetivos da pesquisa quanto a revisão dos conceitos foram alcançados e foi possível sintetizar as definições compiladas de diversas fontes. Quanto ao objetivo de visualizar o panorama nacional, foi identificado que há ações sendo desenvolvidas e a conclusão que emerge desta análise é que o movimento de adoção da GDP ainda está evoluindo lentamente, especialmente entre as agências de fomento. No entanto, nos institutos de pesquisa, observa-se um envolvimento significativo e consistente, com potencial para impulsionar o serviço no Brasil, dada a forte influência e a conexão direta com as IES. Além disso, há universidades engajadas que, mesmo iniciando o processo sob a força da Fapesp, já estão desempenhando um papel pioneiro na adoção da GDP em nível nacional.

Quanto à busca por exemplos positivos de políticas, os modelos consultados contribuíram significativamente para a construção do Quadro 18, que sintetiza todas as ações para alcançar o objetivo geral.

Com base nos resultados compilados, foram delineadas opções para atribuição de funções e responsabilidades, os elementos e abordagens possíveis para a gestão de dados de pesquisa. Essa representação teve como propósito facilitar a compreensão dos componentes envolvidos e das diversas abordagens possíveis na formulação de uma política institucional.

Apresenta-se como limitações deste estudo o fato de que as análises foram realizadas exclusivamente nas políticas institucionais para GDP de abrangência geral, com normativas para toda a IES. Não foram consultados os protocolos individuais de Faculdades/Institutos para identificar elementos que pudessem estar relacionados de forma mais específica com suas áreas, o que poderia contribuir para um gerenciamento mais abrangente e compartilhamento mais seguro. Consequentemente, existe a possibilidade de que elementos importantes tenham passado despercebidos e não tenham sido incluídos no Quadro 18. Essa lacuna pode impactar a representação completa da GDP nas diversas unidades das instituições, que têm o potencial de contribuir para a abrangência dos resultados deste estudo.

Futuras pesquisas podem ser realizadas para estudar os protocolos individuais, de áreas específicas do conhecimento, mas em diferentes IES para que se possa verificar como cada Faculdade/Instituto tem buscado gerenciar os dados de pesquisa.

## 5.1 ALGUMAS CONSIDERAÇÕES SOBRE A CIÊNCIA DA INFORMAÇÃO E OS DADOS DE PESQUISA

Embora as discussões em torno da importância dos dados tenham ganhado protagonismo nos últimos anos, pesquisadores da CI já estudavam as perspectivas conforme a realidade de sua época. Em 1994, Le Coadic (2004, p. 9) apontou que, “Em ciência da computação, chamamos de dados a representação convencional, após a codificação, de informações em um formulário que permite o processamento eletrônico”. Davenport (1998) definiu dados como simples observações sobre o estado do mundo, facilmente estruturado, facilmente obtido por máquina, frequentemente quantificado e facilmente transferível. Essas duas conceituações chamam a atenção por sua ênfase no processamento por meio da tecnologia. Em termos da “expansão da área da Ciência da Informação, destacam-se os avanços tecnológicos, de nova temática, como as bibliotecas virtuais/digitais e mineração de dados” (Pinheiro *et al.*, 2019, p. 150).

A CI promove a organização, a disponibilização e também a proteção dos recursos de informação. Logo, os dados de pesquisa, como recurso de informação, envolve a categoria. Os profissionais da área oferecem orientações e capacitação para estudantes, colaboradores, pesquisadores e outros interessados há milênios. Numa forma mais moderna para a pesquisa em rede, Sales (2014, p. 167) define que “[...] as demandas por profissionais da *e-science* recebem acolhidas dos cursos tradicionalmente vocacionados para tal como Biblioteconomia, Arquivologia e Ciência da Informação”.

Com os conhecimentos e as habilidades dos profissionais da CI, é possível estabelecer práticas eficazes de gestão de dados que promovam a transparência, a integridade e o uso responsável dos dados. Como salientado por Sales (2014, p. 18):

Investigar este novo fenômeno de geração de dados e informações no âmbito da ciência bem como propor novas formas de coletar, organizar, armazenar, recuperar estes dados e ainda atividades como interpretação, transmissão, transformação e uso da informação – que atualmente vem sendo denominada como reuso de informação – é papel da Ciência da Informação.

Embora alguns elementos da gestão de dados sejam exclusivamente realizados por profissionais e especialistas de outras áreas, como desenvolvedores e gerenciadores de TI, equipes de tecnologia de informação e administradores de dados, as áreas de atuação de bibliotecários e arquivistas são indispensáveis no processo da gestão.

Dentro das atividades da gestão da informação destaca-se a catalogação, que envolve o registro do objeto em um catálogo (*Repositórios*), com informações como título, autor, assunto, data, palavras-chave e outros elementos para descrever o conteúdo (agora, chamados de *Metadados*). A CI garante que os processos necessários para cuidar e proteger o conhecimento, permite que estes sejam encontrados, acessados e reutilizados. O trabalho dos profissionais que compõem a área é determinante para que as vantagens e benefícios garantidos à sociedade, à ciência e ao meio ambiente descritas nesta dissertação, sejam reais.

Sugerem-se estudos futuros sejam realizados para interligar nomenclaturas utilizadas para a GDP e as tradicionais da CI para promover e adaptar serviços nos espaços de informação. Isso pode melhorar a compreensão e facilitar a assimilação de novos termos e serviços.

Uma sugestão para futuras pesquisas é explorar a necessidade de sua reinvenção da biblioteca diante das atividades no contexto digital, onde o seu significado na sociedade seria de um espaço integrativo entre bibliotecário, arquivistas, museólogos e especialistas em informática, em uma abordagem multidisciplinar e colaborativa.

É fato que, de acordo com os dados coletados nesta pesquisa, 54,4% das bibliotecas holandesas assumem alguns serviços administrativos e apoio à comunidade acadêmica nos processos de gestão de dados de pesquisa. Com a obsolescência crescente de certos serviços, como a circulação de materiais físicos, as bibliotecas e bibliotecas precisam assumir novos papéis para preservar sua essência milenar na sociedade.

*Há Futuro para as Bibliotecas de Pesquisa no Ambiente de eScience?*, questionam Sales e Sayão (2015). A mudança de paradigma questiona o significado das bibliotecas tradicionais. A resposta a essa pergunta é importante para compreender a abrangência e o impacto das bibliotecas na era digital, em que o gerenciamento de dados de pesquisa desempenha um papel cada vez maior nas academias de pesquisa.

Estas questões não preveem a morte das bibliotecas como eram e nem como são. Pelo contrário. Ao longo dos milênios, como santuários do conhecimento, bibliotecas foram muitas vezes percebidas como ameaças por ideologias lideradas por políticos, grupos religiosos e extremistas. Foram desvalorizadas, invadidas, destruídas e até consumidas pelo fogo. Contudo, em vão são tais ações, pois as bibliotecas eternamente ressurgem, como fênix da sabedoria.

## 6 CONSIDERAÇÕES FINAIS

Com base nos resultados da pesquisa, é possível gerenciar os dados de pesquisa quando há documentação, recursos de tecnologia e de pessoal especializado. Embora existam centenas de IES no Brasil, apenas três apresentaram documentos com normativas para a promoção da GDP. Esses dados revelam um atraso nacional em comparação com países estrangeiros, mais avançados na promoção de trabalhos colaborativos e transparentes, conforme destacado por Sayão e Sales (2015).

A Fapesp foi a primeira agência de fomento a solicitar o PGD no Brasil. Não por acaso, das oito IES que vêm prestando serviços, seis são do estado de São Paulo, onde três são da autarquia estadual e quatro da autarquia federal.

É preciso que mais agências de fomento entrem ao movimento e passem a solicitar o PGD a fim de impulsionar esse serviço no país. Caso contrário, é possível que os pesquisadores brasileiros apresentem o planejamento somente quando houver estímulos externos. Nestes casos, o gerenciamento seria realizado sem a vigilância da instituição que financiou total ou parcialmente a pesquisa.

Ressalta-se que **o PGD não depende da Ciência Aberta**, pois tem valor por si só. A *e-science*, intrinsecamente ligada ao trabalho em rede, busca a máxima agilidade no compartilhamento de dados de pesquisa e os padrões precisam ser claros e documentados. Os benefícios da abertura e reuso de dados vêm acompanhados de grandes responsabilidades, uma vez que a ciência sem fronteiras exige a excelência dos pesquisadores e das ferramentas de trabalho, o PGD informa a intenção, os meios e os objetivos da investigação. Assim como a pesquisa, o PGD também é dinâmico.

A prática da pesquisa é uma atividade cara. Além dela própria, uma série de demandas envolvem custos de manutenção da infraestrutura física, tecnológica, materiais de consumo, de pessoal especializado em diversas áreas e funções que são comumente excluídas nas estimativas de despesas, como o pessoal de limpeza e encarregados de serviços gerais.

Os dados podem ter um valor simbólico, ou relevância social, econômica, comercial e outros ativos. Ignorá-los, guardá-los “em gavetas”, descartá-los ou ceder os direitos a terceiros, é uma atitude irresponsável e, em certos casos, criminosa, pois lança-se mão de direitos e ganhos, como no caso de publicações de artigos científicos.

De acordo com o Índice Global de Inovação, que avalia a capacidade de um país em gerar e aplicar novas ideias, o Brasil ocupa a 49ª posição no ranking de 2023, enquanto a Holanda se destaca entre os dez primeiros colocados (Site da Serra, 2023). No *QS World University Rankings 2024: Top global universities*, a USP assume a liderança na América Latina, ocupando a 85ª geral, obtendo 62.8 pontos. A primeira universidade federal a aparecer na lista é a UFRJ, classificada na 371ª posição, com 29.9 pontos. Em comparativo com as universidades holandesas, a Universidade de Tecnologia de Delft ocupou a 47ª posição com 76.2 pontos, as demais também estão à frente das IES do Brasil.

Em síntese, apesar da significativa lacuna em termos de desenvolvimento na inovação tecnológica e educacional entre a Holanda e o Brasil, é possível avançar e desenvolver a gestão de dados de pesquisa nas IES do nacionais.

Em 2018 o MCTI, no relatório “Estratégia Nacional de Ciência, Tecnologia e Inovação 2016-2022”, estabeleceu como meta para o Brasil tornar-se um líder global em ciência aberta até 2030. No entanto, o mesmo documento informa que, de acordo com o levantamento da OCDE de 2015, o Brasil ainda está distante dos países mais avançados, tanto no investimento em pesquisa e desenvolvimento como em recursos humanos (Brasil, 2018b). Embora o governo federal tenha o pleno poder para induzir a GDP, as instituições também devem se comprometer, ao invés de apenas esperar as circunstâncias favoráveis (ou compelidas).

Destaca-se, por fim, que antes de iniciar a pesquisa, ou mesmo o seu PGD, **é essencial que os pesquisadores busquem em repositórios de dados a disponibilidade dos dados que planejam gerar/coletar**, antes mesmo de elaborarem seus PGD. Sem essa prática, a GDP não fará sentido. Os pesquisadores devem receber orientações sobre a necessidade dessa pesquisa até que ela se torne habitual. No entanto, as habilidades necessárias para conduzir uma busca por dados de pesquisa podem não ser tão simples. Em algumas situações, a biblioteca poderá oferecer assistência; no entanto, em outras, apenas um profissional especializado será capaz de fornecer a capacitação necessária.

Por fim, em meio às adversidades, as heroicas IES do Brasil, com recursos limitados, nunca fugiram à luta. Gigante pela própria natureza, com dimensões continentais, o Brasil é reconhecido por suas pesquisas inovadoras e de alta qualidade. Líder na América Latina na publicação de resultados de pesquisa com Acesso Aberto, resplandece a criatividade e capacidade de adaptação ao ambiente, mesmo no novo mundo de gerenciar para reusar os dados, que fulguras em nosso horizonte. Isso permite



que as IES nacionais, mesmo enfrentando desafios, impávidas, acompanhem as inovações globais da ciência e da tecnologia, ainda que em um ritmo próprio e singular, “*quae sera tamen*”.

## REFERÊNCIAS

ABNT. **NBR 15472**: Sistemas espaciais de dados e informações – Modelo de referência para um sistema aberto de arquivamento. Rio de Janeiro, 2007.

ALBAGLI, Sarita; CLINIO, Anne; RAYCHTOCK, Sabryna. Ciência aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, p. 434-450, 2014. DOI: <https://doi.org/10.18617/liinc.v10i2.749>. Disponível em: <https://revista.ibict.br/liinc/article/view/3593/3072>. Acesso em: 12 jul. 2022.

ALLEA – All European Academies. **The European Code of Conduct for Research Integrity**. Berlin: ALLEA, 2017. Disponível em: <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf>. Acesso em: 22 set. 2022.

ALLEA – All European Academies. **The European Code of Conduct for Research Integrity** - Revised Edition. Berlin: ALLEA, 2023. Disponível em: [https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity\\_horizon\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf). Acesso em: 15 jul. 2023.

ALMEIDA, Fernanda Gomes. **Suporte à gestão de dados de pesquisa: uma ampliação dos serviços oferecidos pelas bibliotecas**. 2019. 249 f. Tese (Doutorado) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019. Orientadora: Beatriz Valadares Cendón. Disponível em: [https://repositorio.ufmg.br/bitstream/1843/31596/4/Tese\\_PPGGOC\\_Fernanda%20Gomes%20Almeida\\_11out.pdf](https://repositorio.ufmg.br/bitstream/1843/31596/4/Tese_PPGGOC_Fernanda%20Gomes%20Almeida_11out.pdf). Acesso em: 05 fev. 2023.

ANGUES, Ivan. **III Relatório: Projeto RDP Brasil** – Acesso Aberto a Dados de Pesquisa no Brasil (Relatório do Website). 2019. Disponível em: <https://dadosdepesquisa.rnp.br/?p=420>. Acesso em: 04 ago. 2022.

ARGOS. **Plataforma para Gestão de Dados de pesquisa**. [202-?]. União Europeia: OpenAIRE. Desenvolvido pelo software aberto: OpenDMP. Disponível em: <https://argos.openaire.eu/splash/index.html>. Acesso em: 11 jul. 2023.

BALL, Alex. Metadata: The Spirit of Research Data Management. **Research Infrastructures in Social Sciences and Humanities**, Bath, v. 29, n. 6, 2017.

BANKS, James. Policies and strategies to facilitate secondary use of research data in the health sciences. **Oxford Journals Current Collection Medline Complete International journal of epidemiology**, Oxford, v. 46, n. 6, p. 1729-1733, 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29025140/>. Acesso em: 30 maio 2022.

BARDIN, Laurence. **Análise de Conteúdo**. Lisboa: Edições 70, 1977. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/7684991/mod\\_resource/content/1/BARDIN\\_\\_L.\\_1977](https://edisciplinas.usp.br/pluginfile.php/7684991/mod_resource/content/1/BARDIN__L._1977). Acesso em: 16 ago. 2022.

BNDES – Banco Nacional de Desenvolvimento Econômico e Social. **Visão 2035:** Brasil, país desenvolvido – Agendas setoriais para alcance da meta. Rio de Janeiro: BNDES, 2018. Disponível em: [https://web.bndes.gov.br/bib/jspui/bitstream/1408/16040/3/PRLiv214078\\_Visao\\_2035\\_compl\\_P.pdf](https://web.bndes.gov.br/bib/jspui/bitstream/1408/16040/3/PRLiv214078_Visao_2035_compl_P.pdf). Acesso em: 30 jun. 2023.

BOAI – Budapest Open Access Initiative. **Declaração Internacional Sobre o Acesso Aberto**. 2002. Disponível em: <http://openaccessprod.wpengine.com/read/>. Acesso em: 12 jul. 2022.

BOAI – Budapest Open Access Initiative. **The Budapest Open Access Initiative: 20th Anniversary Recommendations**. 2022. Disponível em: <https://www.budapestopenaccessinitiative.org/boai20/>. Acesso em: 12 jul. 2022.

