



UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO (FIC)
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO (PPGCOM)

LARISSA BÁRBARA BORGES DRUMOND

**Caminhos para consolidação da Ciência Aberta no âmbito dos
periódicos científicos brasileiros: uma proposta de abertura de dados
por meio do repositório da Universidade Federal de Goiás**

GOIÂNIA

2023



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Larissa Bárbara Borges Drumond

3. Título do trabalho

CAMINHOS PARA CONSOLIDAÇÃO DA CIÊNCIA ABERTA NO ÂMBITO DOS PERIÓDICOS CIENTÍFICOS BRASILEIROS: UMA PROPOSTA DE ABERTURA DE DADOS POR MEIO DO REPOSITÓRIO DA UNIVERSIDADE FEDERAL DE GOIÁS”.

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(à) autor(a) e ao(à) orientador(a);
 - b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
- O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Laura Vilela Rodrigues Rezende, Professora do Magistério Superior**, em 01/08/2023, às 15:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Larissa Barbara Borges Drumond, Discente**, em 01/08/2023, às 16:31, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3919475** e o código CRC **A23A842A**.

LARISSA BÁRBARA BORGES DRUMOND

**Caminhos para consolidação da Ciência Aberta no âmbito dos
periódicos científicos brasileiros: uma proposta de abertura de dados
por meio do repositório da Universidade Federal de Goiás**

Dissertação apresentada ao Programa de Pós-Graduação em Comunicação da Faculdade de Informação e Comunicação da Universidade Federal de Goiás, como requisito para obtenção de título de Mestra em Comunicação.

Área de Concentração: Comunicação, Cultura e Cidadania.

Linha de Pesquisa: Mídia e Informação.

Orientadora: Professora Doutora Laura Vilela Rodrigues Rezende.

GOIÂNIA

2023

Drumond, Larissa Barbara Borges.

Caminhos para consolidação da Ciência Aberta no âmbito dos periódicos científicos brasileiros [manuscrito]: uma proposta de abertura de dados por meio do repositório da Universidade Federal de Goiás / Larissa Barbara Borges Drumond. – Goiânia, 2023.

CXXV, 125 f.: il.

Orientadora: Professora Doutora Laura Vilela Rodrigues Rezende.

Dissertação (Mestrado) - Programa de Pós Graduação em Comunicação, Faculdade de Informação e Comunicação, Universidade Federal de Goiás, Goiânia, 2023.

Inclui siglas, abreviaturas, gráfico, tabelas, lista de figuras, lista de tabelas.

1. comunicação científica. 2. periódicos científicos. 3. Ciência Aberta. 4. dados abertos. 5. repositórios de dados científicos. I. Rezende, Laura Vilela Rodrigues, orient. II. Título.

CDU 007



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE INFORMAÇÃO E COMUNICAÇÃO

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº **19/2023** da sessão de Defesa de Dissertação de **Larissa Bárbara Borges Drumond**, que confere o título de Mestra em **Comunicação**, na área de concentração em **Comunicação, Cultura e Cidadania**.

Aos **vinte e seis dias de junho de dois mil e vinte e três**, a partir das **nove horas e trinta minutos**, realizou-se a sessão pública de Defesa de Dissertação intitulada **“CAMINHOS PARA CONSOLIDAÇÃO DA CIÊNCIA ABERTA NO ÂMBITO DOS PERIÓDICOS CIENTÍFICOS BRASILEIROS: UMA PROPOSTA DE ABERTURA DE DADOS POR MEIO DO REPOSITÓRIO DA UNIVERSIDADE FEDERAL DE GOIÁS”**. Os trabalhos foram instalados pela Orientadora, Professora Doutora **Laura Vilela Rodrigues Rezende (PPGCOM/FIC/UFG)** com a participação dos demais membros da Banca Examinadora: Professora Doutora **Suely Henrique de Aquino Gomes (PPGCOM/FIC/UFG)**, avaliadora titular interna; e Professor Doutor **Fabiano Couto Corrêa da Silva (PPGCIN/FABICO/UFRGS)**, avaliador titular externo, com a participação de todos por **webconferência**. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido a candidata **aprovada** pelos seus membros. Proclamados os resultados pela Professora Doutora **Laura Vilela Rodrigues Rezende**, Presidenta da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos **vinte e seis de junho de dois mil e vinte e três**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Laura Vilela Rodrigues Rezende, Professora do Magistério Superior**, em 06/07/2023, às 10:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Suely Henrique de Aquino Gomes, Usuário Externo**, em 06/07/2023, às 12:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **FABIANO COUTO CORREA DA SILVA, Usuário Externo**, em 06/07/2023, às 14:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3873829** e o código CRC **7EB2CACC**.

Referência: Processo nº 23070.030416/2023-29

SEI nº 3873829

Dedico esse trabalho a três pessoas...

Ao Júlio (em memória), meu primeiro esposo, cuja partida me obrigou a ressignificar toda a vida que eu conhecia e a buscar por novos sonhos.

Aos meus filhos, João e Maria, que me desafiam todos os dias a ser a pessoa que quero que eles se espelhem.

AGRADECIMENTOS

À Universidade Federal de Goiás (UFG), lócus de trabalho e estudo, mas também de muitas lutas. Pela oportunidade proporcionada pelo Programa Qualificar e pela incansável busca institucional pela democratização do saber, base do desenvolvimento e transformação dos indivíduos e da sociedade.

Ao Laboratório de Estudos e Pesquisas em Educação Geográfica (LEPEG/UFG), minha lotação, minha segunda família. Pela flexibilização do atendimento às demandas que possibilitaram a realização deste estudo, sem a necessidade de afastamento e sem o comprometimento dos serviços públicos que ofertamos.

À professora Laura Vilela Rodrigues Rezende, minha inspiração de mulher cientista. Pela relação horizontal construída, pela parceria que culmina na qualidade do trabalho apresentado.

À equipe editorial do Boletim Goiano de Geografia, nas pessoas dos professores Karla Maria Silva de Faria e Ivanilton José de Oliveira, que aceitou o desafio de embarcar nessa mudança, a abertura dos dados científicos. Pela disposição e seriedade que lidaram com meu problema de pesquisa.

Ao meu noivo, Rogério Rodrigues de Moraes, que chegou à minha vida um pouco antes da qualificação, período mais desafiador desse percurso formativo, sua admiração, amor e apoio me ajudaram a persistir.

À grande amiga Bruna de Castro Cruvinel que me apresentou o Programa Qualificar e me encorajou nesse processo. Eu estava me afogando em dor e o mestrado veio como um bote salva vidas, mas eu ainda precisava nadar muito para chegar à terra seca. Bruna foi uma presença reconfortante e um apoio diário. Minha gratidão e carinho duradouros.

Aos demais amigos e familiares que apoiaram e me sustentaram emocionalmente nessa jornada. Sem citar nomes, cada um sabe cotidianamente o quanto sou grata por suas existências em minha vida. Obrigada, amo cada um!

Por fim, meu agradecimento a você que chega para ler essa obra, que ela seja útil em algum aspecto e que avancemos na temática proposta.

Não é o mais forte que sobrevive, nem o mais inteligente, mas o que melhor se adapta às mudanças.

Leon Megginson, falando sobre “A Origem das Espécies” de Charles Darwin, em 1963.

RESUMO

Este trabalho foi construído a partir de reflexões sobre a comunicação científica, especificamente, aquela praticada por meio dos periódicos científicos geridos por Instituições de Ensino Superior e com o viés da Ciência Aberta, que propõe a abertura e a democratização de todas as fases da pesquisa científica. Dito isso, o recorte do estudo é a abertura dos dados provenientes dos artigos científicos publicados nos periódicos hospedados e geridos pela Universidade Federal de Goiás. O objetivo geral foi traçar um conjunto de orientações que sirvam para a criação de coleções de dados abertos para periódicos científicos e que poderão servir de iniciativa piloto para a UFG e outras instituições, contemplando a definição de normativas de abertura de dados científicos, povoamento e uso. Trata-se de uma pesquisa social aplicada, de abordagem qualitativa. Trata-se de uma pesquisa exploratória em relação aos objetivos propostos e documental quanto aos procedimentos, abrangendo também dimensões da pesquisa-ação. Para efetivá-la inicialmente realizou-se um levantamento das características das normativas de abertura de dados científicos de periódicos de diferentes países que utilizam um repositório aberto e público, especificamente o Dataverse da Universidade de Harvard. Tais características foram parametrizadas com base nos requisitos do processo avaliativo do selo de qualidade de repositórios digitais confiáveis: Core Trust Seal. As informações obtidas neste levantamento, conjuntamente à pesquisa bibliográfica, subsidiaram a realização de etapas da pesquisa-ação junto aos editores do periódico da UFG selecionado para o estudo, o Boletim Goiano de Geografia. Nesta etapa foram formuladas as normativas de abertura de dados do referido periódico, a criação de uma coleção no repositório de dados abertos da UFG e a elaboração de peças de divulgação científica para sensibilizar os leitores e possíveis autores do Boletim sobre os temas da Ciência Aberta e abertura de dados científicos. Tais ações compõem os resultados do estudo, a partir do qual outros periódicos poderão refletir sobre os caminhos necessários para aderir ao movimento de abertura dos dados adjacentes aos artigos que publicam, ampliando as práticas científicas concernentes ao movimento pela Ciência Aberta e a democratização do conhecimento científico.

Palavras-chave: comunicação científica, periódicos científicos, Ciência Aberta, dados abertos, repositórios de dados científicos.

ABSTRACT

This work was built based on reflections from the scientific communication process, specifically, related to scientific journals managed by Higher Education Institutions and with the bias of Open Science, which proposes the opening and democratization of all phases of scientific research. The focus of the study is the opening of data from scientific articles published in journals hosted and managed by the Federal University of Goiás. The main objective was to outline a set of guidelines that serve to create open data collections for scientific journals that may serve as a pilot initiative for the UFG and other institutions, contemplating the definition of regulations for the opening of scientific data, population and use. This is applied social research with a qualitative approach. As for the objectives, the research is exploratory and according to the procedures, it is documental and also used action-research dimensions. In order to carry it out, a survey was conducted aiming to identify the characteristics of the regulations for opening scientific data from journals in different countries that use an open and public repository, the Dataverse of Harvard University. Such characteristics were parameterized based on the requirements of the Core Trust Seal trustworthy repositories quality seal. The information obtained in this survey, together with the bibliographical research, supported the implementation of action-research steps with the editors of the UFG journal selected for the study, the *Boletim Goiano de Geografia*. In this stage, the regulations for opening data from the aforementioned journal were formulated, the creation of a collection in the UFG open data repository and the preparation of scientific dissemination pieces to sensitize readers and possible authors of the journal on the themes of Open Science and opening of scientific data. Such actions make up the results of the study, from which other journals will be able to reflect on the paths necessary to join the movement of openness of data adjacent to the articles they publish, expanding scientific practices concerning the movement of Open Science and the democratization of scientific knowledge.

Keywords: scientific communication, scientific journals, Open Science, open data, scientific data repositories.

LISTA DE FIGURAS

Figura 1 – Modelo de comunicação científica: processos	38
Figura 2 – Ciclo da comunicação científica	39
Figura 3 - Taxonomia de Ciência Aberta	48
Figura 4 – Aspectos a serem considerados nas práticas de Ciência Aberta	50
Figura 5 – Plano de Gestão de Dados integrado ao ciclo de vida dos dados científicos	66
Figura 6 - The research data repository lifecycle	68
Figura 7 - Arquitetura de um repositório que atenda os princípios FAIR	69
Figura 8 – Fluxograma de envio de dados científicos pelos autores	92
Figura 9 – Coleção de dados do BGG	95
Figura 10 – Adicionando conjuntos de dados na coleção do BGG	96
Figura 11 – Carregando arquivos para a coleção do BGG	97
Figura 12 – Carregando arquivos ao conjunto de dados na coleção do BGG	97
Figura 13 – Aplicando restrições a arquivo(s) específico(s) dentro do conjunto de dados	98
Figura 14 – Selecionando uma <i>tag</i> de identificação ao(s) arquivo(s)	98
Figura 15 – Vista do conjunto de dados antes da publicação	99
Figura 16 – Alterando as licenças para acesso e uso ao conjunto de dados	99
Figura 17 – Alterando o status de curadoria dos dados antes da publicação.....	100

LISTA DE QUADROS

Quadro 1 – Experiências internacionais em prol da Ciência Aberta.....	52
Quadro 2 – Marcos na esfera do Governo Federal relacionados com a Ciência Aberta.....	53
Quadro 3 - Dados da cabeça e da cauda longa: principais diferenças.....	58
Quadro 4 – Exemplos de variações nos tipos de evidências, fontes de coleta e formatos de dados científicos.....	59
Quadro 5 – Os princípios orientadores FAIR.....	63
Quadro 6 - Proposta de template de Plano de Gestão de Dados do ICICT/Fiocruz.....	64
Quadro 7 – Algumas licenças para permissão de reuso de dados científicos em bases legais.....	70
Quadro 8 - Caracterização da pesquisa.....	72
Quadro 9 - Diferenças entre pesquisa-ação e pesquisa participante.....	75
Quadro 10 – Requisitos <i>Core Trust Seal</i>	79
Quadro 11 – Ações para implantação de uma coleção de dados de periódico científico no repositório da UFG.....	90
Quadro 12 – Metadados a serem preenchidos pela equipe editorial do BGG.....	94

LISTA DE TABELAS

Tabela 1 – Temáticas dos trabalhos com o descritor “comunicação científica” (2013 -2021) 25

Tabela 2 – Elementos observados nas políticas dos periódicos científicos selecionados 86

LISTA DE GRÁFICOS

Gráfico 1 - Panorama das teses e dissertações defendidas entre 2013 e 2021	23
Gráfico 2 - Teses e dissertações que versam sobre Ciência Aberta – 2013 a 2021.....	24
Gráfico 3 - Teses e dissertações que versam sobre abertura de dados na Ciência da Informação	27
(2013 a 2021).....	27
Gráfico 4 - Teses e dissertações que versam sobre periódicos (revistas) científicos (as)	28
na Ciência da Informação (2013 a 2021).....	28