BONINA, Carla M. New business models and the value of open data: definitions, challenges and opportunities. **Nemode–3K Small Grants Call**, São Paulo, v. 10, p. 31-70, 2013. Disponível em: <http://www.nemode.ac.uk/wp-content/uploads/2013/11/Bonina-Opendata-Report-FINAL.pdf>. Acesso em: 05 ago. 2022.

BORGMAN, Christine L. Open Data, Grey Data, and Stewardship: Universities at the Privacy Frontier. **Berkeley Technology Law Journal**, Berkeley, v. 33, n. 2, p. 365-412, 2018. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1802/1802.02953.pdf>. Acesso em: 28 jun. 2022.

BORGMAN, Christine L. The conundrum of sharing research data. **Journal of the American Society for Information Science and Technology**, Hoboken, v. 63, n. 6, p. 1059-1078, 2012.

BORGMAN, Christine L. Research Data: Who Will Share What, with Whom, When, and Why? **RatSWD Working Paper**, Berlin, n. 161, 2010. Disponível em: <https://ssrn.com/abstract=1714427>. Acesso em: 09 ago. 2022.

BORGMAN, Christine L.; BOURNE, Philip E. Why it takes a village to manage and share data. **ArXiv preprint**. 2021. Disponível em: <https://arxiv.org/abs/2109.01694>. Acesso em: 10 set. 2021.

BRASIL. Presidência da República. **Lei nº 13.243**, de 11 de janeiro de 2016. Institui o Marco Legal da Ciência, Tecnologia e Inovação. Diário Oficial da União. Brasília, DF, 12 jan. 2016. Seção 1, p. 1. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2016/Lei/L13243.htm](https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Lei/L13243.htm). Acesso em: 19 mar. 2023.

BRASIL. Presidência da República. **Lei nº 10.973**, de 2 de dezembro de 2004. Estabelece medidas de incentivo à pesquisa científica e tecnológica no ambiente produtivo, com vistas à capacitação e ao alcance da autonomia tecnológica e ao desenvolvimento industrial do País, nos termos dos arts. 218 e 219 da Constituição. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2004/lei/l10.973.htm](https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.973.htm). Acesso em: 08 abr. 2023. Lei da Inovação.

BRASIL. **Constituição** (1988). Constituição da República Federativa do Brasil. Brasília, DF: Senado Federal, 1988.

BRASIL. Controladoria Geral da União. **Portal da Transparência**. 2023. Disponível em: <https://portaldatransparencia.gov.br/sobre/legislacao>. Acesso em: 25 abr. 2023.

BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. **Estratégia Nacional de Ciência, Tecnologia e Inovação 2016-2022**. 2018b. Disponível em: [https://antigo.mctic.gov.br/mctic/export/sites/institucional/ciencia/SEPED/Arquivos/PlanosDeAcao/PACTI\\_Sumario\\_executivo\\_Web.pdf](https://antigo.mctic.gov.br/mctic/export/sites/institucional/ciencia/SEPED/Arquivos/PlanosDeAcao/PACTI_Sumario_executivo_Web.pdf). Acesso em: 01 jun. 2023.

BRASIL. Ministro de Estado da Ciência e Tecnologia. **Portaria nº 693**, de 20 de agosto de 2009. Institui, no âmbito do Programa de Pesquisa em Biodiversidade no Brasil – PPBio, a Política de Dados. Brasília: DOU, 2009. Disponível em: [https://ppbio.inpa.gov.br/sites/default/files/politica\\_dou.pdf](https://ppbio.inpa.gov.br/sites/default/files/politica_dou.pdf). Acesso em: 01 jun. 2023.

BRASIL. Presidência da República. Casa Civil. **Lei nº 12.527**, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm). Acesso em: 10 ago. 2022. Lei de Acesso à Informação (LAI).

BRASIL. Presidência da República. **Lei nº 14.129**, de 29 de março de 2021. Dispõe sobre princípios, regras e instrumentos para o Governo Digital e para o aumento da eficiência pública. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2019-2022/2021/lei/L14129.htm](https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14129.htm). Acesso em: 01 ago de 2023. Lei do Governo Digital (E-gov).

BRASIL. Ministério da Ciência, Tecnologia e Inovação. Rede Nacional de Pesquisa. **O que é Ciência Aberta e como ela pode facilitar a vida de cientistas**. 2021. Disponível em: <https://www.rnp.br/noticias/o-que-e-ciencia-aberta-e-como-ela-pode-facilitar-vida-de-cientistas>. Acesso em: 20 set. 2022.

BRASIL. Ministério da Ciência, Tecnologia e Inovações. Secretaria de Empreendedorismo e Inovação. **Programa Nacional de Acesso Aberto a Dados de Pesquisa**. Brasília: RNP/IBICT, 2018a. Disponível em: [https://dadosdepesquisa.rnp.br/?page\\_id=76](https://dadosdepesquisa.rnp.br/?page_id=76). Acesso em: 15 jun. 2023.

BRASIL. Ministério da Ciência, Tecnologia e Inovações. Secretaria de Empreendedorismo e Inovação. **RNP, IBICT. RDP-Brasil: sobre nós**. Brasília: RNP/IBICT, 2018b. Disponível em: [https://www.prp.unicamp.br/wp-content/uploads/sites/4/2022/05/cgpd\\_delibera\\_ccp\\_006\\_2020\\_cria\\_e\\_regulamenta\\_o\\_repositorio\\_de\\_dados.pdf](https://www.prp.unicamp.br/wp-content/uploads/sites/4/2022/05/cgpd_delibera_ccp_006_2020_cria_e_regulamenta_o_repositorio_de_dados.pdf)[https://dadosdepesquisa.rnp.br/?page\\_id=10](https://dadosdepesquisa.rnp.br/?page_id=10).

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Instrução Normativa No - 4, de 12 de abril de 2012. **Institui a Infraestrutura Nacional de Dados Abertos - INDA**. Brasília, DOU, 2012. Disponível em:

<https://www.gov.br/governodigital/pt-br/dados-abertos/InstrucaoNormativaINDA42012.pdf>. Acesso em: 15 mar. 2023.

BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos. **Plano de Dados Abertos**. [2017?]. Disponível em: <https://www.gov.br/governodigital/pt-br/sisp/guia-do-gestor/pda>. Acesso em: 13 jul. 2023.

BRASIL. Portal Gov.br. Ministério da Gestão e da Inovação em Serviços. **Portal Brasileiro de Dados Abertos**: O Portal Brasileiro de Dados Abertos é o ponto central para a busca e o acesso aos dados públicos no Brasil. Brasília, 2019. Disponível em: <https://www.gov.br/governodigital/pt-br/dados-abertos/portal-brasileiro-de-dados-abertos>. Acesso em: 12 fev. 2023.

BRASIL. Presidência da República. **Decreto nº 1.171**, de 22 de junho de 1994. Código de Ética Profissional do Servidor Público Civil do Poder Executivo Federal. Diário Oficial da União, Brasília, DF, 23 jun. 1994. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/decreto/d1171.htm](https://www.planalto.gov.br/ccivil_03/decreto/d1171.htm). Acesso em: 01 jul. 2023.

BRASIL. Presidência da República. **Decreto nº 8.777**, de 11 de maio de 2016. Institui a Política de Dados Abertos do Poder Executivo federal. Brasília, DF: Presidência da República, 2016. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2016/decreto/d8777.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm). Acesso em: 24 set. 2022.

BRASIL. Presidência da República. **Lei nº 12.527**, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm). Acesso em: 24 set. 2022.

BRASIL. Presidência da República. Secretaria Geral. Subchefia para Assuntos Jurídicos. **Lei nº 13.709**, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). Lei Geral de Proteção de Dados Pessoais (LGPD). (Redação dada pela Lei nº 13.853, de 2019) Vigência. Brasília, DOU, 2018. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/113709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm). Acesso em: 12 fev. 2023.

BRASIL. Presidência da República. **Lei nº 13.709**, de 14 de agosto de 2018c. Dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/113709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm). Acesso em: 25 abr. 2023.

BRASIL. Universidade Federal de São Paulo (UNIFESP). **Política Institucional de Informação para o Repositório Institucional Unifesp**. São Paulo: UNIFESP, 2017. Portaria Reitoria n° 4845 de 18 de dezembro de 2017. Disponível em: <https://repositoriodedados.unifesp.br/>. Acesso em: 10 maio 2023.

BRASIL. Universidade Federal de Santa Catarina. **Comissão de Concepção do Serviço de Suporte à Pesquisa e Gestão de Dados Científicos**. 2019. Disponível em: <https://portal.bu.ufsc.br/conheca-a-bu/comissoes-de-trabalho/concepcaooservico/>. Acesso em: 02 jul. 2023.

BRASIL. Universidade Federal de São Carlos (UFSCar). **Manual de Autodepósito de Dados de Pesquisa**. São Carlos: RI-UFSCar, 2019. Disponível em: <https://repositorio.ufscar.br/static/ri-ufscar-manual-autodeposito-dados.pdf>. Acesso em: 06 de maio de 2023.

BRASIL. Universidade Federal do Paraná. **Base de Dados Científicos da Universidade Federal do Paraná: Diretrizes**. Curitiba: UFPR, 2018. Disponível em: <https://bibliotecas.ufpr.br/2021/06/base-de-dados-cientificos/>. Acesso em: 15 fev. 2023.

BRIET. **Grupo de Estudos em Biblioteconomia, Representação, Interoperabilidade, E-Science e Tecnologia**. Ano de Formação: 2017. Líderes: Luana Farias Sales Marques e Luis Fernando Sayão. Localização: Rio de Janeiro. Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT. Disponível em: <https://grupobriet.com/>.

BRYMAN, Alan. **Social research methods**. 4th. ed. Oxford: Oxford university press, 2012.

CALIARI, Thiago; SANTOS, Ulisses Pereira; MENDES, Philipe Scherrer. Geração de tecnologia em universidades/institutos de pesquisa e a importância da interação com empresas: constatações através da base de dados dos grupos de pesquisa do CNPQ. **Análise Econômica**, Porto Alegre, v. 34, n. 66, 2016.

CLINIO, Anne *et al.* **Panorama histórico da Ciência Aberta**. Rio de Janeiro: Campus Virtual Fiocruz/Escola Corporativa Fiocruz, 2018. Material multimídia (Microcurso 2/3, Série 1 – Formação Modular em Ciência Aberta).

CHAWINGA, Winner Dominic; ZINN, Sandy. Global perspectives of research data sharing: A systematic literature review. **Library & Information Science Research**, v. 41, n. 2, p. 109-122, Apr. 2019. Disponível em: <https://doi.org/10.1016/j.lisr.2019.04.004>. Acesso em: 22 jun. 2022.

CHOO, Chun Wei. **A organização do conhecimento**: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. São Paulo: Senac, 1998. 425 p.

CODATA – COMITÊ DE DADOS. **Missão do CODATA**. [20--b]. Disponível em: <https://codata.org/about-codata/our-mission/>. Acesso em :27 abr. 2022.

CODATA – COMITÊ DE DADOS. **Glossary of Terms**. [20--a]. Disponível em: <https://datatree.org.uk/mod/glossary/print.php?id=230&mode=letter&hook=D&sortkey&sortorder&offset=0&pagelimit=1000>. Acesso em: 11 ago. 2022.

COMISSÃO EUROPEIA. **Horizon Europe: 2021-2027**. Disponível em: [https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe\\_en](https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en). Acesso em: 10 de out. 2023.

COMISSÃO EUROPEIA. **Programa Horizon 2020**. Disponível em: [https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020\\_en](https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en)>. Acesso em: 10 jan. 2023. Programa de Financiamento de Pesquisa e Inovação da UE [2014 a 2020].

COSTA, Belkiz; OLIVEIRA, Marlene; ARAÚJO, Ronaldo. Impactos das teses e dissertações do programa de pós-graduação em ciência da informação da UFMG: dos dados de leitura no Mendeley às citações no Google Scholar. **Informação em Pauta**, Fortaleza, v. 4, n. 2, jul./dez. 2019. DOI: <https://doi.org/10.32810/2525-3468.ip.v4i2.2019.42444.11-31>. Disponível em: <http://www.periodicos.ufc.br/informacaoempauta/article/view/42444>. Acesso em: 13 maio 2022.

COSTA, Maira Murrieta. **Diretrizes para uma política de gestão de dados científicos no Brasil**. 2017. 288f. Dissertação (Mestrado em Ciência da Informação). Universidade de Brasília, Brasília. Disponível em: <https://repositorio.unb.br/handle/10482/24895>. Acesso em: 02 mar. 2023.

COSTA, M. M.; CUNHA, M. B. da. A necessidade de uma política nacional para a gestão de dados de pesquisa no Brasil. **Liinc Em Revista**, Rio de Janeiro, v. 2, n. 15, p. 287-309, 2019. Disponível em: <https://doi.org/10.18617/liinc.v15i2.4763>. Acesso em: 02 fev. 2023.

CUNHA, Murilo Bastos da; AMARAL, Sueli Angelica do; DANTAS, Edmundo B. **Manual de Estudo de Usuários da Informação**. São Paulo: Atlas, 2015. E-book. ISBN 9788522498789. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522498789/>. Acesso em: 22 maio 2023.

CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. Dicionário de biblioteconomia e arquivologia. Briquet de Lemos Livros. **Revista Ibero-americana de Ciência da Informação (RICI)**, Brasília, v.1 n.2, p. 225-226, jul./dez. 2008 Disponível em: <https://scholar.archive.org/work/3a3wjsxzvfgbfwy5x42mjbhyi/access/wayback/https://periodicos.unb.br/index.php/RICI/article/download/1357/1195>. Acesso em: 09 maio 2023.

CWTS – Centre for Science and Technology Studies. **Leiden Ranking 2022**. Disponível em: <https://www.leidenranking.com/ranking/2022/list>. Acesso em: 16 set. 2022.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, p. 101-107, 1978. Disponível em: <https://doi.org/10.18225/ci.inf.v7i2.115>. Acesso em: 24 set. 2022.

DANS – Data Archiving and Networked Services. **Apresentação**. 2023. Disponível em: <https://dans.knaw.nl/nl/over/>. Acesso em: 08 maio 2023.

DATA'S SHAMEFUL NEGLECT. *Nature*, Londres, v. 461, n. 145, 2009. Disponível em: <https://doi.org/10.1038/461145a>. Acesso em: 08 ago. 2022.

DATAVERSENO. **What is DataverseNO?**.2020. Disponível em: <https://site.uit.no/dataverseNO/about/>. Acesso em: 08 ago. 2022.

DAVENPORT, Thomas H. **Ecologia da informação**: por que só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 1998.

DCC – Digital Curation Centre **Using DMPonline to support data management and FAIR data**. 2019. Disponível em: <https://www.dcc.ac.uk/events/workshops/using-dmponline-support-data-management-and-fair-data>. Acesso em: 20 out. 2022.

DECLARAÇÃO de Berlim sobre Acesso Livre ao Conhecimento nas Ciências e Humanidades. 2003. Disponível em: [https://openaccess.mpg.de/67693/BerlinDeclaration\\_pt.pdf](https://openaccess.mpg.de/67693/BerlinDeclaration_pt.pdf). Acesso em: 26 set. 2022.

DILLO, Ingrid; DOORN, Peter. **The front office-back office model**: supporting research data management in the Netherlands. Semantic Scholar, Seattle, 2014. Disponível em: <https://www.semanticscholar.org/paper/The-Front-Office-Back-Office-Model%3A-Supporting-Data-Dillo-Doorn/3d1bd6ec35f86d65d57de78688905876d4f0556b>. Acesso em: 02 mar. 2023.

DIJK, E. M. S.; DOORN, P. K. Providing access to research data, publications and current research information at Data Archiving and Networked Services-DANS. In: **7th Conference on Grey Literature and Repositories**: proceedings [online] 2014. National Library of Technology, 2014. Disponível em: <https://pure.knaw.nl/portal/en/publications/providing-access-to-research-data-publications-and-current-resear>. Acesso em: 02 mar. 2023.

DUDZIAK, Elizabeth. Levantamento mostra quem financia a pesquisa no Brasil e na USP. *Jornal da USP*, São Paulo, 14 out. 2020. Disponível em: <https://jornal.usp.br/?p=182855>. Acesso em: 16 set. 2022.

ECONOMIC AND SOCIAL RESEARCH COUNCIL. **Research Data Management Policy**. 2021. Disponível em: [https://www.open.ac.uk/library-research-support/sites/www.open.ac.uk.library-research-support/files/files/Open%20University%20Research%20Data%20Management%20Policy%202021%20\(2\).pdf](https://www.open.ac.uk/library-research-support/sites/www.open.ac.uk.library-research-support/files/files/Open%20University%20Research%20Data%20Management%20Policy%202021%20(2).pdf). Acesso em: 01 jul. 2022.

ENSERINK, Martin. In dramatic statement, european leaders call for ‘immediate’ open access to all scientific papers by 2020. *Science*, 27 May 2016. Disponível em: [www.sciencemag.org/news/2016/05/dramatic-statement-european-leaders-call-immediate-open-access-all-scientific-papers?et rid=34985979&et cid=547163#](http://www.sciencemag.org/news/2016/05/dramatic-statement-european-leaders-call-immediate-open-access-all-scientific-papers?et rid=34985979&et cid=547163#). Acesso em: 07 jun. 2022.

EOSC – European Open Science Cloud. **Sobre**. 2023. Disponível em: <https://eosc-portal.eu/>. Acesso em: 02 jan. 2023.



EOSC – EUROPEAN COMMISSION. **Open Science**. 2019. Disponível em: [https://ec.europa.eu/info/sites/default/files/research\\_and\\_innovation/knowledge\\_publications\\_tools\\_and\\_data/documents/ec\\_rtd\\_factsheet-open-science\\_2019.pdf](https://ec.europa.eu/info/sites/default/files/research_and_innovation/knowledge_publications_tools_and_data/documents/ec_rtd_factsheet-open-science_2019.pdf). Acesso em: 12 jul. 2022.

ESTERLING, Kevin; NEBLO, Michael; LAZER, David. **Replication Data for: Politics With the People**. Cambridge: Cambridge Press, 2018. Disponível em: <https://doi.org/10.7910/DVN/5FCUPY>, UNF:6:iLmQpyyqcES5SluNgwnxOg== [fileUNF]. Acesso em: 03 abr. 2022.

EUROPEAN COMMISSION. Directorate-General for Research and Innovation. **Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group**. Bruxelas: Publications Office, 2020. Disponível em: <https://data.europa.eu/doi/10.2777/986252>. Acesso em: 10 jan. 2023.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **Diretrizes para Planos de Gestão de Dados (PGD) para Propostas de Centros**. 2021a. Disponível em: <https://fapesp.br/14974/diretrizes-para-planos-de-gestao-de-dados-pgd-para-propostas-d-e-centros>. Acesso em: 05 ago. 2022.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **Elaboração e esclarecimento de dúvidas sobre Planos de Gestão de Dados (PGD)**. 2020. Disponível em: <https://fapesp.br/14648/elaboracao-e-esclarecimento-de-duvidas-sobre-planos-de-gestao-de-dados-pgd>. Acesso em: 09 ago. 2022.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **FAPESP em números**. São Paulo: FAPESP, 2023. 10 p. <https://bv.fapesp.br/pt/10/>. Acesso em: 2 jun. 2023.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **Gestão de Dados**. [201-a]. Disponível em: <https://fapesp.br/gestaodedados>. Acesso em: 18 abr. 2022.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **Metabusador de dados de pesquisa**. [201-b]. Disponível em: <https://metabusador.uspdigital.usp.br/>. Acesso em: 24 set. 2022.

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo. **Política para Propriedade Intelectual da FAPESP**. 2021b. Disponível em: <https://fapesp.br/15332/politica-para-propriedade-intelectual-da-fapesp-valida-de-20042021-a-16022022>. Acesso em: 05 ago. 2022.

FUNDAÇÃO ESPÍRITO-SANTENSE DE TECNOLOGIA. **Plano de Gestão de dados indicado pelo IBAMA**. Programa de Monitoramento da Biodiversidade Aquática da Área Ambiental I – Porção Capixaba do Rio Doce e Região Marinha e Costeira Adjacente. Plano de Gestão e Dados. Vitória, ES: Rede Rio Doce/Fest, 2019. Disponível em: <https://fest.org.br/projetos/programa-de-monitoramento-da-biodiversidade-ambiental-pmba/>. Acesso em: 16 jun. 2023.

FIOCRUZ – Fundação Oswaldo Cruz. **Ciência Aberta**: conheça os oito cursos on-line e gratuitos da Fiocruz. 2021. Disponível em: <https://portal.fiocruz.br/es/noticia/ciencia-aberta-conheca-os-oito-cursos-line-e-gratuitos-da-fiocruz>. Acesso em: 05 maio 2023.

FIOCRUZ – Fundação Oswaldo Cruz. **Glossário do Acesso Aberto**. [20--]. Disponível em: <https://portal.fiocruz.br/glossario-acesso-aberto/C>. Acesso em: 02 jan. 2023.

FIOCRUZ – Fundação Oswaldo Cruz: Campus Virtual. **Panorama Histórico da Ciência Aberta**. 2018. Disponível em: <https://mooc.campusvirtual.fiocruz.br/rea/ciencia-aberta/serie1/curso2/aula1.html>. Acesso em: 09 dez. 2022.

FIOCRUZ – Fundação Oswaldo Cruz. **PDA – Plano de Gestão de Dados da Fundação Oswaldo Cruz**. [Ferramenta para elaborar um plano de gestão].2022. Disponível em: <https://fiopgd.icict.fiocruz.br/>. Acesso em: 15 set. 2022.

FIOCRUZ – Fundação Oswaldo Cruz. **FioDMP – Plano de Gestão de Dados da Fundação Oswaldo Cruz**. 2023. [Ferramenta aberta para elaborar um plano de gestão de dados de pesquisa-nova versão]. Disponível em: <https://fiopgd.icict.fiocruz.br/minha-lista-01/190>. Acesso em: 12 ago. 2023.

FIOCRUZ – Fundação Oswaldo Cruz. **Política de gestão, compartilhamento e abertura de dados para pesquisa**: princípios e diretrizes. 2020. Rio de Janeiro: Fiocruz/VPEIC, 2020. Disponível em: [https://www.arca.fiocruz.br/bitstream/icict/46408/2/VPEIC\\_versao\\_PORTUGUES\\_2021-03-22.pdf](https://www.arca.fiocruz.br/bitstream/icict/46408/2/VPEIC_versao_PORTUGUES_2021-03-22.pdf). Acesso em: 05 ago. 2022.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

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

GIL, Antonio C. **Como Elaborar Projetos de Pesquisa**. Grupo GEN, 2022. E-book. ISBN 9786559771653. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786559771653/>. Acesso em: 19 maio 2023.

GO FAIR. **Princípios FAIR**. [201-?]. Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 18 fev. 2023.

GO FAIR. GO FAIR Brasil. **Escritório da GO FAIR Brasil**. 2023. Disponível em: <https://www.go-fair.org/go-fair-initiative/go-fair-offices/go-fair-brazil-office/>. Acesso em: 18 jul. 2023.

GO FAIR. **A Internet de Dados e Serviços FAIR**. [2023?]. Disponível em: <https://www.go-fair.org/resources/internet-fair-data-services/>. Acesso em: 25 mar. 2023.

GRÁCIO, J. C. A.; MÁRDERO ARELLANO, M. Ángel. A gestão da preservação digital de dados de pesquisa: proposta de um modelo processual. **Revista Brasileira de Preservação Digital**, Campinas, SP, v. 1, p. 1-19, 2021. DOI:

10.20396/rebpred.v1i00.13223. Disponível em:  
<https://econtents.bc.unicamp.br/inpec/index.php/rebpred/article/view/13223>. Acesso em: 22 set. 2022.

GRAY, J. **eScience**: a transformed scientific method. Transcrição de palestra ministrada por Jim Gray no Conselho Nacional de Pesquisa (EUA), 11 Jan. 2007. In: HEY, T.; TANSLEY, S.; TOLLE, K. (Ed.). *The fourth paradigm: data-intensive scientific discovery*. Redmond: Microsoft Research, 2009.

HIGMAN, Rosie; BANGERT, Daniel; JONES, Sarah. Three Camps, One Destination: The Intersections of Research Data Management, FAIR and Open. **Insights**, v. 32, n. 1. 2019. DOI: <http://doi.org/10.1629/uksg.468>. Disponível em:  
<https://insights.uksg.org/articles/10.1629/uksg.468/>. Acesso em: 06 set. 2022.

HIGMAN, Rosie; PINFIELD, Stephen. Research data management and openness: The role of data sharing in developing institutional policies and practices. **Program: electronic library and information systems**, Bingley, v. 49, n. 4, p. 364-381, 2015.

HOLANDA. **Gerenciamento de Dados de Pesquisa UvA**. 2019. Disponível em:  
<https://rdm.uva.nl/en>. Acesso em: 28 set. 2022.

HOLANDA. **Portal of Research Output from the Netherlands**. 2023. Disponível em:  
<https://openresearch.nl/>. Acesso em: 06 jul. 2023.

HOLANDA. **Lei do Ensino Superior e da Pesquisa Científica** - WHW. Overheid, 2002. Disponível em:  
[https://wetten.overheid.nl/BWBR0005682/2020-08-01/#Hoofdstuk1\\_Titeldeel1\\_Artikel1.5](https://wetten.overheid.nl/BWBR0005682/2020-08-01/#Hoofdstuk1_Titeldeel1_Artikel1.5). Acesso em: 15 jul. 2023.

HOLANDA. **Lei do Governo Aberto** - Woo. Overheid, 2023. Disponível em:  
<https://wetten.overheid.nl/BWBR0045754/2023-04-01>. Acesso em: 17 jun. 2023.

HOLANDA. Programa Nacional de Ciência Aberta da Holanda. **Open Science 2030 in the Netherlands (NPOS)**. NPOS2030 Ambition Document and Rolling Agenda. Disponível em: <https://doi.org/10.5281/zenodo.7433767>. Acesso em: 15 jul. 2023.

HOLANDA. Universidade de Erasmus de Rotterdam. **Política de Gerenciamento de Dados de Pesquisa da Erasmus University Rotterdam**. 2020. Disponível em:  
<https://www.eur.nl/en/research/research-services/research-data-management>. Acesso em: 01 dez. 2022.

HOLANDA. Universidade de Groningen. **Gerenciamento de Dados de Pesquisa**. 2021. Disponível em: <https://www.rug.nl/digital-competence-centre/research-data/>. Acesso em: 03 out. 2022.

HOLANDA. Universidade de Leiden. **Regulamento de gerenciamento de dados da Universidade de Leiden**. 2021. Disponível em:  
<https://www.library.universiteitleiden.nl/researchers/data-management>. Acesso em: 18 out. 2022.

HOLANDA. Universidade de Radboud. **General Policy on Research Data Storage and Management**. 2021. Disponível em:

<https://www.ru.nl/en/about-us/policies-and-regulations/research-data-management/research-data-storage-and-management>. Acesso em: 15 dez. 2022.

HOLANDA. Universidade de Radboud. **Política da universidade para armazenar e gerenciar dados de pesquisa**. 2021. Disponível em: <https://www.ru.nl/rdm/vm/policy-documents/>. Acesso em: 25 out. 2022.

HOLANDA. Universidade de Tilburg. **Regulamento de Gerenciamento de Dados de Pesquisa**. 2020. Disponível em: <https://www.tilburguniversity.edu/about/conduct-and-integrity/privacy-and-security/research-data/policy>. Acesso em: 15 jan. 2023.

HOLANDA. Universidade de Twente. **Política de Gerenciamento de Dados de Pesquisa**. 2018. Disponível em: <https://www.utwente.nl/en/service-portal/university-library/research-data-management>. Acesso em: 06 jan. 2023.

HOLANDA. Universidade de Utrecht. **Política de Gerenciamento de Dados de Pesquisa**. 2019. Disponível em: <https://www.uu.nl/en/research/research-data-management/guides/policies-codes-of-conduct-and-laws#ownership>. Acesso em: 10 out. 2022.

HOLANDA. Universidade de Utrecht. **University Policy Framework for Research Data**. 2019. Disponível em: <https://www.uu.nl/en/research/research-data-management/guides/policies-codes-of-conduct-and-laws>. Acesso em: 15 dez. 2022.

HOLANDA. Universidade de Vrije. **Política de Gerenciamento de Dados de Pesquisa**. 2020. Disponível em: <https://www.eur.nl/en/research/research-services/research-data-management>. Acesso em: 10 dez. 2022.

HOLANDA. Universidade de Wageningen. **Política de Dados na WUR**. 2019, Disponível em: <https://www.wur.nl/en/value-creation-cooperation/collaborating-with-wur-1/wdcc/research-data-management-wdcc/data-policy.htm>. Acesso em: 28 dez. 2022.

HOLANDA. Universidade Técnica de Delft. **Política de Estrutura de Dados de Pesquisa da TU Delft**. 2020. Disponível em: <https://www.tudelft.nl/library/research-data-management>. Acesso em: 19 dez. 2022.

HOLANDA. **Universiteiten van Nederland**. 2021. Haarlem. Disponível em: [https://www.universiteitenvannederland.nl/nl\\_NL/universiteiten.html](https://www.universiteitenvannederland.nl/nl_NL/universiteiten.html). Acesso em: 01 jun. 2023.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Índice Geral de Cursos (IGC)**. 2020. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior/indice-geral-de-cursos-igc>. Acesso em: 24 set. 2022.

IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia. **Termos de Compromisso de Gestão**. 2018. Disponível em: <https://www.gov.br/mcti/pt-br/aceso-a-informacao/unidades-de-pesquisa/termos-de-compromisso-de-gestao/ibict-instituto-brasileiro-de-informacao-em-ciencia-e-tecnologia>. Acesso em: 02 ago. 2023.

JONES, Sarah; PRYOR, Graham; WHYTE, Angus. **How to Develop Research Data Management Services** – a guide for HEIs. Edinburgh: Digital Curation Centre, 2013. Disponível em: <https://repository.oceanbestpractices.org/handle/11329/250>. Acesso em: 02 ago. 2022.

LE COADIC, Yves-François. **A Ciência da Informação**. 2. ed. Tradução de Maria Yeda F. S. F. Gomes. Brasília: Briquet de Lemos, 2004. 124 p.

LE COADIC, Yves-François. **La science de l'information**. 2<sup>e</sup> éd. corr. Paris: Universitaires de France, 1994. 127 p. Série: Que sais-je?

LEITE, Fernando César Lima; COSTA, Sely. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 206-219, maio/ago. 2006. Disponível em: <https://www.scielo.br/pdf/pci/v11n2/v11n2a05.pdf>. Acesso em: 16 jul. 2021.

LEITE, Valentina. **Fiocruz publica Termos de Uso de seu novo repositório de dados, o Arca Dados**. Rio de Janeiro: Fiocruz, 2022. Disponível em: <https://portal.fiocruz.br/noticia/fiocruz-publica-termos-de-uso-de-seu-novo-repositorio-de-dados-o-arca-dados>. Acesso em: 24 set. 2022.

LEWIS, Martin J. Libraries and the management of research data. In: MCKNIGHT, Sue. (ed.). **Envisioning future academic library services: initiatives, ideas and challenges**. London: Facet Publishing, 2010. p. 145-168. Disponível em: [https://eprints.whiterose.ac.uk/11171/1/LEWIS\\_Chapter\\_v10.pdf](https://eprints.whiterose.ac.uk/11171/1/LEWIS_Chapter_v10.pdf). Acesso em: 21 set. 2022.

MACIEL, Maria Lucia; APPEL, Andre Luiz; ALBAGLI, Sarita. **E-science e ciência aberta: questões em debate**. 2013. Disponível em: <https://brapci.inf.br/index.php/res/download/184798>. Acesso em: 29 set. 2022.