LISTA DE ABREVIATURAS E SIGLAS

APA - *American Psychological Association*
APC - Taxas de Processamento de Artigos
BDTD - Banco de Teses e Dissertações
BGG – Boletim Goiano de Geografia
BN – Biblioteca Nacional
BOAI - *Budapest Open Access Initiative*
CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CBPF - Centro Brasileiro de Pesquisas Físicas
CCA – Centro de Ciência Aberta
Cercomp - Centro de Recursos Computacionais
Cidarq - Centro de Informação, Documentação e Arquivo
CNEN - Comissão Nacional de Energia Nuclear
CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
CONSUNI – Conselho universitário
DOI – Identificador de objeto digital
Embrapa - Empresa Brasileira de Pesquisa Agropecuária
FAIR - *Findable, accessible, interoperable and reusable*
FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo
FIC – Faculdade de Informação e Comunicação
FIOCRUZ - Fundação Oswaldo Cruz
GT – Grupo de trabalho
IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia
IES – Instituições de Ensino Superior
IESA – Instituto de estudos socioambientais
ISBN - *International Standard Book Number*
ISSN - *International Standard Serial Number*
LAI – Lei de acesso à informação
MCTIC - Ministério da Ciência, Tecnologia, Inovações e Comunicações
OJS - *Open Journal Systems*
PGD – Plano de Gestão de Dados
PKP - *Public Knowledge Project*
PRPG – Pró Reitoria de Pós-Graduação
PRPI – Pró Reitoria de Pesquisa e Inovação
RNP - Rede Nacional de Ensino e Pesquisa
SciELO - *Scientific Electronic Library*
SEER - Sistema Eletrônico de Editoração Eletrônica
SiBBR - Sistema de Informação sobre a Biodiversidade Brasileira
SIBI - Sistema Integrado de Bibliotecas
TICs – Tecnologias de Informação e Comunicação
TOP - *Transparency and Openness Promotion*
UFC - Universidade Federal do Ceará
UFG – Universidade Federal de Goiás
UFRA - Universidade Federal Rural da Amazônia
UNESCO - Organização das Nações Unidas para a Educação, a Ciência e a Cultura

SUMÁRIO

1 INTRODUÇÃO	17
1.1 JUSTIFICATIVA	21
1.2 OBJETIVOS	30
1.2.1 Geral	30
1.2.2 Específicos	30
2 FUNDAMENTAÇÃO TEÓRICO-CONCEITUAL	31
2.1 OS MEANDROS DA CIÊNCIA	31
2.2 A COMUNICAÇÃO CIENTÍFICA	34
2.3 O PERIÓDICO CIENTÍFICO	41
2.4 A CIÊNCIA ABERTA	46
2.4.1 Cenário Internacional	51
2.4.2 Cenário Brasileiro	53
2.5 DADOS CIENTÍFICOS	57
2.5.1 Gestão de dados científicos: os princípios FAIR fomentando a sua abertura em repositórios	60
2.5.2 Direitos autorais sobre dados científicos	69
3 CAMINHOS METODOLÓGICOS	72
3.1 COLETA DOS DADOS	75
3.2 ANÁLISE DOS DADOS.....	77
4 RESULTADOS	82
4.1 ANÁLISE DESCRITIVA DOS PERIÓDICOS CIENTÍFICOS QUE POSSUEM COLEÇÕES DE DADOS NO REPOSITÓRIO DATAVERSE DA UNIVERSIDADE DE HARVARD.....	83
4.2 PROPOSTA DE IMPLANTAÇÃO DA COLEÇÃO DE DADOS DO BOLETIM GOIANO DE GEOGRAFIA	88
5 CONSIDERAÇÕES FINAIS	102
6 REFERÊNCIAS	106
APÊNDICE A - Texto para página <i>web</i> do BGG, menu “Sobre”	117
APÊNDICE B – Texto de apresentação da coleção de dados do BGG	118
APÊNDICE C - Declaração de Disponibilidade de Dados a ser preenchida pelos autores	119
APÊNDICE D – Conteúdo para guia de Instruções aos Autores – Preparação e envio dos dados	120
APÊNDICE E - Conjunto de metadados para coleção do BGG	122
APÊNDICE F – Carrossel de divulgação sobre Ciência Aberta e compartilhamento de dados científicos	124

1 INTRODUÇÃO

Primeiramente é preciso destacar que este estudo pode ser considerado como um convite à mudança, não somente para os pesquisadores que publicam os resultados de suas pesquisas nos períodos científicos hospedados no Portal de Periódicos da Universidade Federal de Goiás (UFG), mas também, para os editores dessas revistas, que são os responsáveis por apoiar e consolidar as mudanças concernentes às suas políticas editoriais.

A Ciência Aberta é uma realidade, sobretudo, nos países mais desenvolvidos do mundo. Seu estabelecimento é fonte de preocupação e recomendação da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO). Discuti-la, encontrar mecanismos para que ela ocorra e criar normativas para a conscientização e efetivação de uma cultura institucional científica mais aberta, colaborativa e sustentável são os pilares aos quais um grupo de trabalho (GT) da UFG tem se dedicado.

A partir deste GT, a UFG se prepara para a criação de um Centro de Ciência Aberta (CCA) com vistas a promover um debate amplo com toda a comunidade acadêmica sobre os movimentos que compõem a Ciência Aberta e gerar um entendimento comum sobre ela, de modo que, também, se dedique à formação dos recursos humanos sobre práticas científicas de abertura, sem desconsiderar as proteções legais já previstas.

Nesse ínterim, cabe frisar que, em 2021, a UFG foi contemplada por uma chamada aberta¹ de incubação de repositório de dados científicos, juntamente com a Universidade Federal do Ceará (UFC), a Universidade Federal Rural da Amazônia (UFRA) e o Centro Brasileiro de Pesquisas Físicas (CBPF). A referida chamada, fomentada pela Rede Nacional de Ensino e Pesquisa (RNP), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), previa suporte tecnológico, técnico e de assessoria para a implantação inicial dos repositórios de dados das instituições contempladas, fornecimento de um conjunto limitado de identificadores persistentes de objetos digitais (DOIs), a partir de parceria da organização *DataCite* e CNPq, com vistas a identificar os conjuntos de dados armazenados no repositório das instituições contempladas que obtivessem êxito durante as atividades de incubação previstas na chamada. Este evento será tratado mais adiante.

Os repositórios de dados científicos (ou repositórios de dados de pesquisa) são estruturas de *software* essenciais para o delinear do presente estudo, uma vez que permitem o

1

https://www.rnp.br/arquivos/documents/Edital_incubacao_repositorios_2021_rerratifica%C3%A7ao1.pdf?iuYXJcNWPPHtD19Wpa4BItrSzwR_5GWA=

armazenamento, publicação, acesso e uso dos dados, a longo prazo, com políticas definidas de disponibilização, acesso e uso previamente definidas, além de normatizações, padronização de metadados e interoperabilidade com outros sistemas.

Quando inicialmente se menciona sobre o convite à mudança, se pretende sugerir que a utilização de repositórios de dados científicos passe a integrar a cultura universitária, já que podem ser utilizados para o armazenamento de dados oriundos das atividades do tripé universitário pesquisa, ensino e extensão. Como recorte para o estudo, tem-se sua utilização pelos periódicos científicos hospedados no Portal de Periódicos da UFG. Sabidamente, na contemporaneidade, os periódicos são os canais mais utilizados para a comunicação das descobertas científicas, principalmente, da chamada cauda longa da ciência, aquela realizada individualmente ou em pequenos grupos nas universidades e institutos de pesquisa e que, juntas acumulam um expressivo número de dados gerados em investigações científicas.

Entretanto, antes de tecer reflexões sobre comunicação científica, em especial a praticada pelos periódicos científicos, embasa-se em uma discussão anterior, iniciando com a própria definição de ciência, tomando-a nesse estudo como um processo, tal qual uma pesquisa em andamento e cujos resultados necessitam ser disseminados. Tal processo passou por modificações ao longo dos séculos, promovendo as chamadas revoluções científicas e que foram experienciadas por todos os campos disciplinares, com rapidez variadas.

Dito isso, é importante compreender que em cada um desses momentos de revolução científica, estabeleceram-se diferentes fenômenos para comunicar o conhecimento científico, que, à sua maneira, apresentaram-se como imprescindíveis para a produção de novos estudos, visto que todo processo de pesquisa científica se inicia com o embasamento de estudos realizados anteriormente.