MARCONI, Marina de A.; LAKATOS, Eva M. **Metodologia Científica**. Grupo GEN, 2010. E-book. ISBN 9786559770670. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786559770670/>. Acesso em: 05 jun. 2023.

MATUSIAK, Krystyna K.; SPOSITO, Frank A. Types of research data management services: An international perspective. **Proceedings of the Association for Information Science and Technology**, v. 54, n. 1, p. 754-756, 2017. <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pra2.2017.14505401144>. Acesso em: 29 set. 2022.

MICHEL, Maria H. **Metodologia e Pesquisa Científica em Ciências Sociais**. 3. ed. São Paulo, Atlas, 2015. E-book. ISBN 978-85-970-0359-8. Disponível em:

<https://integrada.minhabiblioteca.com.br/#/books/978-85-970-0359-8/>. Acesso em: 19 maio 2023.

MIROWSKI, Philip. The future (s) of open science. **Social studies of science**, Londres, v. 48, n. 2, p. 171-203, 2018. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/0306312718772086>. Acesso em: 10 ago. 2022.

MONS, Barend. Fair Science for Social Machines: Let's Share Metadata Knowlets in the Internet of Fair Data and Services. **Data Intelligence**, Amsterdam, v. 1, n. 1, p. 22-42, 2019. Disponível em: [https://doi.org/10.1162/dint\\_a\\_00002](https://doi.org/10.1162/dint_a_00002). Acesso em: 29 set. 2022:

MULHOLLAND, Caitlin Sampaio. Dados pessoais sensíveis e a tutela de direitos fundamentais: uma análise à luz da Lei Geral de Proteção de Dados (Lei 13.709/18). **Revista de Direitos e Garantias Fundamentais**, Vitória, v. 19, n. 3, p. 159-180, 2018.

NATIONAL ACADEMIC RESEARCH AND COLLABORATIONS INFORMATION. **About Narcis**. [2011?]. Disponível em: <https://www.narcis.nl/about/Language/en>. Acesso em: 29 set. 2022.

NATIONAL RESEARCH COUNCIL. **A Question of Balance**: Private Rights and the Public Interest in Scientific and Technical Databases. Washington DC: National Academic Press, 1999. Citados por HIGMAN, Rosie; PINFIELD, Stephen. Research data management and openness: The role of data sharing in developing institutional policies and practices. *Program: electronic library and information systems*, v. 49, n. 4, p. 364-381, 2015. Disponível em: <https://nap.nationalacademies.org/download/9692>. Acesso em: 01 ago. 2022.

NETHERLANDS code of conduct for research integrity. 2018. (Elaborado por diversas universidades e instituições de pesquisa holandesas). Disponível em: [https://www.fearp.usp.br/images/pesquisa/NetherlandsCodeofConductforResearchIntegrity\\_2018\\_UK.pdf](https://www.fearp.usp.br/images/pesquisa/NetherlandsCodeofConductforResearchIntegrity_2018_UK.pdf). Acesso em: 22 set. 2022.

NETHERLANDS National Programme Open Science. [2022]. **Ambition Document Netherlands National Programme Open Science**. Disponível em: <https://doi.org/10.5281/zenodo.7010402>. Acesso em: 10 jul. 2023

OCDE – Organização para Cooperação e Desenvolvimento Econômico. **Enhanced Access to Publicly Funded Data for Science, Technology and Innovation**. DOI: <https://doi.org/10.1787/e045cd3d-en>. Paris: OCDE, 2020. Disponível em: [https://www.oecd-ilibrary.org/science-and-technology/enhanced-access-to-publicly-funded-data-for-science-technology-and-innovation\\_e045cd3d-en](https://www.oecd-ilibrary.org/science-and-technology/enhanced-access-to-publicly-funded-data-for-science-technology-and-innovation_e045cd3d-en). Acesso em: 15 dez. 2022.

OCDE – Organização para Cooperação e Desenvolvimento Econômico. **Principles and Guidelines for Access to Research Data from Public Funding**. Paris: OCDE, 2007. Disponível em: <https://www.oecd.org/sti/inno/38500813.pdf>. Acesso em: 10 jun. 2022.

OPEN ACCESS NETWORK. **History of the Open Access Movement**. [2019?]. Disponível em: <https://open-access.network/en/information/open-access-primers/history-of-the-open-access-movement>. acesso em: 10 ago. 2022.

OPENAIRE. OpenAIRE – **Open Access Infrastructure for Research in Europe**. [201-?]. Disponível em: <https://www.openaire.eu/>. Acesso em: 01 jul. 2023.

OPEN DATA INSTITUTE. **What makes data open?** 2013. Disponível em: <https://theodi.org/article/what-makes-data-open/>. Acesso em: 12 ago 2022.

OPEN KNOWLEDGE FOUNDATION. **Open Science**. [20--?]. Disponível em: <https://blog.okfn.org/category/open-science/>. Acesso em: 12 jul. 2022.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Open Science**. [201-?]. Disponível em: <https://www.oecd.org/sti/inno/open-science.htm#:~:text=Open%20science%20encompasses%20unhindered%20access,by%20ICT%20tools%20and%20incentives>. Acesso em: 12 jul. 2022.

PASQUETTO, Irene V.; RANDLES, Bernadette M.; BORGMAN, Christine L. On the reuse of Scientific Data. **Data Science Journal**, Berlin, v. 16, p. 01-08, 2017. DOI: <http://doi.org/10.5334/dsj-2017-008>. Disponível em: <https://datascience.codata.org/articles/10.5334/dsj-2017-008/>. Acesso em: 24 set. 2022.

PATRICE, M. Building open government. **Government Information Quarterly**, Amsterdam, v. 27, n. 4, p. 401-13, 2010.

PINHEIRO, Lena Vania Ribeiro *et al.* Memória Colóquios Interdisciplinares GT1 ENANCIB 2018: as mandalas da interdisciplinaridade da Ciência da Informação. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 14, n. 3, p. 145-161, 2019. Disponível em: <https://brapci.inf.br/index.php/res/download/151518>. Acesso em: 26 set. 2022.

PONTIKA, Nancy *et al.* Fostering open science to research using a taxonomy and an eLearning portal. In: **Proceedings of the 15th international conference on knowledge technologies and data-driven business**. 2015. p. 1-8. Disponível em: <https://dl.acm.org/doi/10.1145/2809563.2809571>. Acesso em: 15 mar. 2023.

PORTAL BRASILEIRO DE DADOS ABERTOS. **O que são dados abertos?**. [202-?]. Disponível em: <https://dados.gov.br/pagina/dados-abertos>. Acesso em: 19 jul. 2022.

PORTAL FOSTER. **Preparing a Research Data Management response: policy, infrastructure and practice**. 2015. Disponível em: <https://www.fosteropenscience.eu/node/539>. Acesso em: 31 jul. 2023.

PORTAL FOSTER. **Resources: Research Open Data Management**. [20--]. Disponível em: <https://www.fosteropenscience.eu/resources>. Acesso em: 01 jun. 2022.

PRYOR, Graham. Why manage research data? In: PRYOR, Graham (ed.). **Managing Research Data**. London: Facet Publishing, 2012. p. 1-16.

PUBLIC RESOURCE ORG. **The 8 Principles of Open Government Data**. 2007. Disponível em: <https://opengovdata.org/>. Acesso em: 19 jul. 2022.

QS WORLD UNIVERSITY RANKINGS. **QS World University Rankings 2024: Top global universities**. 2023. Disponível em:

<https://www.topuniversities.com/university-rankings/world-university-rankings/2024>. Acesso em: 27 set. 2023.

RDA – RESEARCH DATA ALLIANCE. **The Research Data Alliance (RDA) builds the social and technical bridges to enable the open sharing and re-use of data.** 2016. Disponível em: <https://www.rd-alliance.org/about-rda>. Acesso em: 21 mar. 2023.

RIBEIRO, Nivaldo Calixto; OLIVEIRA, Dalgiza Andrade. Universidades públicas federais brasileiras: ações e estratégias para a abertura da ciência. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, p. 01-12, 2019. Disponível em: <https://periodicos.ufmg.br/index.php/moci/article/view/19184>. Acesso em: 15 jul. 2022.

RNP – Rede Nacional de Ensino e Pesquisa. **Termos de uso da Rede Ipê.** 2019. Disponível em: <https://dadosdepesquisa.rnp.br/?p=442>. Acesso em:

RODE, Sigmar de Mello; SILVA, Eli Lopes da. Ética e integridade na publicação científica. In: PRÍNCIPE, Eloísa; RODE, Sigmar de Mello (org.). **Comunicação científica aberta.** Rio de Janeiro: IBICT, 2022. p. 63-79. (Coleção PPGCI 50 anos). Disponível em: [https://repositorio.ibict.br/bitstream/123456789/1223/1/PrincipeRode\\_ComunicacaoCientificaAberta\\_2022.pdf](https://repositorio.ibict.br/bitstream/123456789/1223/1/PrincipeRode_ComunicacaoCientificaAberta_2022.pdf). Acesso em: 22 set. 2022.

SAGAN, Carl. **Cosmos.** São Paulo: Cia das Letras, 1980. 560 p.

SALES, Luana Farias. **Integração semântica de publicações científicas e dados de pesquisa:** proposta de modelo de publicação ampliada para a área de Ciências Nucleares. 2014. 265 f. Tese (Doutorado em Ciência da Informação). Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia/Universidade Federal do Rio de Janeiro, Rio de Janeiro. Disponível em: <https://repositorio.ibict.br/bitstream/123456789/874/1/LUANA%20SALES%20D.pdf>. Acesso em: 11 jul. 2022.

SALES, Luana Farias. **Repositório Digital de Dados de Pesquisa: uma experiência na área de ciências nucleares.** [2017]. Disponível em: [https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/19044/2Apresenta%C3%A7%C3%A3o\\_Luana.pdf](https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/19044/2Apresenta%C3%A7%C3%A3o_Luana.pdf). Acesso em: 06 jul. 2022.

SALES, Luana Farias; SAYÃO, Luis Fernando; MARANHÃO, Ana Maria Neves; DRUMOND, Geisa Meirelles; SILVA, Maria Helena Ferreira Xavier da. Competências dos bibliotecários na gestão dos dados de pesquisa. **Ciência da Informação**, Brasília, v. 48, n. 3, p. 303-313, 2019. Suplemento. Disponível em: <https://revista.ibict.br/ciinf/article/view/4973/4458>. Acesso em: 24 set. 2022.

SALES, Luana Farias; SAYÃO, Luís Fernando. Há futuro para as bibliotecas de pesquisa no ambiente de e-Science. **Informação & Tecnologia**, Brasília, v. 2, n. 1, p. 30-52, 2015. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/26029/14677>. Acesso em: 12 ago.

SANSONE, Susanna-Assunta et al. FAIRsharing as a community approach to standards, repositories and policies. **Nature Biotechnology**, Londres, v. 37, p. 358-367, 2019. Disponível em: <https://doi.org/10.1038/s41587-019-0080-8>.



SANTOS, Paula Xavier dos; ALMEIDA, Bethânia de Araújo; HENNING, Patricia (organizadoras). **Livro Verde - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional**. Rio de Janeiro: Fiocruz, 2017. 140 p. Disponível em: <https://www.arca.fiocruz.br/bitstream/handle/icict/24117/Livro-Verde-07-06-2018.pdf?sequence=2&isAllowed=y>. Acesso em: 04 dez. 2022.

SÃO PAULO. Universidade de São Paulo. **A USP no Leiden Ranking**. 2017. Disponível em: <https://egida.usp.br/leiden/>. Acesso em: 29 set. 2022.

SÃO PAULO. Universidade de São Paulo. **Portaria PRP/PRPG/STI/ÁGUIA N° 001**, de 12 de dezembro de 2019. Regulamenta, no âmbito da Universidade de São Paulo, a Resolução n° 7900, de 11.12.2019. São Paulo: USP, 2019.

SÃO PAULO. Universidade de São Paulo. **Resolução n° 7900**, de 11 de dezembro de 2019 que estabelece normas para a Gestão de Dados Científicos na Universidade de São Paulo. Disponível em: <http://leginf.usp.br/?resolucao=resolucao-no-7900-de-11-de-dezembro-de-2019>. Acesso em: 01 jun. 2022.

SÃO PAULO. Universidade de São Paulo. **Resolução n° 8094**, de 09 de junho de 2021, altera dispositivo da Resolução n° 7900, de 11 de dezembro de 2019, que estabelece normas para a Gestão de Dados Científicos na Universidade de São Paulo. Disponível em: <http://leginf.usp.br/?resolucao=resolucao-no-8094-de-09-de-junho-de-2021>. Acesso em: 01 jun. 2022.

SÃO PAULO. Universidade Estadual de Campinas. **Deliberação CONSU-A-050/2020**. Institui a Política Institucional de Acesso Aberto à Produção Intelectual e Científica da Universidade Estadual de Campinas e estabelece os repositórios oficiais de depósito das produções. 2020. Disponível em: <https://www.pg.unicamp.br/norma/23869/0>. Acesso em: 24 set. 2022.

SÃO PAULO. Universidade Estadual de Campinas. **Política Institucional de Acesso Aberto à Produção Intelectual e Científica**. 2020. Disponível em: <https://www.pg.unicamp.br/norma/23869/0>. Acesso em: 02 de julh. 2023.

SÃO PAULO. Universidade Federal do ABC. **Repositório de Dados de Pesquisa da UFABC**. Disponível em: <http://dataverse.ufabc.edu.br/>. Acesso em: 15 jul. 2023.

SÃO PAULO. Universidade Estadual Paulista. **Repositório Institucional UNESP**. Disponível em: <https://repositorio.unesp.br/>. Acesso em: 02 jul. 2023.

SAYÃO, Luís Fernando; SALES, Luana. **A ciência invisível: revelando os dados da cauda longa da pesquisa**. Janeiro 2019. In: XIX ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação, Londrina/Paraná. Disponível em: [https://www.researchgate.net/publication/330449075\\_A\\_CIENCIA\\_INVISIVEL\\_REVELANDO\\_OS\\_DADOS\\_DA\\_CAUDA\\_LONGA\\_DA\\_PESQUISA](https://www.researchgate.net/publication/330449075_A_CIENCIA_INVISIVEL_REVELANDO_OS_DADOS_DA_CAUDA_LONGA_DA_PESQUISA). Acesso em: 23 jan. 2023.

SÃO PAULO. Universidade Estadual de Campinas. Comissão de Pesquisa da Universidade Estadual de Campinas. **Criação e regulamentação do repositório de dados**. 2020. Disponível em:

<https://www.sbu.unicamp.br/sbu/normas-e-regulamentos-externos/>. Acesso em: 5 ago. 2023.

SÃO PAULO. Universidade Estadual de Campinas. **Norma de Gerenciamento de Dados de Pesquisa**. Campinas: UNICAMP, 2020. Disponível em: <https://www.pg.unicamp.br/norma/23869/0>. Acesso em: 10 mar. 2023.

SAYÃO, Luís Fernando. Modelos teóricos em Ciência da Informação: abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, 2001.

SAYÃO, Luís Fernando *et al.* Princípios FAIR aplicados à gestão de dados de pesquisa. **Coleção PPGCI 50 anos**, 2021. Disponível em: [https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT\\_Principios%20FAIR%20aplicados%20a%20gest%C3%A3o%20de%20dados%20de%20pesquisa\\_2021.pdf](https://ridi.ibict.br/bitstream/123456789/1182/2/IBICT_Principios%20FAIR%20aplicados%20a%20gest%C3%A3o%20de%20dados%20de%20pesquisa_2021.pdf). Acesso em: 10 mar. 2023.

SAYÃO, Luis Fernando; MARCONDES, Carlos Henrique. Softwares livres para repositórios institucionais: alguns subsídios para a seleção. In: SAYÃO, L. (Org.) *et al.* **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

SAYÃO, Luis Fernando; SALES, Luana Farias. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **RECIIS: Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 8, n. 2, p. 76-92, 2014. Disponível em: DOI: 10.3395/reciis.v8i2.934.pt. Acesso em: 02 jun. 2022.

SAYÃO, Luís Fernando; SALES, Luana Farias. Curadoria digital e dados de pesquisa. **AtoZ: Novas Práticas em Informação e Conhecimento**, Curitiba, v. 5, n. 2, p. 121-130, 2016. Disponível em: <https://brapci.inf.br/index.php/res/download/156514>. DOI: 10.5380/atoz.v5i2.49708. Acesso em: 01 fev. 2023.

SAYÃO, Luis Fernando; SALES, Luana Farias. Dados de pesquisa: contribuição para o estabelecimento de um modelo de curadoria digital para o país. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 6, n. 1, p. 01-26, 2013. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/119469>. Acesso em: 14 jun. 2022.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: [http://www.cnen.gov.br/images/CIN/PDFs/GUIA\\_DE\\_DADOS\\_DE\\_PESQUISA.pdf](http://www.cnen.gov.br/images/CIN/PDFs/GUIA_DE_DADOS_DE_PESQUISA.pdf). Acesso em: 20 jul. 2022.

SAYÃO, Luis Fernando; SALES, Luana Farias. Gestão de dados como serviço: proposta de um modelo. In: SALES; Luana Farias; VIOLA, Carla Maria Martellote (Orgs.). **Informação digital e suas diversas abordagens pela ótica de um cientista da informação**. Rio de Janeiro: Ibict, 2021. p. 287-334. (Coleção PPGCI 50 anos). Disponível em: [https://www.gov.br/cnen/pt-br/aceso-rapido/centro-de-informacoes-nucleares/material-didatico-1/ibict\\_informacao-digital-e-suas-diversas-abordagens-pela-otica-de-um-cientista-da-informacao\\_2021\\_ibict.pdf](https://www.gov.br/cnen/pt-br/aceso-rapido/centro-de-informacoes-nucleares/material-didatico-1/ibict_informacao-digital-e-suas-diversas-abordagens-pela-otica-de-um-cientista-da-informacao_2021_ibict.pdf). Acesso em: 11 nov. 2022.

SAYÃO, Luis Fernando; SALES, Luana Farias. Subsídios para a construção de um modelo de avaliação de sistemas de gestão de dados de pesquisa. **Ponto de Acesso**, Salvador, v. 12, n. 3, p. 80-108, dez. 2018. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/28965>. Acesso em: 15 jun. 2022.

SITE DA SERRA. **Brasil lidera ranking de inovação na América Latina**. 2023. Disponível em: <https://www.sitedaserra.com.br/noticia/brasil-lidera-ranking-de-inovacao-na-america-latina>. Acesso em: 27 set. 2023.

SCIENCE-METRIX. **Analytical Support for Bibliometrics Indicators**: Open access availability of scientific publications. Montreal, 2018. Disponível em: [http://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix\\_open\\_access\\_availability\\_scientific\\_publications\\_report.pdf](http://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix_open_access_availability_scientific_publications_report.pdf). Acesso em: 2 mar. 2023.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. **Guia de citação de dados de pesquisa**. 2018. Disponível em: [https://wp.scielo.org/wp-content/uploads/guia-de-citacao-de-dados\\_pt.pdf](https://wp.scielo.org/wp-content/uploads/guia-de-citacao-de-dados_pt.pdf). Acesso em: 05 ago. 2022.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. **Guia de curadoria de dados de pesquisa para equipes editoriais**. 2022b. Disponível em: [https://wp.scielo.org/wp-content/uploads/Guia\\_curadoria\\_pt.pdf](https://wp.scielo.org/wp-content/uploads/Guia_curadoria_pt.pdf). Acesso em: 05 ago. 2022.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. **Guia de preparação de dados de pesquisa**. 2022a. Disponível em: [https://wp.scielo.org/wp-content/uploads/Guia\\_preparacao\\_pt.pdf](https://wp.scielo.org/wp-content/uploads/Guia_preparacao_pt.pdf). Acesso em: 05 ago. 2022.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. **Linhas prioritárias de ação 2019-2023**. 2018. Disponível em: [https://www.scielo20.org/redesciolo/wp-content/uploads/sites/2/2018/09/L%C3%ADneas-prioritarias-de-acci%C3%B3n-2019-2023\\_pt.pdf](https://www.scielo20.org/redesciolo/wp-content/uploads/sites/2/2018/09/L%C3%ADneas-prioritarias-de-acci%C3%B3n-2019-2023_pt.pdf). Acesso em: 05 ago. 2022.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. **Termos e condições de uso do repositório SciELO Data**. 2021. Disponível em: <https://www.scielo.org/pt/sobre-o-scielo/scielo-data-pt/termos-data/>. Acesso em: 05 ago. 2022.

SEGATTO-MENDES, Andréa Paula; MENDES, Nathan. University-industry technological cooperation for energy efficiency: a case study. **BAR-Brazilian Administration Review**, Curitiba, v. 3, p. 31-45, 2006.

SHEARER, Kathleen. **Comprehensive brief on research data management policies**. 2015. Disponível em: <http://docplayer.net/17594465-Comprehensivebrief-on-research-datamanagement-policies.html>. Acesso em: 09 set. 2022.

SINGH, Neeraj Kumar; MONU, Harpreet; DHINGRA, Navjyoti. Research data management policy and institutional framework. 2018. In: International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS), 5. 2018. **Anais [...]**. Piscataway: IEEE, 2018. p. 111-115.

SPINAK, Ernesto. Sobre as vinte e duas definições de revisão por pares aberta... e mais **SciELO em Perspectiva** [online]. 2018. Disponível em: <https://blog.scielo.org/blog/2018/02/28/sobre-as-vinte-e-duas-definicoes-de-revisao-por-pares-aberta-e-mais/>. Acesso em: 03 jun. 2022.

STEINER, João. Entrevista concedida ao Canal Univesp. **Cientistas do Brasil**. Youtube, em 6 de fev. de 2014. Disponível em: [https://www.youtube.com/watch?v=1eqHGwT5bcU&ab\\_channel=UNIVESP](https://www.youtube.com/watch?v=1eqHGwT5bcU&ab_channel=UNIVESP). Acesso em: 17 maio 2023.

SVEINSDOTTIR, Thordis; WESSELS, Bridgette A.; SMALLWOOD, Rod; LINDE, Peter; KALA, Vasso; TSOUKALA, Victoria; SONDERVAN, Jeroen. Policy recommendations for Open Access to research data in Europe: Stakeholder values and ecosystems. **Information services & use**, Bingley, v. 34, n. 3-4, p. 331-333, 2014. Disponível em: <https://www-periodicos-capes-gov-br.ez32.periodicos.capes.gov.br/index.php/buscaador-primo.html>. Acesso em: 29 jun. 2022.

THE ECONOMIST (2017, May 6). **The World's Most Valuable Resource Is No Longer Oil, but Data**. Disponível em: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. Acesso em: 15 nov. 2022.

TOMAÉL, Maria Inês; SILVA, Terezinha Elisabeth. Repositórios Institucionais: diretrizes para políticas de informação. In: Encontro Nacional de Pesquisa em Ciência da Informação, 8, 2007, Salvador. **Anais [...]**. Salvador: UFBA, 2007, p. 1-12.

TORINO, Emanuelle; ROA-MARTÍNEZ, Sandra Milena; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Dados de pesquisa: disponibilização ou publicação?. In: SHINKAKU, Milton; SALES, Luana; COSTA, Michelli (orgs.). **Tópicos sobre dados abertos para editores científicos**. Rio de Janeiro: Ibict, 2020. p. 191-207.

UK RESEARCH INTEGRITY OFFICE. **Code of Practice for Research**: Promoting good practice and preventing misconduct. 2021. Disponível em: <https://ukrio.org/wp-content/uploads/UKRIO-Code-of-Practice-for-Research.pdf>. Acesso em: 22 set. 2022.

UNESCO. **Recomendação da UNESCO sobre Ciência Aberta**. 2022. 39 p. Disponível em: [https://unesdoc.unesco.org/ark:/48223/pf0000379949\\_por](https://unesdoc.unesco.org/ark:/48223/pf0000379949_por). Acesso em: 7 fev. 2023.

UNIÃO EUROPEIA. European Research Council (ERC). European Research Council (ERC). **Guidelines on Implementation of Open Access to Scientific Publications and Research Data**: in projects supported by the European Research Council under Horizon 2020. 2017. Disponível em:

[https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf). Acesso em: 5 jan. 2023.

UNIÃO EUROPEIA. **General Data Protection Regulation**. 2016. Disponível em: <https://gdpr-info.eu/>. Acesso em: 2 abr. 2023.

UNIVERSITY OF GRONINGEN. **UG Research Data Policy**. 2021. Disponível em: <https://www.rug.nl/digital-competence-centre/ug-research-data-policy-2021.pdf>. Acesso em: 30 set. 2022.

THE NEW ENGLAND JOURNAL OF MEDICINE. More on Data Sharing. **N Engl J Med**, 12 mai. 2016 [Editorial]. Disponível em: [https://www.nejm.org/doi/full/10.1056/NEJMc1602586?query=recirc\\_curatedRelated\\_article](https://www.nejm.org/doi/full/10.1056/NEJMc1602586?query=recirc_curatedRelated_article). Acesso em: 13 maio 2023.

WALLIS, J.C.; ROLANDO, E.; BORGMAN, C.L. If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. **Plos One**, San Francisco, v. 8, n. 7, p. 1-17, 2013. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0067332>. Acesso em 08 de set. 2022.

WILKINSON, M. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Sci Data**, Londres, v. 160018, n. 3, 2016. Disponível em: <https://doi.org/10.1038/sdata.2016.18>. Acesso em: 07 jul. 2022.

## APÊNDICE A – Plano de Gestão de Dados, modelo FioDMP



O desafio da gestão de dados de pesquisa em instituições de ensino superior públicas brasileiras: diretrizes para o desenvolvimento de uma política institucional

Plano de Gestão de Dados - FioDMP

ANA LUCIA DA SILVA ARAUJO

<http://lattes.cnpq.br/8311213970578250>

Informação Administrativa

Título do Projeto

O desafio da gestão de dados de pesquisa em instituições de ensino superior públicas brasileiras: diretrizes para o desenvolvimento de uma política institucional

Plano Público ou Privado?

Público

Etapas do Plano de Gestão de Dados

Durante a execução do projeto

Data de início do projeto?

07/2022

Indique o idioma do projeto.

Português

Resumo do projeto

Este estudo abrangeu a evolução global do compartilhamento de dados de pesquisa, emergindo a necessidade de gerenciar esses dados de maneira organizada e responsável. Com a convicção de que uma política institucional para a gestão de dados de pesquisa, é essencial para estabelecer práticas obrigatórias e padronizadas, esta pesquisa buscou os elementos e as abordagens cruciais para viabilizar tais serviços. A apresentação dos resultados foi embasada em uma investigação na literatura da área, em documentos legais e institucionais, diretrizes e padrões de organizações especializadas nacionais e internacionais, realizada por meio de uma metodologia de pesquisa qualitativa, exploratória e aplicada. Com o intuito de compreender a abrangência e as lacunas no cenário nacional em relação ao tema, foram coletadas informações de instituições de ensino superior brasileiras, bem como de outras instituições com as quais se relacionam, como institutos de pesquisa e agências de fomento. Para atingir o principal objetivo da pesquisa, foram analisadas políticas de gestão de dados de pesquisa em países que já têm políticas instituídas, nas quais instituições de ensino e pesquisa serviram como modelos de referência.

Dentre essas instituições, as universidades da Holanda, reconhecidas por seu comprometimento com os princípios da Ciência Aberta, foram escolhidas para extrair elementos e abordagens de seus documentos. Por meio de análises individuais e comparativas, foi elaborada uma planilha para identificar os elementos comuns e singulares dessas políticas. A partir desse levantamento, foi construído um quadro com elementos e abordagens possíveis que podem ser aplicáveis na construção de políticas para as instituições de ensino e pesquisa no Brasil. Foi constatado que há diferenças significativas entre os dois países em relação a apoio de órgãos públicos, infraestrutura, regulamentações de nível nacional e serviços federados para viabilização e apoio aos serviços. Concluiu-se que a elaboração de políticas institucionais e gestão de dados nas instituições de ensino superior do Brasil carecem de apoio dos órgãos públicos, do conhecimento e reconhecimento do valor dos dados de pesquisa de suas lideranças institucionais.

Palavras-chave: Gestão de dados de pesquisa. Políticas para gestão de dados de pesquisa - Brasil

O projeto tem financiamento?

MCTIC - Ministério da Ciência, Tecnologia, Inovações e Comunicações

Sobre o Pesquisador

Nome completo do pesquisador principal

ANA LUCIA DA SILVA ARAUJO

Afiliação do pesquisador principal

BIBLIOTECÁRIA

E-mail do pesquisador principal

[analuciaeci@gmail.com](mailto:analuciaeci@gmail.com)

Link do Currículo Lattes do pesquisador

<http://lattes.cnpq.br/8311213970578250>

Descrição dos dados coletados ou reuso de dados existentes

Seus dados serão:

Coletados e/ou Produzidos

Critérios para selecionar instituições por meio de índices de avaliação. Pesquisa exploratória nos sites das instituições.

Quais os tipos de dados que serão coletados, produzidos ou reutilizados?

Imagem Numéricos Textuais

Quais os formatos de dados que serão coletados, produzidos ou reutilizados?

CSV - Comma-Separated Values DOC - Document Format

PDF - Portable Document Format PNG - Portable Network Graphics TXT - Text Format

XLS - Excel Binary File Format

Qual o volume aproximado dos dados coletados, produzidos ou reutilizados?

Maior que 1 GB e menor que 10 GB

Indique quais metadados serão utilizados para identificar os dados, de forma a facilitar que outras pessoas os encontrem.

Títulos: texto; imagem; planilhas. Autor: nome da autora e orientadora. Data de publicação dos arquivos; data de início e fim da pesquisa. Palavras-chave de cada conjunto de dados. Descrição do conteúdo; Formatos abertos; Proprietário; Licenças e identificadores persistentes.

Armazenamento e Backup durante o processo de pesquisa

---

Onde os dados serão armazenados?

Discos Rígidos Autônomos; Laptops; Pcs; Serviços de Nuvem Privado

Requisitos Legais e Éticos

---

A equipe tem ciência das normas relacionadas à avaliação ética na pesquisa?

Sim

A pesquisa envolve tratamento de dados em que se exija avaliação ética?

Não

A equipe está familiarizada com a legislação sobre o tratamento e proteção dos dados pessoais?

Sim

A pesquisa envolve tratamento de dados pessoais?

Não

Há previsão de período de embargo relacionado ao direito autoral?

Não

A pesquisa envolve outras categorias de dados que exijam período de embargo para sua disponibilização?

Não

A pesquisa envolve outras categorias de dados que exijam período de embargo para sua disponibilização?

Não

Informe a licença a ser aplicada à base de dados

CC-0 - Creative Commons Public Domain Dedication

Compartilhamento de Dados



---

A pesquisa prevê compartilhamento de dados?

Sim

Após a conclusão da pesquisa

Existem possíveis motivações para submeter os dados à períodos de embargo no compartilhamento de dados?