Meadows (1999) traz que as evoluções da comunicação científica podem ser analisadas a partir do meio utilizado para efetivá-la (canais formais e informais), que foram sendo aperfeiçoados com a criação da imprensa e, posteriormente, surgimento e constantes progressos das Tecnologias de Informação e Comunicação (TICs); e, também, pelas necessidades dos próprios membros da comunidade científica em comunicar suas descobertas e assimilar a de outros, diante do crescente número de pesquisadores e informações publicadas.

Como mencionado, essa pesquisa se situa na especificidade dos periódicos científicos, como principal canal de comunicação entre os pares. Cabe esclarecer que os pares são os pesquisadores e estudantes dentro de um mesmo campo científico. Quando a comunicação é

voltada ao público não especializado, também chamado de público leigo, denomina-se divulgação científica, que não está contemplada neste estudo.

Sabe-se que, tanto os primeiros periódicos a serem criados no mundo, quanto no Brasil, apresentavam-se no formato impresso, o que tornava a comunicação formal dos resultados das pesquisas um processo lento. Com a popularização da internet, esse cenário começa a se modificar, permitindo que pesquisadores do mundo todo passem a utilizar os periódicos científicos *online* como um canal mais ágil de divulgação de suas pesquisas (Vanz; Silva Filho, 2019), ao mesmo passo que se iniciam as discussões sobre a democratização do acesso ao conhecimento científico e a utilização de práticas científicas mais abertas, em consonância com um movimento maior, o da Ciência Aberta.

Albagli (2015) assinala que o movimento pela Ciência Aberta encontra caminhos para consolidar-se, principalmente, em meio às discussões sobre a livre circulação da informação e com isso seu avanço e difusão mais ágil sem barreiras técnicas. Entretanto, a autora reflete que tal movimento é muito mais amplo do que a adoção de práticas científicas abertas, como acesso aberto às publicações científicas; ciência cidadã; utilização de redes sociais científicas; abertura de dados científicos, revisão por pares, códigos de softwares etc., ele perpassa, também, por discussões sobre apropriação do conhecimento, capitalismo e política, visto que o movimento pela Ciência Aberta abala “fontes estabelecidas de autoridade e reputação, colocando foco nas relações entre ciência e poder, e, mais amplamente entre saber e poder” (Albagli, 2015, p. 10).

Assim, quando se fala em Ciência Aberta há de se considerar essas duas perspectivas: a do movimento mundial em curso, que sugere reflexões sociais e políticas mais densas; e as das práticas científicas, que possibilitam debates em torno de um fazer científico mais transparente e colaborativo culminando na democratização do conhecimento científico. Observando num contexto integral, não é possível dissociar essas duas perspectivas, afinal órgãos governamentais, instituições científicas, econômicas e privadas e população em geral têm interesse e se beneficiam de práticas científicas abertas, sobretudo, se considerar a questão dos dados. Dados são a matéria-prima da economia mundial e das políticas internacionais; são bens muito preciosos a todos os setores da sociedade. Dados provenientes de pesquisas científicas podem ser considerados bens sociais, uma vez que agilizam descobertas científicas, tornam os estudos mais transparentes, entre outros benefícios que serão elucidados neste estudo.

Desse modo, embora cada uma das perspectivas sobre Ciência Aberta possua uma infinidade de possibilidades de serem investigadas cientificamente, o presente estudo possui

foco no segundo caso, das práticas científicas abertas, sem desconsiderar completamente as implicações da primeira perspectiva sobre elas.

Sabe-se que é possível afirmar que essas práticas foram preconizadas por ações anteriores como o Projeto Gutenberg e o movimento pelo acesso aberto. O primeiro, criado em 1971, com o objetivo de viabilizar o acesso a livros através de sua digitalização e que sobrevive até os dias atuais, com a disponibilização de milhares de *e-books*². O segundo surgiu na década de 1980 para atender à necessidade de disseminação livre das publicações científicas. O movimento pelo acesso aberto será elucidado mais detalhadamente na seção sobre os periódicos científicos, uma vez que estes, geridos por Instituições Brasileiras de Ensino Superior (IES), recorte do estudo, em sua maioria, já disponibilizam gratuitamente, para qualquer pessoa, os resultados de pesquisas científicas sem que sejam imputadas barreiras de qualquer ordem, a não ser a própria conexão com a internet.

Segundo Santos, Almeida e Henning (2017), no âmbito dos periódicos científicos que já disponibilizam acesso aos conteúdos publicados, sem a necessidade de cadastros ou pagamento de taxas, as discussões centram-se, agora, principalmente, nas políticas e infraestruturas necessárias para a abertura dos dados científicos, gerados pelas investigações cujos resultados constam em seus artigos científicos.

É nesse cenário, que esta pesquisa se propõe a estudar o tema da Ciência Aberta no contexto dos periódicos científicos brasileiros e internacionais, com foco nos repositórios de dados científicos. Geralmente, os dados gerados durante a realização das pesquisas, os quais os resultados são veiculados por meio dos periódicos científicos, acabam sendo “esquecidos” nos computadores pessoais dos pesquisadores ou das instituições. O compartilhamento desses dados faz parte do que vem a ser uma das vertentes da Ciência Aberta, com vistas a possibilitar a reprodutibilidade ou replicabilidade da ciência, a economia de recursos, a transparência e colaboração das investigações científicas.

Diante disso, identificou-se neste estudo a necessidade de compreender como os periódicos científicos brasileiros e de outros países têm desenvolvido práticas científicas correlatas à Ciência Aberta para que seja feita a proposição de um conjunto de orientações que sirvam para que periódicos científicos criem coleções de dados abertos em repositórios.

Mediante as finalidades pretendidas, este estudo classifica-se como uma pesquisa social aplicada, de abordagem qualitativa. Quanto aos procedimentos e fonte de coleta de dados, a pesquisa é exploratória, com levantamento documental, e etapas de pesquisa-ação,

² <https://www.gutenberg.org/>

junto aos editores do Boletim Goiano de Geografia (BGG), periódico selecionado no estudo para ser o piloto, ou seja, o primeiro a ter uma coleção de dados no repositório da UFG.

Nas subseções que se seguem apresentam-se a justificativa e os objetivos do estudo. Além disso, o trabalho está organizado em outras três seções: referencial teórico, abordagem metodológica e resultados. Esta última está estruturada em dois momentos, o primeiro traz o levantamento documental das especificidades das normativas de abertura de dados científicos de periódicos nacionais e internacionais que já utilizam um repositório de dados científicos, o *Dataverse* da Universidade de *Harvard*. A análise desse levantamento teve como intuito propor os aspectos necessários para a construção das normativas da coleção de dados do BGG.

No segundo momento, a partir de todo constructo das seções anteriores, apresentam-se as normativas de acesso e uso, questões de preservação e curadoria dos dados a serem armazenados na coleção de dados do BGG e que poderá servir como modelo para as demais a serem implantadas na UFG.

Por fim, trazem-se as considerações finais, com reflexões de todo estudo desenvolvido, as lacunas ou limitações e as perspectivas para novas pesquisas a partir desta. O texto foi construído na tentativa de se estabelecer um diálogo entre os principais autores aqui referenciados, à luz dos objetivos inicialmente propostos.