Não

## APÊNDICE B – Plano de Gestão de Dados, modelo DMPtool (sem atualização)

---

### Plan Overview

A Data Management Plan created using DMPTool

Title: Política Institucional de Gestão e Compartilhamento de Dados de Pesquisa: construção teórica de um modelo para as universidades brasileiras

Creator: Ana lúcia da silva Araújo

Affiliation: Federal University of Rio de Janeiro

(uffj.br) Funder: Digital Curation Centre

(dcc.ac.u)

Template: Digital Curation Centre

### Project abstract:

A demanda crescente pelo Acesso aos Dados de Pesquisa se tornou uma tendência mundial e formas para gerenciar este recurso passou a ser uma preocupação de instituições de diversas categorias, principalmente, das que coletam e geram dados por meio de pesquisas científicas. O processo de gerenciamento requer uma prévia elaboração de Plano de Gestão de Dados, onde o pesquisador deve apresentar informações mínimas sobre quais dados serão gerados/coletados e gerenciados. A questão será problematizada dentro da realidade brasileira, com consulta a diversas instituições nacionais e internacionais. Este projeto pretende apresentar/propor um modelo teórico de Política Institucional para a Gestão e Compartilhamento de Dados de Pesquisa, voltando para as especificidades das universidades públicas brasileiras. Como maiores produtoras de pesquisa e, conseqüentemente, maiores geradoras de dados, as universidades têm muito a contribuir com a sustentabilidade e responsabilidade social na abertura e compartilhamento de dados, promovendo o acesso à informação e economia de recursos, tempo e energia, evitando o retrabalho para novas pesquisas. Para alcançar os objetivos do projeto, pretende-se conhecer os conceitos e abordagens que envolvem a gestão de dados de pesquisa para construir uma política eficiente e qualificada. Serão levantadas as normativas de instituições nacionais sobre gestão e compartilhamento de dados de pesquisa para extrair seus elementos; selecionar um conjunto de políticas internacionais para levantar seus itens e contextos. Ao reunir os elementos essenciais, serão conhecer os itens em comum e distintos entre as instituições para análise individual e comparativa. Assim, será averiguada a possibilidade de sua inclusão e/ou adaptação. Como resultado, pretende-se apresentar uma proposta de modelo teórico de políticas para gestão de dados de pesquisa para as universidades públicas brasileiras.

Start date: 10-07-2022

End date: 03-31-2023

Last modified: 10-07-2022

### APÊNDICE C – A reunião dos conceitos

Termos	Conceitos	Fontes
Acesso aberto	Acordo em que autores e detentores de direitos autorais garantem a todos os usuários o direito irrevogável, universal e vitalício de acesso grátis, e uma licença para copiar, usar, distribuir, transmitir e apresentar publicamente o trabalho e de produzir e distribuir trabalhos dele derivados em qualquer meio digital para qualquer propósito responsável, sujeito à atribuição de autoria apropriada, bem como ao direito de fazer poucas cópias impressas para seu uso pessoal.	Decs - Descritores da Área da Saúde
Administrador de dados	Pessoa responsável por manter a qualidade, integridade e acordos de acesso de dados e metadados de forma consistente com a lei aplicável, política institucional e permissões individuais. Visa garantir que os dados sejam tratados adequadamente em todas as etapas do ciclo de pesquisa (ou seja, design, coleta, processamento, análise, preservação, compartilhamento de dados e reutilização)	LCRDM (201-)
Administrador de dados	Alguém responsável pelo cuidado profissional e cuidadoso dos dados durante todas as fases da pesquisa. O Data Steward garante o gerenciamento de dados sustentável e de longo prazo.	Universidade de Utrecht (2019)
Administrador de dados	Os administradores de dados ajudam os pesquisadores em seus institutos ou faculdades com questões relacionadas ao GDP	Universidade de Amsterdam (2019)
Administrador de dados	Funcionário, geralmente um pesquisador, responsável pelo gerenciamento de dados de pesquisa em um ou mais projetos de pesquisa. É responsável pelo desenvolvimento e alinhamento da implementação do regulamento de Gestão de Dados de Investigação e aconselha os investigadores na recolha, armazenamento, arquivo e reutilização de dados na Escola	Universidade de Tilburg (2020)
Administrador de dados (Gestor de dados; Gerenciador de dados)	Alguém que lida com um ou mais aspectos do gerenciamento de dados. O conteúdo das tarefas pode, portanto, variar muito dependendo da função do gerenciador de dados. Esses processadores estão cientes do conteúdo desta política, leis e regulamentos atuais, possibilidades e facilidades na área de gerenciamento de dados. Consulte também Administrador de Dados.	Universidade de Utrecht (2019)
Administrador de dados	Pessoa responsável pelo gerenciamento de objetos de dados, incluindo metadados. Devem ser especialistas em informação, como arquivistas, bibliotecários e responsáveis pela conformidade dos dados	LCRDM (201-)
Anonimização	Desvinculação irreversível de informações de identificação de dados associados.	LCRDM (201-)
Anonimização	Processo pelo qual passam dados que contêm informações pessoais para que não revelem a identidade de indivíduos e evitem que esses dados não estejam vinculados a outros dados que possam revelar a identidade desses indivíduos	Sayão e Sales (2015)
Anonimização	Utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo	BRASIL. LGPD (2018)
Armazenamento	Ação de dispor ou guardar documentos nos lugares predeterminados	Cunha e Cavalcanti (2008)
Armazenamento	A colocação e guarda de coisas em um lugar especial para uso no futuro	Dicionário Cambridge
Armazenamento	Serviços e funções usados ??para o armazenamento e recuperação de pacotes de informações arquivísticas	Data Curation Centre
Armazenamento	Refere-se à série de atividades gerenciadas necessárias para garantir o acesso contínuo a materiais digitais pelo tempo que for necessário	Universidade de Leiden (2021)
Armazenamento	Gestão de arquivos de backup que podem ser criados manualmente ou por sistemas automatizados de armazenamento. Eles são exigidos para restaurar arquivos que tenham sido corrompidos, perdidos, alterados irreversivelmente ou destruídos	Sayão e Sales (2015)
Armazenamento	Atividades organizadas relacionadas com a estocagem, localização, busca e recuperação de informação. Fontes de dados: Sistemas de registro, arquivos, banco de dados, documentos a partir dos quais se pode obter dados referentes à determinada população.	Decs - Descritores da Área da Saúde
Armazenamento	Conjunto de dados que está submetido a um processo de arquivamento é geralmente um conjunto de registros que não sofrerá mais mudanças, é como	Sayão e Sales (2015)

	uma fotografia histórica. Ele será preservado para usos futuros, portanto ele precisa ser recuperável e estar armazenado em formatos de arquivos estáveis, amplamente usados, padronizados e abertos	
Armazenamento	Um local de armazenamento físico onde um objeto de dados será armazenado após a ingestão em um repositório de dados. Isso requer a identificação do endereço IP e o nome do caminho físico dentro do local de armazenamento onde um objeto de dados será armazenado. A sequência dessas atividades encadeadas é conceituada como um objeto de fluxo de trabalho. Para recuperação, o local do objeto de dados é especificado pelo local de armazenamento e pelo nome do caminho físico.	CODATA (2022)
Arquivamento	Uma atividade de curadoria que garante que os dados sejam adequadamente selecionados, armazenados e possam ser acessados, e para os quais a integridade lógica e física seja mantida ao longo do tempo, incluindo segurança e autenticidade	CODATA (2022)
Arquivamento	Um arquivo de dados é uma instalação que move os dados para um ambiente para retenção de longo prazo. Um arquivo de dados é indexado e possui recursos de pesquisa, permitindo que os dados sejam recuperados	LCRDM (201-)
Arquivamento	O termo é usado quando um arquivo deve ser preservado como está, ou quando se deseja um registro do histórico do arquivo. Geralmente, o arquivamento lida com registros que estão na sua versão final e faz parte das etapas requeridas para a preservação dos dados para necessidades futuras, ou seja, para a preservação de longo prazo. Nessa direção, o arquivamento é realizado, via de regra, quando o projeto termina.	Sayão e Sales (2015)
Arquivamento	Atividades de gestão necessárias à manutenção de dados de pesquisa de forma a assegurar que eles sirvam para uso contemporâneo e que estejam disponíveis para descoberta e reuso	Decs - Descritores da Área da Saúde
Arquivamento	Preservação de documentos em que a versão digital é utilizada para arquivamento ou preservação	Cunha e Cavalcanti (2008)
Artigo de dados / Data paper	Tipo de publicação cujo principal objetivo é descrever uma coleção ou conjunto de dados de pesquisa. Contêm somente informação sobre os dados, não se estendendo a hipóteses, argumentos, interpretações, deduções e conclusões	SciELO
Ativo de informação	Qualquer dado, informação ou conhecimento a que esteja associado um valor para o negócio da organização	Embrapa (2021) (2019)
Auditoria de Dados	Verificação das práticas institucionais de coleta, preservação e divulgação de conjuntos de dados, resultando em sugestões para melhoria dos procedimentos atuais.	LCRDM (201-)
Auditoria de Dados	Uma avaliação independente de uma organização, sistema, processo, projeto ou produto.	CODATA (2022)
Backup	Várias cópias de vários arquivos tendo conhecimento que os arquivos podem mudar. Dessa forma, as cópias de backup podem ser guardadas por certo período de tempo, mas podem ser descartadas quando for conveniente	Sayão e Sales (2015)
Backup	Ampla variedade de tópicos, incluindo backup, arquivamento, preservação e proteção física, criptografia e ainda as leis que governam a proteção dos dados	LCRDM (201-)
Catálogo	Um processo intelectual de descrição de objetos de acordo com os princípios aceitos da biblioteca, particularmente aqueles de assunto e ordem de classificação.	CODATA (2022)
Ciclo de vida dos dados	Etapas que serão realizadas durante o processo de vida dos registros de dados de pesquisa, desde o seu planejamento até o seu arquivamento ou descarte, para garantir que eles possam ter o seu uso, reuso e compartilhamento otimizado e estendido	Fiocruz (2017)
Ciência Aberta	Construção inclusiva que combina vários movimentos e práticas visando tornar o conhecimento científico multilíngue disponível, acessível e reutilizável para todos, aumentar as colaborações científicas e o compartilhamento de informações em benefício da ciência e da sociedade e abrir os processos de criação, avaliação e comunicação do conhecimento científico a atores sociais além da comunidade científica tradicional.	UNESCO (2021)
Ciência de dados	Campo interdisciplinar que envolve processos, teorias, conceitos, ferramentas e tecnologias que permitem a revisão, análise e a extração de	Decs - Descritores da Área da Saúde

	conhecimento e informação de valor a partir de dados estruturados e não estruturados (dados não tratados).	
Cientista de dados	São uma nova geração de especialistas analíticos que têm as habilidades técnicas para resolver problemas complexos - e a curiosidade de explorar os problemas que precisam ser resolvidos. Podem ser matemáticos, estatísticos, cientistas da computação e/ou analistas de tendências	Fiocruz (2017)
Cientista de dados	Um praticante da ciência de dados. É um termo genérico que abrange muitos campos de especialização. No relatório atual, analistas de dados, engenheiros de dados, administradores de dados e engenheiros de software de pesquisa são considerados subgrupos de cientistas de dados. Em certos contextos, o cientista de dados às vezes também é usado de maneiras mais limitadas que o tornam equivalente às funções de analista de dados ou engenheiro de software	LCRDM (201-)
Classificação dos dados	Ação de definir o grau de sigilo e os critérios adequados para a proteção da informação, observado seu teor, criticidade e valor	Embrapa (2021)
Comitê de direção	Grupo responsável por garantir que as metas do programa sejam alcançadas e fornecer suporte para lidar com os riscos e problemas do programa	CODATA (2022)
Comitê Gestor	Colegiado responsável pela deliberação sobre questões técnicas, administrativas, infra-estruturais e operativas que venham a ocorrer durante a operacionalização e gerenciamento de dados e informações	PPBio (2009)
Compartilhamento de Dados	Acesso aos dados para fins de pesquisa ou análise por outros pesquisadores ou pessoas interessadas	LCRDM (201-)
Compartilhamento de dados	O ato de disponibilizar dados científicos para uso por outros (por exemplo, a grande comunidade de pesquisa, instituições, o público em geral), por exemplo, por meio de um repositório estabelecido.	Borgam e Bourne (2021)
Confidencialidade	Os deveres e práticas de pessoas e organizações para garantir que as informações pessoais dos indivíduos fluam apenas de uma entidade para outra de acordo com a legislação ou normas e políticas amplamente aceitas.	CODATA (2022)
Conjunto de dados	Termo usado para se referir aos arquivos de dados de pesquisa e seus metadados	SciELO
Curadoria de Dados	Um local de armazenamento físico onde um objeto de dados será armazenado após a inserção em um repositório de dados. Isso requer a identificação do endereço IP e o nome do caminho físico dentro do local de armazenamento onde um objeto de dados será armazenado. A sequência dessas atividades encadeadas é conceituada como um objeto de fluxo de trabalho. Para recuperação, o local do objeto de dados é especificado pelo local de armazenamento e pelo nome do caminho físico.	CODATA (2022)
Curadoria de Dados	Processo de revisão dos dados e metadados realizado por um curador que garante que os conjuntos de dados estejam adequadamente descritos e atendam aos critérios definidos pelo repositório de dados de pesquisa	SciELO
Curadoria de Dados	A curadoria de dados de pesquisa é o processo de gerenciamento de dados de pesquisa durante todo o seu ciclo de vida para disponibilidade em longo prazo e reusabilidade .	Dudziak (2019)
Curadoria de Dados	Atividade de gerenciamento do uso de dados desde seu ponto de criação para garantir que estejam disponíveis para serem encontrados e reutilizados no futuro. Ou o processo de seleção, anotação, manutenção, arquivamento e rastreamento de dados	LCRDM (201-)
Curadoria de dados	serviço que envolve todos os processos aplicados a objetos digitais durante o seu ciclo de vida. Compreende as atividades de estabelecimento de padrões para conjunto de dados, adição de valor, gestão de risco e boas práticas na gestão de dados digitais. Trata-se de um conceito mais inclusivo que o arquivamento digital e a preservação digital.	Costa (2017)
Curadoria de dados	Ações realizadas pelos curadores com o objetivo de garantir que o conjunto de dados esteja estruturado e documentado da forma mais completa possível e seguindo as boas práticas	SciELO
Curadoria de Dados	Atividades de gestão necessárias à manutenção de dados de pesquisa de forma a assegurar que eles sirvam para uso contemporâneo e que estejam disponíveis para descoberta e reuso.	Decs - Descritores da Área da Saúde
Curadoria de Dados	Gerenciamento e aprimoramento de ativos de informação digital para uso atual e futuro	National Research Council (2015)

Curadoria de Dados	Ações voltadas para o gerenciamento de dados de pesquisa durante o seu ciclo de vida; envolve manter, preservar e adicionar valor aos dados.	Sayão e Sales (2015)
Curadoria de Dados	São ações voltadas para o gerenciamento de dados de pesquisa durante o seu ciclo de vida. Envolve manter, preservar e adicionar valor aos dados	Fiocruz (2017)
Dado público	Qualquer dado gerado ou sob a guarda governamental que não tenha o seu acesso restrito por legislação específica	IBGE (2020)
Dados	Sequência de símbolos ou valores, produzidos como resultado de um processo natural ou artificial e representados em qualquer meio	Embrapa (2021)
Dados abertos	São dados publicados em um formato legível por máquina e sem restrição de licenças, patentes ou mecanismos de controle, de modo a estarem livremente disponíveis para serem utilizados e redistribuídos à vontade	Fiocruz (2017)
Dados abertos	Dados disponibilizados à sociedade completos, acessíveis, legíveis por máquina e mantidos no formato mais primário, sempre que possível, de modo a facilitar a reutilização, a criação de valor e novos modelos de negócio para a prestação de serviço	Embrapa (2021)
Dados abertos	Dados públicos representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na rede mundial de computadores e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento	IBGE (2020)
Dados abertos	Dados que podem ser usados por qualquer pessoa sem restrições técnicas ou legais. O uso abrange tanto o acesso quanto a reutilização	SciELO
Dados abertos	Dados públicos representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na rede mundial de computadores e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento.	Jardim Botânico do Rio de Janeiro (2021)
Dados abertos	Dados estruturados que são acessíveis, legíveis por máquina, utilizáveis, inteligíveis e compartilhados livremente. Os dados abertos podem ser livremente usados, reutilizados, construídos e redistribuídos por qualquer pessoa sujeitos apenas, no máximo, ao requisito de atribuição e compartilhamento.	CODATA (2022)
Dados brutos	Dado que vem diretamente dos instrumentos científicos ou coletados diretamente da fonte sem sofrer nenhuma manipulação ou processamento.	Sayão e Sales (2015)
Dados brutos	São os que vêm diretamente dos instrumentos científicos ou coletados diretamente da fonte, sem sofrer qualquer manipulação ou processamento. São também chamados de dados crus ou dados primários	Santos et al. (2017)
Dados Brutos	Todos os dados brutos e não processados em registros originais e cópias autenticadas de documentos de registro originais relacionados a descobertas, observações e outras atividades em um estudo necessários para a reconstrução e avaliação do estudo.	Universidade de Utrecht (2019)
Dados compartilhados	Dados de pesquisa cujo acesso, reuso e redistribuição estão limitados a determinadas pessoas ou grupos	SciELO
Dados confidenciais	São os que não estão em domínio público, tais como informações sobre negócios, lucros, saúde, detalhes médicos e opiniões políticas, entregues em confiança ou quando duas partes concordaram em mantê-los confidenciais, isto é, secretos	Santos et al. (2017)
Dados de pesquisa	Todas as informações (independente de forma ou apresentação) necessárias para apoiar ou validar o desenvolvimento, resultados, observações ou descobertas de um projeto de pesquisa, incluindo informações contextuais e dados secundários	Universidade Erasmus de Rotterdam (2020)
Dados de pesquisa	Todos os dados (digitais e não digitais) coletados e gerados durante a pesquisa acadêmica, os instrumentos com os quais os dados da pesquisa foram coletados e outras informações relevantes, como questionários, softwares, roteiros e diários de laboratório	Universidade de Tilburg (2020)
Dados de pesquisa	Dados, tanto físicos quanto digitais, que são coletados, criados e/ou usados no âmbito de um projeto de pesquisa e (parcialmente) destinados a produzir e validar os resultados da pesquisa. No contexto do gerenciamento de dados de pesquisa, também materiais relacionados a dados, como Armazenamento formalmente aprovado como compatível com um ou mais padrões.	Universidade de Twente (2018)

Dados de pesquisa	Registros factuais (pontuações numéricas, registros textuais, imagens e sons) produzidos ou utilizados como fontes primárias para a pesquisa científica e tecnológica e que são necessários para validação dos seus resultados.	Embrapa (2021)
Dados de pesquisa	Informações registradas (digitais) que, independentemente de sua forma, são necessárias para apoiar ou validar os resultados da pesquisa.	Universidade de Amsterdam (2019)
Dados de pesquisa	O material factual registrado comumente aceito na comunidade científica como necessário para validar os resultados da pesquisa	Universidade de Leiden (2021); Jardim Botânico do Rio de Janeiro (2021)
Dados de pesquisa	O material factual registrado comumente aceito na comunidade científica como necessário para validar os resultados da pesquisa	Jardim Botânico do Rio de Janeiro (2021)
Dados de pesquisa	Unidades de informação criadas ou coletadas no curso da pesquisa científica, que são frequentemente formatadas de maneira a torná-las adequadas à comunicação, à interpretação e ao processamento por computador.	Santos et al. (2017)
Dados de Pesquisa	Unidades de informação criadas ou coletadas no curso da pesquisa científica, e que são frequentemente formatadas de maneira a torná-las adequadas à comunicação, interpretação e processamento por computador. São exemplos de dados de pesquisa: planilhas de estatísticas, uma série de mensagens de email, um registro sonoro de uma entrevista, um registro descritivo de um espécime de rocha, uma coleção de imagens digitais.	Sayão e Sales (2015)
Dados de Pesquisa	Material factual registrado comumente aceito na comunidade científica como de qualidade suficiente para validar e replicar os resultados da pesquisa, independentemente de os dados serem usados para apoiar publicações acadêmicas	Borgam e Bourne (2021)
Dados digitais	Dados na forma de materiais digitais	CODATA (2022)
Dados fechados	Dados de pesquisa que não podem ser livremente usados, reutilizados e redistribuídos	SciELO
Dados pessoais	Qualquer informação relativa a uma pessoa singular identificada ou identificável. Uma pessoa é identificável se sua identidade puder ser razoavelmente determinada, sem esforço desproporcional. Isso se refere à identidade física, não de pessoas falecidas	Universidade de Utrecht (2019)
Dados pessoais	Relacionados à pessoa natural identificada ou identificável aquela que pode ser reconhecida, direta ou indiretamente, a partir de um identificador como um nome, número de identificação, dados de localização, identificador online ou um ou mais fatores específicos da identidade física, fisiológica, genética, mental, econômica, cultural ou social dessa pessoa natural	Embrapa (2021)
Dados pessoais	Dados pessoais referem-se a qualquer informação que possa ser rastreada até uma pessoa. Essas informações podem ser um nome, endereço ou localização, mas também podem ser números de contas bancárias, números de telefone ou códigos postais com números de residência.	Universidade de Leiden (2021)
Dados pessoais	Informação relacionada a pessoa natural identificada ou identificável	BRASIL. LGPD (2018)
Dados pessoais	Relacionados à pessoa natural identificada ou identificável aquela que pode ser reconhecida, direta ou indiretamente, a partir de um identificador como um nome, número de identificação, dados de localização, identificador online ou um ou mais fatores específicos da identidade física, fisiológica, genética, mental, econômica, cultural ou social dessa pessoa natural	Embrapa (2021)
Dados pessoais	São informações relacionadas a indivíduos vivos, que podem ser identificados a partir deles ou a partir da combinação com outras informações	Santos et al. (2017)
Dados pessoais	Os dados que, sozinhos ou em combinação com outros dados, podem razoavelmente identificar um indivíduo.	LCRDM (201-)
Dados processados	Resultados (incluindo resultados negativos e inconclusivos) após o processamento de dados (recodificados, visualizados, categorizados, agregados, etc.)	Universidade de Rotterdam (2020)
Dados sensíveis	Dados preliminares ou consolidados que, se liberados ao acesso público, possam resultar em efeito adverso ao local e/ou às comunidades de origem da mesma e por isso, passível de restrição.	PPBio (2009)

Dados sensíveis	São informações sobre raça, origem étnica, opinião política, religião ou crenças similares, filiação sindical, doença física ou mental, vida sexual etc.	Santos et al. (2017)
Dataset (Conjunto de Dados)	Uma coleção de dados que pode ser um subconjunto em um banco de dado	LCRDM (201-)
Desidentificação	Remoção ou alteração de quaisquer dados que identifiquem uma pessoa singular ou possam, previsivelmente, vir a identificar uma pessoa singular no futuro	LCRDM (201-)
Destruição de dados	O processo de destruição de dados armazenados em fitas, discos rígidos e outras formas de mídia eletrônica para que fiquem completamente ilegíveis e não possam ser acessados ??ou usados	CODATA (2022)
Digital Object Identifier (DOI)	Um tipo de identificador persistente (dentre outros) usado para identificar objetos. O sistema DOI é usado, particularmente, para documentos eletrônicos.	LCRDM (201-)
Digital Object Identifier (DOI)	Identificador persistente que é usualmente assinalado a itens digitais como um artigo de periódico ou uma coleção de dados, com o objetivo de identificá-los univocamente e dessa forma serem descobertos e citados.	Sayão e Sales (2015)
Diretor de dados	Um executivo corporativo responsável pela governança em toda a empresa e pela utilização de informações como um ativo, por meio de processamento de dados, análise, mineração de dados, troca de informações e outros meios.	CODATA (2022)
Acesso restrito	disponibilização de dados sensíveis permitida mediante autorização ou senha de acesso.	PPBio (2009)
Documentação	Informação sobre a coleção de dados; pode estar incorporado aos dados ou existir separadamente; metadados podem descrever, por exemplo, a autoria, direitos de propriedade, propósitos, métodos, organização e condições de uso dos dados, informações técnicas dos dados e outras informações necessárias à compreensão dos dados	Sayão e Sales (2015)
Documentação	Informações estruturadas padronizadas explicando itens de dados como, mas não limitados a: finalidade, origem, referências de tempo, localização geográfica, criador, condições de acesso e termos de uso de uma coleta de dados.	LCRDM (201-)
Documentação	Quaisquer arquivos digitais, como um codebook, relatório técnico ou metodológico ou guia do usuário, que expliquem a produção, proveniência, processamento ou interpretação dos dados da pesquisa.	CODATA (2022)
E-science	Conceito relacionado às novas práticas de compartilhamento da produção do conhecimento científico e uso de plataformas tecnológicas para a pesquisa colaborativa. Representa uma ciência atuante no uso de dados por meio de computação avançada	Santos et al. (2017)
Embargo	dados de pesquisa cujo acesso e reuso será liberado após um período de tempo predeterminado	SciELO
Embargo	período no qual, dados sob restrições de uso e acesso, não são disponibilizados pelo portal, mas são passíveis de visualização pelo Comitê Gestor de Informação	PPBio (2009)
Embargo	Período de tempo no qual o acesso e o uso dos dados e da informação podem estar restritos por um período determinado. Tem o objetivo de proteger o interesse dos proprietários dos dados, pesquisadores e organizações e editores científicos	Santos et al. (2017)
Embargo	Restrição ao acesso e publicação temporária e necessária para garantir o tratamento, análise e utilização em publicação original por parte dos seus autores	Jardim Botânico do Rio de Janeiro (2021)
Embargo	período no qual, os dados sob restrições de uso e acesso, não são disponibilizados, mas são passíveis de visualização pelo Comitê Gestor de Informação	Sayão e Sales (2015)
Entrevistador de Dados	Conversa que um curador/consultor de dados tem com um pesquisador para obter as informações necessárias sobre um conjunto de dados	LCRDM (201-)
Fluxo de trabalho	Um tipo de fluxo de trabalho que inclui etapas ativas para selecionar dados como uma ajuda para o gerenciamento contínuo de dados ao longo de seu ciclo de vida	CODATA (2022)



Fluxo de trabalho	Atividades e processos em um ambiente digital que levam à publicação de dados de pesquisa, metadados associados e documentação que os acompanha e código de software na Web.	CODATA (2022)
Formato aberto	Formato de arquivo não proprietário, cuja especificação esteja documentada publicamente e seja de livre conhecimento e implementação, livre de patentes ou qualquer outra restrição legal quanto à sua utilização	Universidade de Leiden (2021); Jardim Botânico do Rio de Janeiro (2021)
Formato Aberto	Organização específica da informação em um arquivo digital aberto, que permite edição para reutilização	Sayão e Sales (2015)
Formato Aberto	A maneira como os dados/informações são codificados e armazenados.	LCRDM (201-)
Gestão de dados	Processos que incluem a aquisição, validação, armazenamento, proteção e processamento de dados para garantir a acessibilidade, confiança e pontualidade para usuários.	Decs - Descritores da Área da Saúde
Gestão de dados	Processo que contempla as atividades de planejamento, aquisição, organização, estruturação, curadoria e análise de dados, utilizando para isso ferramenta computacional apropriada para o armazenamento de dados, levando em consideração as questões relativas à preservação, à organização, ao compartilhamento, à proteção e à confidencialidade dos mesmos, bem como o acesso e disponibilização para a sociedade quando cabível.	Embrapa (2021)
Segurança da informação	Conjunto de processos que permitem identificar e implementar as medidas de proteção necessárias para minimizar ou eliminar os riscos a que estão sujeitos os ativos de informação da Embrapa, e equilibrá-los com os custos operacionais e financeiros envolvidos.	Embrapa (2021)
Gestão de dados	Uma declaração formal descrevendo como os dados de pesquisa serão gerenciados e documentados ao longo de um projeto de pesquisa e os termos relativos ao depósito subsequente dos dados em um repositório de dados para gerenciamento e preservação a longo prazo	Universidade de Leiden (2021)
Gestão de dados	Qualquer ação ou conjunto de ações relativas a dados, incluindo coletar, organizar, armazenar, atualizar, modificar, recuperar, consultar, usar, fornecer por meio de encaminhamento, divulgação ou qualquer outra forma de disponibilizar, reunir, inter-relacionar, bem como como a proteção, apagamento ou destruição de dados	Universidade de Utrecht (2019)
Gestão de Dados	Organização dos dados de pesquisa, desde sua entrada até o ciclo de pesquisa até a disseminação e arquivamento de resultados valiosos	Universidade de Amsterdam (2019)
Gestão de dados	Conjunto de práticas de gestão voltadas para o tratamento de dados de pesquisa durante o seu ciclo de vida; inclui todos os aspectos de manutenção, compartilhamento, segurança e preservação	Santos et al. (2017)
Gestão de dados	Refere-se ao armazenamento, acesso e preservação dos dados produzidos a partir de uma determinada investigação. Abrange o ciclo de vida dos dados e incluem: Nomenclatura de arquivos (a maneira correta de nomear arquivos de computador); controle de qualidade de dados e garantia de qualidade; Acesso de dados; documentação de dados (incluindo níveis de incerteza); criação de metadados e vocabulários controlados; armazenamento de dados; arquivamento e preservação de dados; compartilhamento e reutilização de dados; integridade de dados; segurança de dados; dados privados; direitos de dados; protocolos de notebook (laboratório ou campo).	CODATA (2022)
Gestão de dados	Conjunto de atividades intrínsecas ao processo de tratamento técnico (curadoria), armazenamento, recuperação, disseminação e preservação dos dados coletados pela e-science.	Costa (2017)
Gestão de dados	gerenciamento de dados de pesquisa é parte integrante do processo de pesquisa, que diz respeito à maneira como você coleta, analisa, armazena, compartilha, arquiva e publica dados de pesquisa, para satisfazer as necessidades dos usuários de dados atuais e futuros	Universidade de Leiden (2021)
Governança	Organização e implementação de políticas, procedimentos, estruturas, culturas, papéis e responsabilidades que delinham, viabilizam e transformam os processos de gestão para atender às necessidades atuais e futuras dos públicos interno e externo da organização.	Embrapa (2021)
Governança	O processo de formulação e gestão de políticas que orienta e supervisiona a pesquisa de maneira consistente e estruturada	LCRDM (201-)

Governança	Envolve políticas, processos e pessoas, além de uma estrutura organizacional para apoiar o gerenciamento de dados institucionais.	Fiocruz
Governança	Garante que os benefícios da pesquisa para a sociedade superam quaisquer riscos, tanto do ponto de vista ético quanto legal.	CODATA (2022)
Governança	Um modelo que descreve como são investidas as responsabilidades pelas várias fases do Gerenciamento de Dados de Pesquisa, quais documentos e procedimentos de política se seguem, quem é responsável por eles, qual infraestrutura foi desenvolvida e qual suporte é oferecido.	Universidade de Utrecht (2019)
Identificador persistente	Uma referência a longo prazo de um documento, arquivo, página da Web ou outro objeto (real ou abstrato)	LCRDM (201-)
Identificador persistente	Sequência de caracteres que identifica uma entidade. O termo 'identificador persistente' é geralmente usado no contexto de objetos digitais acessíveis pela Internet. Normalmente, tal identificador não é apenas persistente, mas também acionável, ou seja, é um Uniform Resource Identifier (URI)?, geralmente do tipo http/s, que você pode colar em uma barra de endereço do navegador da Web para ser levado para o fonte identificada	Wilkinson, M.D et al. (2016)
Identificador persistente	Referência duradoura a um objeto digital que fornece informações sobre esse objeto, independentemente do que aconteça com ele. Pode fornecer uma representação apropriada de um objeto, quer esse objeto mude sua localização online ou fique offline.	CODATA (2022)
Identificador persistente	É usualmente assinalado a itens digitais como um artigo de periódico ou uma coleção de dados, com o objetivo de identificá-los univocamente e dessa forma serem descobertos e citados	Santos et al. (2017)
Infraestrutura de gerenciamento de dados de pesquisa	A configuração da equipe, serviços e ferramentas reunidas para apoiar o gerenciamento de dados em todo o ciclo de vida da pesquisa e, mais especificamente, para fornecer uma cobertura abrangente dos estágios que compõem o ciclo de vida dos dados. Pode ser organizado local e/ou globalmente para apoiar atividades de dados de pesquisa em todo o ciclo de vida da pesquisa.	CODATA (2022)
Infraestrutura para dados científicos	O que é necessário para permitir que os pesquisadores criem, armazenem e compartilhem os dados resultantes de seus experimentos e encontrem, acessem e processem os dados de que precisam	CODATA (2022)
Integridade dos dados	Garantia de que a informação seja mantida em seu estado original, visando protegê-la, na guarda ou transmissão, contra alterações indevidas, intencionais ou acidentais.	Embrapa (2021)
Interoperabilidade	É a capacidade de diversos sistemas (informatizados ou não) e conteúdos se comunicarem de forma transparente (ou o mais próximo disso) com outro sistema (semelhante ou não). Para um sistema ser considerado interoperável, é necessário que ele trabalhe com padrões abertos e/ou ontologias	Santos et al. (2017)
Interoperabilidade	A capacidade de dados ou ferramentas de recursos não cooperativos para integrar ou trabalhar em conjunto com o mínimo de esforço	LCRDM
Licença de uso	Acordo de fornecimento de dados que conceda amplo acesso para que qualquer pessoa os utilize, os reutilize, e os redistribua, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença	IBGE (2020); Jardim Botânico do Rio de Janeiro (2021)
Metadados	Informação que descreve características de determinado dado, explicando-o em certo contexto de uso	IBGE (2020)
Metadados	Os metadados fornecem as informações necessárias para permitir que uma ampla comunidade científica encontre, compartilhe e compreenda o conteúdo e o contexto de um conjunto de dados	Universidade de Utrecht (2019)
Metadados	Informações estruturadas padronizadas explicando itens de dados como, mas não limitados a: finalidade, origem, referências de tempo, localização geográfica, criador, condições de acesso e termos de uso de uma coleta de dados. Documentação e explicação dos dados.	Universidade de Leiden (2021)
Metadados	Dados que fornecem informações sobre um conjunto de dados, como o(s) criador(es), o assunto dos dados, o tipo de arquivo e informações geográficas	Universidade de Twente (2018)
Metadados	Documentação e/ou informação sobre dados de pesquisa, necessária para compreender o conteúdo dos dados e o contexto em que os dados foram coletados	Universidade de Tilburg (2020)