1.1 JUSTIFICATIVA

Essa pesquisa se justifica em pelo menos duas questões: a primeira seria sua relevância institucional, especialmente em se tratando de incremento da visibilidade e transparência e a segunda seria a relevância de sua realização para as áreas da Comunicação e Informação, sobretudo destacando a comunicação científica, temática largamente discutida nas pesquisas do campo interdisciplinar da Ciência da Informação. Esta pesquisa de mestrado está inserida em um Programa de Pós-Graduação em Comunicação, linha Mídia e Informação. Assim, pretendeu-se traçar um panorama das produções acadêmicas (teses e dissertações) que versam sobre os temas correlatos ao problema de pesquisa proposto. Não se teve o interesse de criar uma barreira entre Comunicação e Ciência da Informação, mas, pelo contrário, encontrar o que as une no sentido da comunicação científica e desenvolver um trabalho que possibilite neutralizar possíveis lacunas encontradas, como se elucidará adiante.

No que concerne à justificativa institucional, cabe mencionar que, com a importante missão de democratizar o livre acesso à produção científica, em 2004, a UFG iniciou os

esforços para estruturação de seu Portal de Periódicos, uma parceria da Pró Reitoria de Pós-Graduação (PRPG) e Biblioteca Central. No levantamento realizado em janeiro de 2023, contabiliza-se que o Portal hospeda 34 periódicos agrupados em nove áreas do conhecimento, a saber: 1 - Ciências Agrárias; 2 - Ciências Biológicas; 3 - Ciências da Saúde; 4 - Ciências Exatas; 5 - Ciências Humanas; 6 - Ciências Sociais e Aplicadas; 7 – Engenharias; 8 - Linguística, Letras e Artes; e 9 - Multidisciplinar.

Assim, o estudo se assenta no contexto do Portal de Periódicos da UFG e, também, do Programa Qualificar, cumprindo a Resolução nº 21/2015 do Conselho Universitário (Consuni) da UFG, que objetiva qualificar o corpo administrativo desta Instituição, por meio da reserva de vagas em programas de pós-graduação, que queiram integrar o programa. Por conseguinte, a autora é técnica administrativa em educação desta universidade, atuando em um periódico científico vinculado ao Instituto de Estudos Socioambientais (IESA).

Paralelamente ao interesse da administração pública na capacitação da autora, do impulso pessoal pela profissionalização sobre a temática dos periódicos científicos geridos pelas universidades, observou-se em leituras recentes que, a partir de 2014, começaram a se intensificar as discussões sobre abertura de outras vertentes do “fazer científico”, para além da disponibilização dos resultados finais das pesquisas científicas. Vislumbra-se que esse movimento de abertura e democratização de todas as fases da comunicação científica, e que culmina no conceito de Ciência Aberta, pode estabelecer-se também na UFG, sobretudo, nos periódicos científicos hospedados em seu portal.

Isso esclarecido parte-se para reflexões que justificam a relevância do estudo, a partir da qual se realizou um levantamento no Painel de informações quantitativas do Catálogo de Teses e Dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a fim de traçar um panorama de publicações acadêmicas (teses e dissertações) sobre as principais temáticas correlatas ao estudo. Tal levantamento permitiu mapear a bibliografia mais recente e relevante e que foram utilizadas para a construção da seção 2, tanto os conteúdos dos próprios trabalhos, quanto as referências teóricas apresentadas neles, além de elucidar os principais aspectos e dimensões que vêm sendo privilegiados nas pesquisas e apontar potencialidades e lacunas.

Desse modo, a partir do referido catálogo da CAPES foram realizadas duas buscas, especificamente nos links referentes às áreas do conhecimento. No primeiro momento, obteve-se a tabela com todas as teses e dissertações defendidas, e disponíveis no catálogo, na área da Comunicação e, em seguida, na área da Ciência da Informação.

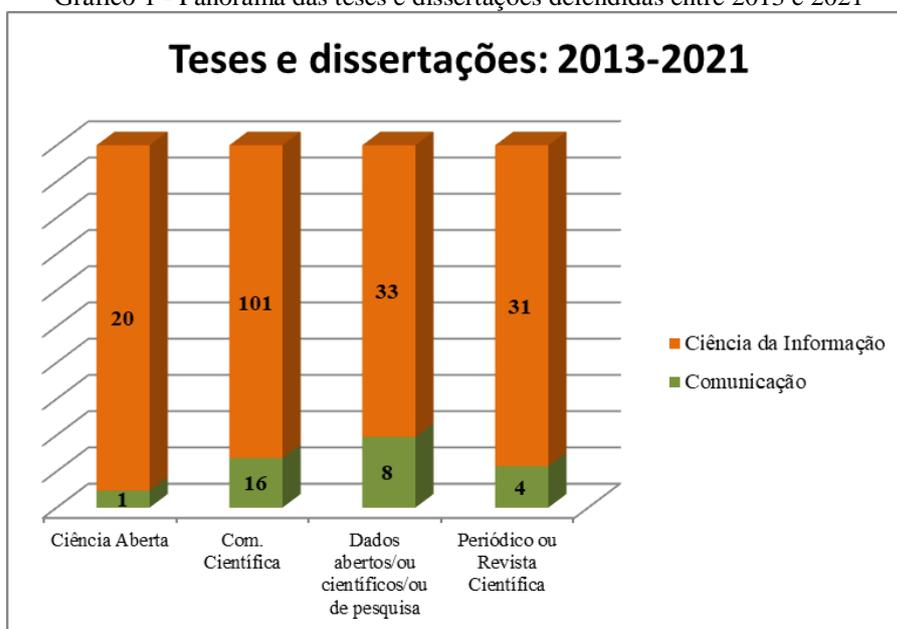
O Catálogo de Teses e Dissertações da CAPES apresenta os trabalhos defendidos desde 2013 e que tenham sido informados pelos programas de pós-graduação de todo o Brasil em seus relatórios anuais.

Optou-se pela escolha dessas duas áreas por dois motivos: primeiro porque, como já mencionado, este estudo está inserido em um programa de pós-graduação em Comunicação; segundo, porque se observou que os temas relacionados ao foco do estudo são abordados tanto na Ciência da Informação quanto na Comunicação, como será explanado seguir.

Conforme o levantamento realizado em agosto de 2022, entre os anos de 2013 e 2021 foram defendidas 2.153 dissertações e 666 teses nos programas de pós-graduação da área do conhecimento Ciência da Informação, com temáticas variadas. Em paralelo, no mesmo período, na área da Comunicação, foram defendidas 5.961 dissertações e 1.893 teses. Em ambas as áreas do conhecimento, a área de avaliação da CAPES é a Comunicação e Informação.

Nas planilhas obtidas a partir do catálogo de teses e dissertações, criou-se um esquema de análise, no qual os trabalhos eram marcados com uma cor diferente caso apresentassem no título ou nas palavras-chave um dos seguintes termos, buscados individualmente nas publicações: Ciência Aberta; comunicação científica; dados abertos (ou dados de pesquisa ou dados científicos); periódico científico (ou sua variação no plural); revista científica (ou sua variação no plural). Em quantitativo, tem-se o apresentado no Gráfico 1.