Metadados	Informação que descreve características de determinado dado, explicando-o em certo contexto de uso.	IBGE (2020); Jardim Botânico do Rio de Janeiro (2021)
Metadados	Conjunto de dados estruturados e codificados, a partir de protocolos comuns, que descrevem e permitem acessar, gerenciar, distribuir, compreender, recuperar e/ou preservar outros dados.	Embrapa (2021)
Metadados	Conjunto de informações que acompanham e descrevem as características dos dados biológicos, ambientais, socioambientais e espaciais e as condições de sua coleta, por exemplo: local de coleta, data de coleta, nome do coletor, latitude e longitude, imagens digitais ou fotos, entre outras	PPBio (2009)
Metadados	Informações sobre os dados, como o protocolo usado para criar os dados ou o tipo de moléculas que são o foco do estudo	Wilkinson, M.D et al. (2016)
Metadados	Informações sobre um conjunto de dados de pesquisa para fins de atribuição de autoria, descrição, gerenciamento, verificação e descoberta	SciELO
Metadados	Dados sobre dados, que descrevem os contextos e o conteúdo dos arquivos de dados	National Research Council (2015)
Metadados	Dados que fornecem informações adicionais destinadas a tornar os dados científicos interpretáveis e reutilizáveis (por exemplo, data, amostra independente e construção e descrição de variáveis, metodologia, proveniência de dados, transformações de dados, qualquer intermediário ou descritivo variáveis observacionais)	Borgam e Bourne (2021)
Mineração de Dados	Uso de ferramentas de análise sofisticada para classificar, organizar, examinar e combinar grandes conjuntos de informação.	Decs - Descritores da Área da Saúde
Mineração de Dados	Processo computacional que utiliza técnicas de análise e extração de dados. Os conjuntos de dados são recuperados por softwares com sistema de busca, que reconhecem entidades, relações e ações, integrando-as de novas formas para gerar novos conhecimentos	Santos et al. (2017)
Sistemas de gerenciamento em nuvem	Ambientes virtuais disponíveis na internet, com livre acesso aos serviços abertos de armazenamento, gerenciamento, análise e reutilização dos dados, ligados às atividades de pesquisa, além das fronteiras e disciplinas científicas	Santos et al. (2017)
Sistemas de gerenciamento em nuvem	Serviço digital que permite o acesso onipresente e sob demanda a um conjunto compartilhado de recursos de computação configuráveis (por exemplo, redes de computadores, servidores, armazenamento, aplicativos e serviços), que podem ser rapidamente provisionados e liberados com esforço mínimo de gerenciamento	LCRDM (201-)
Objeto digital	A sequência de bits que realiza o objeto pode ser identificada e acessada por um identificador exclusivo e persistente ou pelo uso de atributos de referência que descrevem suas propriedades	CODATA (2022)
Pesquisador principal	Pesquisador que tem um papel de liderança em pesquisa e é o ponto de contato para um projeto ou parceria que aplica o método científico, método histórico ou outra metodologia de pesquisa para o avanço do conhecimento, resultando em resultados independentes, objetivos, resultados de alta qualidade, rastreáveis e reproduzíveis.	CODATA (2022)
Pesquisador	A pessoa que gerencia ou coordena recursos, pessoal, instalações e alocações de fundos operacionais em uma organização que realiza pesquisa, desenvolvimento e análise. Determina a natureza, os objetivos prioritários e os recursos comprometidos com sua realização dentro e entre as organizações e avalia os resultados do programa em relação aos objetivos e políticas organizacionais.	CODATA (2022)
Pesquisador Responsável	Pesquisador Responsável pelo Auxílio ou Bolsa concedido e que gerou a Propriedade Intelectual a ser protegida e licenciada.	FAPESP (2017)
Peudonimização	Tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro.	BRASIL. LGPD (2018)
Plano de gestão de dados	Um documento formal que descreve como os dados são tratados durante e após a pesquisa e quem é responsável pelo gerenciamento de dados.	Universidade de Utrecht (2019)
Plano de gestão de dados	Documento formal que descreve como os dados devem ser tratados tanto durante um projeto de pesquisa e após a conclusão do projeto. Normalmente consiste nas seguintes seções: coleta de dados, armazenamento e backup,	Universidade de Twente (2018)

	documentação, acesso, compartilhamento e reutilização e preservação e arquivamento	
Plano de gestão de dados	Descreve os dados usados ??e/ou coletados durante o projeto; onde e como esses dados de pesquisa são armazenados, gerenciados e protegidos durante o projeto; e o que acontece com os dados da pesquisa quando o projeto é concluído.	Universidade de Amsterdam (2019)
Plano de gestão de dados	Uma declaração formal descrevendo como os dados de pesquisa serão gerenciados e documentados ao longo de um projeto de pesquisa e os termos relativos ao depósito subsequente dos dados em um repositório de dados para gerenciamento e preservação a longo prazo.	CODATA (2022)
Plano de gestão de dados	Documento que especifica como os dados serão coletados ou gerados por um projeto de pesquisa, tratados, gerenciados e compartilhados durante todo o ciclo de vida da pesquisa	Santos et al. (2017)
Política para gestão de dados de pesquisa	Os processos declarados de gerenciamento de dados/informações de uma organização projetados para auxiliar e proteger os ativos de pesquisa de dados da organização. É um conjunto de princípios de alto nível que estabelecem uma estrutura orientadora para o gerenciamento de dados. Uma política de dados pode ser usada para abordar aspectos estratégicos, como acesso a dados, questões legais relevantes, questões de administração de dados e deveres de custódia, aquisição de dados e outras questões.	CODATA (2022)
Política para gestão de dados de pesquisa	Uma política institucional relativa ao compartilhamento de dados de pesquisa. Muitas vezes, é escrito como uma carta de intenção declarando que os dados da pesquisa serão submetidos a repositórios dedicados o mais rápido possível, em conformidade com os dados internacionais e formatos de troca.	LCRDM (201-)
Política para gestão de dados de pesquisa	Um documento escrito apoiado pela administração que descreve a política e fornece orientação para garantir que padrões apropriados, diretrizes consistentes e estratégias comuns sejam usadas, fornecendo vínculos e consistência com outros sistemas semelhantes e promovendo uma verdadeira rede em uma organização que produz dados.	CODATA (2022)
Política para gestão de dados de pesquisa	Documento que define as diretrizes para a coleta, armazenamento, gerenciamento, uso e preservação de dados de pesquisa. A política deve ser abrangente e abordar todos os aspectos da GDP, incluindo a identificação de dados, a preservação de dados, a segurança de dados, a acessibilidade de dados e a reutilização de dados. A política deve ser desenvolvida em consulta com todos os stakeholders envolvidos na produção e uso de dados de pesquisa, incluindo pesquisadores, administradores, bibliotecários e especialistas em TI.	Portal Foster
Políticas de Ciência Aberta	São diretrizes das melhores práticas para que a Ciência Aberta alcance seus objetivos fundamentais	Santos et al. (2017)
Preservação dos Dados de Pesquisa	Metodologias utilizadas pelos processos de backup e de arquivamento, entretanto inclui outros itens, tais como: resgate de dados, formatação de arquivos, conversão de dados e atribuição de metadados	Sayão e Sales (2015)
Preservação dos Dados de Pesquisa	Uma atividade dentro do arquivamento em que itens específicos de dados são mantidos ao longo do tempo para que ainda possam ser acessados ??e compreendidos por meio de mudanças na tecnologia.	CODATA (2022)
Preservação dos Dados de Pesquisa	Conjunto de métodos tecnológicos e gerenciais voltados para garantir que os dados permaneçam intactos, acessíveis e compreensíveis ao longo do tempo	Santos et al. (2017)
Preservação dos Dados de Pesquisa a curto prazo	Acesso a materiais digitais por um período de tempo definido enquanto o uso é previsto, mas que não se estende além do futuro previsível e/ou até que se torne inacessível devido a mudanças na tecnologia.	CODATA (2022)
Preservação dos Dados de Pesquisa a longo prazo	Um período de tempo longo o suficiente para que haja preocupação com os impactos das mudanças tecnológicas, incluindo suporte para novas mídias e formatos de dados, e de uma comunidade de usuários em mudança, sobre as informações mantidas em um repositório. Este período se estende até o futuro indefinido.	DCC - Digital Curation Centre
Preservação dos Dados de Pesquisa a longo prazo	Conjunto de métodos tecnológicos e gerenciais voltados para garantir que os dados permaneçam intactos, acessíveis e compreensíveis ao longo do tempo. O termo engloba muitas das metodologias utilizadas pelos processos de backup e de arquivamento, entretanto inclui outros itens, tais como: resgate	Sayão e Sales (2015)

	de dados, formatação de arquivos, conversão de dados e atribuição de metadados	
Princípios FAIR	Conjunto de princípios orientadores voltados para o gerenciamento e administração de dados científicos. Findable, Accessible, Interoperable e Reusable, em português: Encontrável, Acessível, Interoperável, Reusável:	SciELO
Proprietário dos dados	Refere-se a informações (ou outras propriedades) pertencentes a um indivíduo ou organização e cujo uso é restrito por esse indivíduo ou organização	CODATA (2022)
Proteção dos Dados	A proteção da confidencialidade, disponibilidade e integridade dos dados. É o termo mais amplo, pois cobre uma ampla variedade de tópicos, incluindo backup, arquivamento, preservação e proteção física, criptografia e ainda as leis que governam a proteção dos dados	LCRDM (201-)
Proteção dos Dados	Refere-se a um sistema de cópias de segurança (backup) que deve oferecer garantia contra eventual perda ou corrupção dos dados primários contidos no sistema de armazenamento por desastre, mau uso ou roubo	Sayão e Sales (2015)
Protocolo de dados	Elaboração do Regulamento de Gerenciamento de dados ao nível das faculdades e institutos de pesquisa	Universidade de Leiden (2021)
Pseudonimização ou Codificação	Ato de substituir um identificador por um código com o objetivo de evitar a identificação direta do participante/pessoa	LCRDM (201-)
Repositório de Dados de Pesquisa	Estrutura tecnológica e gerencial que permite que pesquisadores depositem seus dados de pesquisa para armazenamento e amplo acesso	Sayão e Sales (2015)
Reprodutibilidade	Obtenção de resultados consistentes utilizando os mesmos dados de entrada (input), passos computacionais, métodos, código, e condições de análise	SciELO
Reuso de dados de pesquisa	Uso dos dados para propósitos diferentes do qual eles foram coletados, geralmente por outros pesquisadores que não os autores dos dados.	Sayão e Sales (2015)
Reuso de dados de pesquisa	Uso dos dados para propósitos diferentes daquele pelo qual foram originalmente coletados. Geralmente são usados por outros pesquisadores que não os autores dos dados	Santos et al. (2017)
Riscos de segurança da informação	Potencial associado à exploração de uma ou mais vulnerabilidades de um ativo de informação ou de um conjunto de tais ativos, por parte de uma ou mais ameaças, com impacto negativo no negócio	Embrapa (2021)
Segurança da informação	Ações que objetivam viabilizar e assegurar o sigilo, integridade, autenticidade, disponibilidade e conformidade de dados e informações.	Embrapa (2021)
Segurança de dados	Termo amplo que engloba uma variedade de tópicos, incluindo backup, arquivamento, preservação e proteção física, criptografia e ainda as leis que governam a proteção dos dados.	Sayão e Sales (2015)
Segurança de rede	Dados que não são armazenados em servidores ou computadores conectados a redes externas, particularmente, em servidores que hospedagem serviços externos. Ex. Dados confidenciais, pessoais, sensíveis.	Sayão e Sales (2015)
Serviços de dados	Serviços para auxiliar as organizações na captura, armazenamento, curadoria, preservação de longo prazo, descoberta, acesso, recuperação, agregação, análise e/ou visualização de dados científicos, bem como nas estruturas legais associadas, para apoiar a pesquisa científica disciplinar e multidisciplinar.	CODATA (2022)
Stakeholders	Indivíduos, grupos ou organizações que tenham interesse ou compartilhem um empreendimento ou relacionamento e seu resultado eles podem ser afetados por ele, impactá-lo ou influenciá-lo e, de alguma forma, ser responsáveis ??por ele.	CODATA (2022)
Stakeholders	Uma pessoa ou grupo de pessoas que possuem uma participação e/ou interesses em um serviço/negócio, que não possuem vínculo administrativo e subordinação entre si.	Dicionário Oxford
Stakeholders	Processo entre uma entidade e aqueles grupos ou indivíduos potencialmente ou realmente impactados pelas ações dessa entidade em uma variedade de atividades e abordagens.	Decs - Descritores da Área da Saúde
Stakeholders	Núcleos Executores, os Núcleos Regionais, coordenadores de projetos e coordenadores de redes temáticas e todos os pesquisadores, colaboradores, estudantes, técnicos e bolsistas vinculados a esses núcleos que assinarem o termo de compromisso com esta Política	PPBio (2009)

Stakeholders internos	Uma pessoa ou grupo de pessoas vinculadas a uma instituição, que possuem uma participação em um serviço/negócio.	Dicionário Oxford
Stakeholders internos	Indivíduos, setores e lideranças da instituição envolvidas em um processo	Decs - Descritores da Área da Saúde
Tipos de dados de pesquisa	Dependem da área. São exemplos de dados de pesquisa: planilhas de estatísticas, dados retirados de bases de dados públicas ou privadas, um registro sonoro de uma entrevista, um registro descritivo de um espécime de rocha, uma coleção de imagens digitais	Santos et al. (2017)
Tipos de dados de pesquisa	considera-se que os dados variam conforme a área do conhecimento e podem estar contidos em documentos textuais, planilhas, estatísticas, cadernos de laboratório, cadernos de campo, diários, questionários, transcrições, arquivos de áudio, vídeo, fotografias, imagens, sequências de proteínas ou genéticas, artefatos, amostras, modelos, algoritmos, scripts, arquivos de log, software de simulação, metodologias e fluxos de trabalho, procedimentos operacionais, padrões e protocolos	Embrapa (2021)