Gráfico 1 - Panorama das teses e dissertações defendidas entre 2013 e 2021

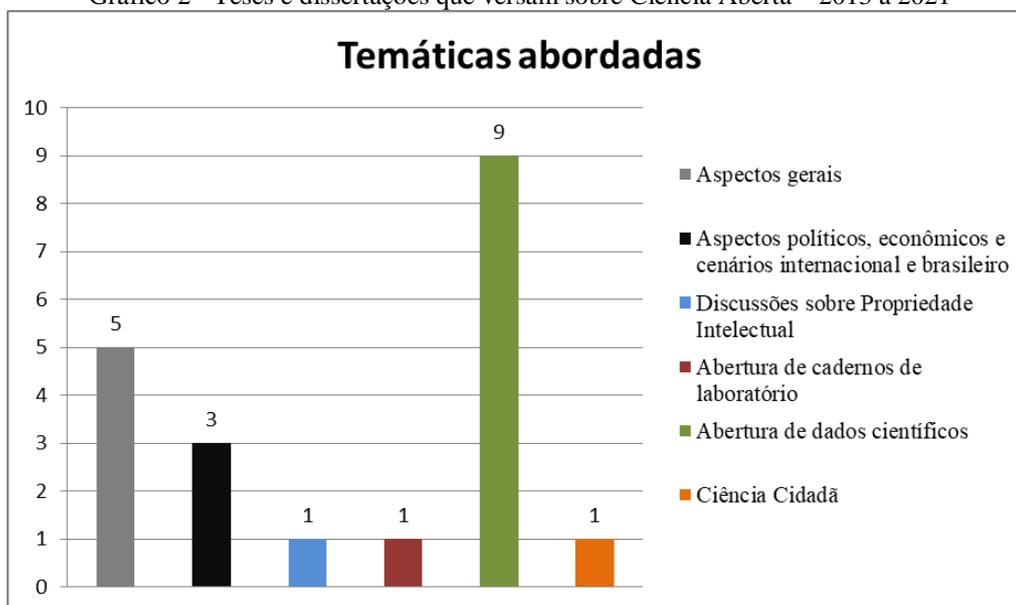


Fonte: organizado pela autora com base no Catálogo de Teses e Dissertações da CAPES (2022).

Como evidenciado no Gráfico 1, os temas correlatos à comunicação científica e Ciência Aberta são mais abordados nas produções acadêmicas provenientes dos trabalhos de conclusão dos programas de pós-graduação da área de Ciência da Informação do que na área da Comunicação. As teses e dissertações enquadradas nas quatro especificações apontadas no Gráfico 1 foram analisadas em quatro momentos do estudo.

Inicialmente partiu-se para a leitura dos resumos dos trabalhos que continham o descritor “Ciência Aberta” no título ou nas palavras-chave. O único trabalho apontado com essa especificação no campo da Comunicação, (Andrade, 2014), versa sobre aspectos políticos e de infraestrutura para o desenvolvimento da Ciência Aberta, principalmente no cenário brasileiro. Sobre os trabalhos com a mesma especificação, na área da Ciência da Informação, tem-se o apresentado no Gráfico 2.

Gráfico 2 - Teses e dissertações que versam sobre Ciência Aberta – 2013 a 2021



Fonte: organizado pela autora com base no Catálogo de Teses e Dissertações da CAPES (2022)

Com base nas principais temáticas abordadas das teses e dissertações que possuem o termo Ciência Aberta no título e/ou palavras-chave (Gráfico 2) pode-se inferir que os trabalhos elaborados debatem, sobretudo, sobre a vertente de abertura de dados em detrimento de outras, como a abertura da revisão por pares, dos códigos, dos recursos educacionais, das redes sociais científicas, entre outras concernentes à Ciência Aberta.

Além disso, mesmo que a abertura de dados tenha sido privilegiada nas produções analisadas, a concomitância da abordagem, especificamente, com periódicos científicos só se verificou no trabalho de Schiessl (2020), no qual se apresenta a relação entre os conjuntos de dados depositados em repositórios de dados e os respectivos artigos científicos publicados em periódicos científicos, num período de dez anos. É o que chamamos de bidirecionalidade

dados/publicações científicas. O trabalho se mostrou muito importante para o embasamento do estudo aqui proposto, pois também se ocupou em investigar a existência de políticas de compartilhamento de dados nas revistas científicas, alinhando-se, assim, à uma das preocupações desta pesquisa.

Outro trabalho categorizado como sendo de aspectos gerais, políticos e econômicos da Ciência Aberta e que no resumo cita os periódicos científicos é a tese de Appel (2019), que esteve centrada nas transformações dos processos de publicação dos resultados das pesquisas científicas, incluindo questões referentes aos periódicos científicos.

Num segundo momento, foi realizada a análise de conteúdo das produções que possuíam o descritor “comunicação científica”, buscado individualmente, no título ou nas palavras-chave. Cabe ressaltar que um trabalho não entrou na análise desse segundo momento, uma vez que este também possuía o descritor “Ciência Aberta”, portanto, já havia sido contabilizado e analisado no primeiro momento.

Com base na leitura dos resumos das produções, foi criada uma lista de principais temáticas abordadas. Como esse conjunto de trabalhos foi o mais expressivo (117 trabalhos), a construção dessa lista teve como subsídio o Tesouro Brasileiro de Ciência da Informação, com as adaptações que se fizeram necessárias de acordo com as especificidades desse estudo e de maneira que fosse possível analisar as produções tanto da Ciência da Informação quanto da Comunicação, em concomitância.

Desse modo, os trabalhos foram categorizados com base nas principais temáticas abordadas nas produções. Um comparativo quantitativo desses trabalhos pode ser verificado na Tabela 1.

Tabela 1 – Temáticas dos trabalhos com o descritor “comunicação científica” (2013 -2021)

Temáticas Abordadas	Ciência da Informação	Comunicação
Softwares de gestão da informação: Bibliotecas digitais e repositórios	16	2
Avaliação da produção científica por meio de estudos métricos	53	7
Comportamento informacional	3	1
Contexto histórico	4	1
Produtividade e Divulgação Científica	5	2
Gestão do conhecimento científico/periódicos	10	0
Mídias e Informação	4	1
Direito à informação e propriedade intelectual	3	0
Revisão por pares	3	0
Gênero e Ciência	0	1

Fonte: organizado pela autora com base no Catálogo de Teses e Dissertações da CAPES (2022).

Da análise da Tabela 1, é possível perceber que em ambas as áreas do conhecimento predominam os estudos métricos (análises bibliométricas ou cientométricas) de um termo, campo, periódico ou base de dados específicos.

Para este estudo, duas temáticas se mostraram mais importantes: Softwares de gestão da informação: Bibliotecas digitais e repositórios; e Gestão do conhecimento científico/periódicos.

Quanto à primeira temática mais relevante, "Software de gestão da informação", notou-se que os dois trabalhos da Comunicação e os 15 da Ciência da Informação focam na questão das bibliotecas digitais e dos repositórios institucionais. Apenas uma dissertação se debruçou sobre o objetivo de identificar as características e contribuições dos repositórios de dados científicos da América Latina, entretanto o trabalho não abordou a questão do uso dos repositórios de dados, especificamente, para o depósito dos dados científicos provenientes das pesquisas comunicadas nos periódicos.

Quanto à segunda temática mais relevante, "Gestão do conhecimento científico/periódicos", os trabalhos abordam prioritariamente as questões de políticas editoriais, gestão dos periódicos e o apoio institucional das Instituições de Ensino Superior para que essa gestão se torne mais eficiente. Um trabalho em específico tratou da gestão do conhecimento científico advindo das atividades de pesquisa, ensino e cultura de uma unidade acadêmica da Universidade Federal de Minas Gerais.

Em um terceiro momento realizou-se a análise dos resumos das teses e dissertações que continham em seus títulos ou palavras-chave as expressões "dados abertos" ou "dados de pesquisa" ou "dados científicos".

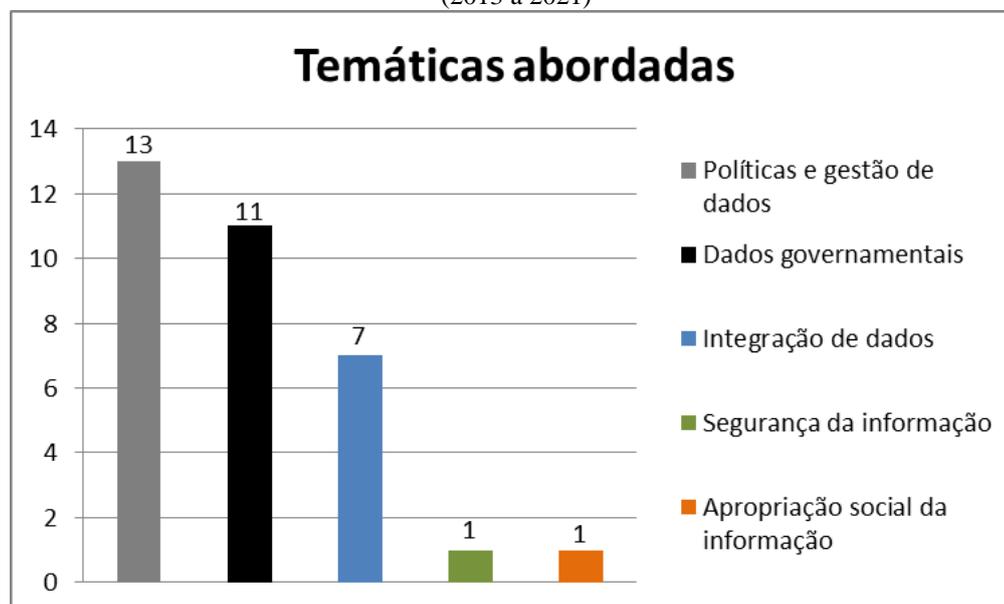
Vale frisar que para a área da Comunicação, um trabalho já havia sido categorizado, no segundo momento das análises, com as produções sobre comunicação científica. Outro trabalho aborda as mudanças sociais e políticas advindas da apropriação dos meios de comunicação, logo o uso do termo "dados abertos" nas palavras-chave, se deve apenas ao fato de que a existência desses tipos de dados na rede, é um dos fatores que favorece tais mudanças. Assim sendo, a produção não versa especificamente sobre a abertura de dados.

Restaram para a análise sobre dados (abertos ou de pesquisa ou científicos) seis trabalhos na área do conhecimento da Comunicação, dos quais dois são estudos sobre a integração de dados, especialmente a bidirecionalidade entre os dados científicos disponibilizados em repositórios de dados e as publicações científicas originadas a partir deles.

Outros três estudos abordam as questões relativas aos dados governamentais abertos: as políticas, a Lei de Acesso à Informação (LAI) e os impactos dessa abertura para a democracia. O último trabalho analisado trata da construção de notícias a partir do jornalismo de dados abertos.

Quanto aos 33 trabalhos com as mesmas especificações nos títulos ou palavras-chave na área do conhecimento da Ciência da Informação e a partir da leitura dos resumos, também se observa como uma das principais temáticas abordadas a questão dos dados governamentais; seguida das políticas e gestão de dados, as quais abordam as políticas dos repositórios, os aspectos de gestão dos dados como, por exemplo, os planos de gestão de dados e, também, versam sobre curadoria de dados digitais, onde aparece a figura do bibliotecário como protagonista. A terceira temática mais abordada nos trabalhos é a integração dos dados, com uma abordagem parecida com a verificada nos trabalhos da área da Comunicação, ou seja, a bidirecionalidade entre dados e publicações científicas, como se pode ver no Gráfico 3.

Gráfico 3 - Teses e dissertações que versam sobre abertura de dados na Ciência da Informação (2013 a 2021)



Fonte: organizado pela autora com base no Catálogo de Teses e Dissertações da CAPES (2022)

O Gráfico 3 também aponta a existência de dois outros trabalhos que trataram, respectivamente, um sobre a violação das propriedades básicas da segurança da informação com base numa análise de redes sociais; e o outro sobre a apropriação social da informação, permitida por meio da visualização de dados abertos.

No quarto e último momento de elaboração deste panorama das teses e dissertações que versam sobre as principais temáticas correlatas ao estudo, analisaram-se os trabalhos que tinham nos títulos e/ou palavras-chave os termos “periódico científico” ou “revista científica”, bem como suas variações no plural.

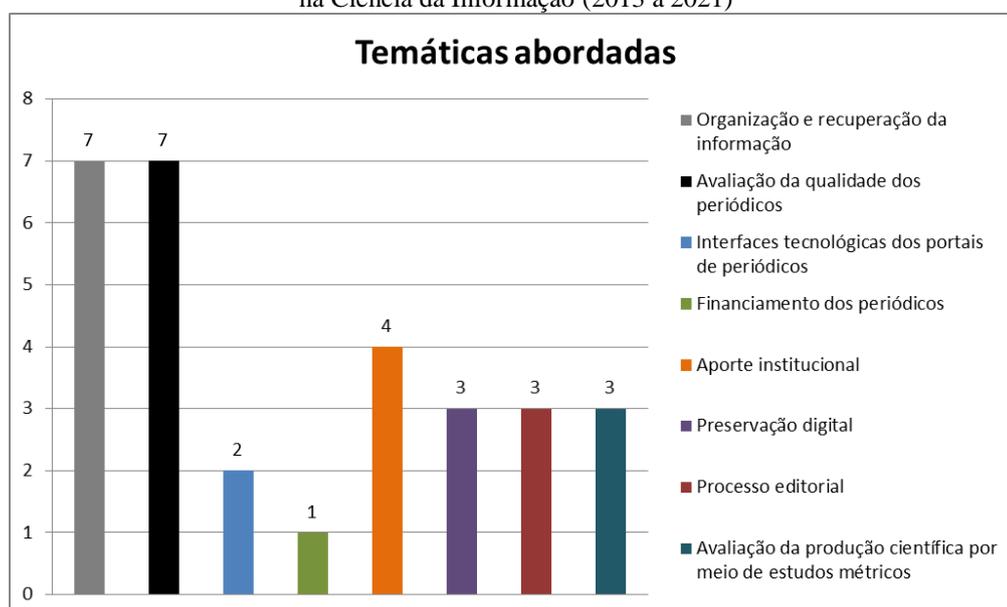
Com a leitura dos resumos dos quatro trabalhos com essa especificação na área da Comunicação observou-se que, assim como ocorreu com outro descritor analisado, três

trabalhos tratam da análise de produções científicas por meio de estudos métricos sobre um periódico específico ou grupo de periódicos com o objetivo de analisar: 1 – a classificação e qualidade enquanto periódico da área da Comunicação; 2 – a produção científica sobre um tema específico; 3 - descrever o perfil do pesquisador brasileiro que publica nos periódicos científicos da área da Comunicação.

O quarto trabalho analisado aborda o conceito de arquitetura da informação como subsídio para fortalecer a colaboração em rede de revistas científicas e para isso também analisou três revistas científicas de importância mundial.

No que se refere aos trabalhos da área da Ciência da Informação (Gráfico 4), foram analisados os resumos de 30 trabalhos que versam sobre periódicos (revistas) científicos (as). Vale ressaltar que um trabalho foi retirado da análise por não se tratar especificamente dessa temática, embora possuísse o termo “periódico científico” nas palavras-chave, tratando-se de um estudo sobre documentos fotográficos em periódicos científicos.

Gráfico 4 - Teses e dissertações que versam sobre periódicos (revistas) científicos (as) na Ciência da Informação (2013 a 2021)



Fonte: organizado pela autora com base no Catálogo de Teses e Dissertações da CAPES (2022).

Se nota, ainda no Gráfico 4 que se destacam os estudos sobre organização e recuperação da informação, principalmente sobre a importância da correta atribuição das palavras-chave nas publicações científicas e sobre os vocabulários controlados; os estudos que tratam da avaliação da qualidade dos periódicos, com foco nos critérios para que ocorra a avaliação dos periódicos, os indicadores bibliométricos, como as citações, e também a qualidade dos periódicos como um dos requisitos para avaliação dos programas de pós-graduação; e trabalhos que versam sobre o aporte institucional aos periódicos científicos,

sobretudo, tratando dos portais de periódicos científicos geridos por Instituições de Ensino Superior, notadamente as universidades federais.

Por fim, ao traçar esse panorama das teses e dissertações sobre as principais temáticas correlatas ao estudo, podem-se fazer algumas inferências:

a) Em consonância com Caribé (2015, p. 90), concebe-se que “a comunicação científica é um processo de comunicação clássico, tal como descrito, em 1949, por Shannon e Weaver (emissor, mensagem/canal e receptor)”. Sendo assim, a comunicação científica, bem como sua evolução no contexto da Ciência Aberta, é um tema com potencialidade para ser mais abordado nos programas de pós-graduação da área da Comunicação;

b) Trabalhos que se debruçam sobre a avaliação da produção científica de um termo, campo, periódico ou base de dados específicos, por meio de estudos métricos (análises bibliométricas ou cientométricas), são expressivos, em termos de quantidade, tanto na área da Comunicação, quanto na área da Ciência da Informação;

c) No geral, os trabalhos analisados, mesmo considerando que os processos de comunicação científica estão em plena revolução, sobretudo com a popularização de práticas correlatas à Ciência Aberta, não trazem uma abordagem específica sobre a abertura dos dados científicos das pesquisas publicadas pelos periódicos científicos, principalmente, os periódicos geridos por Instituições de Ensino Superior.

Assim o estudo aqui proposto se mostra relevante ao considerar essa potencialidade e se dedicar às peculiaridades inerentes aos repositórios de dados científicos provenientes das pesquisas científicas divulgadas em tais canais formais, como as políticas de acesso e uso, descrição dos dados alinhada aos estudos, questões de preservação e curadoria dos dados armazenados.

Diante de todo exposto, o estudo apresenta um panorama atualizado da Ciência Aberta no contexto dos periódicos científicos brasileiros e internacionais, e, também, possibilita que o conhecimento produzido no âmbito dos periódicos científicos da UFG seja difundido livremente, por meio da implantação de coleções de dados científicos no repositório da UFG. Além disso, poderá potencializar a internacionalização das pesquisas publicadas, ao seguir os padrões mundiais de comunicação da ciência.

1.2 OBJETIVOS

1.2.1 Geral

Traçar um conjunto de orientações voltadas para a criação de coleções de dados abertos para periódicos científicos que poderão servir de iniciativa piloto para a UFG e outras instituições, contemplando a definição de normativas de abertura de dados científicos, povoamento e uso.

1.2.2 Específicos

Delinear um panorama teórico sobre Ciência Aberta no contexto da comunicação científica, em especial, a praticada por meio dos periódicos científicos;

Diagnosticar periódicos científicos que já possuem coleções de dados em um repositório de dados, em âmbito nacional e internacional;

Levantar e apresentar as normativas de abertura de dados científicos de periódicos nacionais e internacionais com base na certificação Core Trust Seal ;

Idealizar, juntamente com a equipe editorial de um periódico da UFG, uma proposta de caracterização, normativas e implantação da sua futura coleção de dados científicos no repositório da UFG;

Implantar uma coleção de dados no repositório de dados da UFG para um periódico científico da UFG.

2 FUNDAMENTAÇÃO TEÓRICO-CONCEITUAL

Nesta seção apresentam-se reflexões de diferentes teóricos para uma compreensão ampla do objeto deste estudo, a Ciência Aberta, especificamente, a abertura dos dados científicos.

Para tanto, inicia-se com uma discussão sobre a ciência, o conhecimento científico e seus modos de comunicação. A partir desse embasamento, compreende-se que, na contemporaneidade, vive-se uma revolução científica, a era da *eScience*, ou ainda, a era da explosão dos dados e, com isso, a emergente necessidade de gerenciamento dos dados oriundos das pesquisas científicas, seja para fins de economia de recursos, ou ainda, para validar e, eventualmente, permitir a autocorreção da ciência.

Entretanto, para que isso seja possível, é necessário que os dados das pesquisas sejam abertos, um dos cerne do movimento pela Ciência Aberta. Portanto, outras discussões se fizeram necessárias para a compreensão de como pode ocorrer o compartilhamento dos dados de maneira clara, transparente e sem comprometer a autoria dos pesquisadores, como, por exemplo, uma breve revisão de literatura sobre Ciência Aberta, dados científicos (e sua abertura) o que culminou na importância em se abordar o tema dos repositórios de dados, para um compartilhamento seguro e interoperável dos dados científicos.

Desse modo, as subseções que se seguem embasam teoricamente o atendimento ao objetivo central deste estudo, que se propõe a implantar uma coleção de dados aberto para um periódico científico da Universidade Federal de Goiás (UFG).

2.1 OS MEANDROS DA CIÊNCIA

Muito se ouve, principalmente nas propagandas comerciais de alguns produtos, sobre coisas ou informações comprovadas cientificamente, o que implicitamente na mentalidade coletiva traz a compreensão de veracidade, qualidade e confiabilidade. É o que Chalmers (1993, p. 23) alega ser “uma concepção de senso comum da ciência amplamente aceita”. Mas o que vem a ser a Ciência e o que garante a um produto ou informação serem comprovados por ela? As respostas para perguntas como essas são basilares para se compreender de onde parte a comunicação científica e o motivo pelo qual ela é tão importante para disseminação das descobertas acadêmicas e para a formação dos profissionais das mais diversas áreas.

Sob a ótica de Freire-Maia (2007), a Ciência é vista sob dois aspectos: a ciência-disciplina (tal como um professor ensina aos seus estudantes e que é resultado de teorias, leis

etc. já consolidadas) e a ciência-processo (proveniente das pesquisas em andamento e que serão divulgadas posteriormente). Ambas podem ser as bases para novos estudos.

Embora diferentes autores possam conceituar Ciência de distintas maneiras, aqui cabe citar a definição dada por Freire-Maia (2007) e que se aproxima mais da ciência-processo, pretendida neste estudo e realizada pelos pesquisadores das instituições de pesquisa, como as universidades, e, posteriormente, divulgada nos periódicos científicos:

Ciência é um conjunto de descrições, interpretações, teorias, leis, modelos, etc., visando ao conhecimento de uma parcela da realidade, em contínua ampliação e renovação, que resulta da aplicação deliberada de uma metodologia especial (metodologia científica) (Freire-Maia, 2007, p. 18).

O mesmo autor conclui ainda que é o método que dá *status* científico a determinada área do saber, por meio de proposições testáveis e generalizações comprováveis. Desse modo, a utilização de uma metodologia bem definida garante confiabilidade às pesquisas fazendo com que os conhecimentos gerados, dito como científicos, sejam fundamentalmente distintos dos conhecimentos populares.

Ocorre que o modo de se fazer Ciência atualmente não é o mesmo que há dez anos e menos ainda semelhante às práticas científicas de cem anos. Kuhn (1998) faz refletir que a história da Ciência está permeada de períodos onde se praticou a “ciência normal”, ou seja, os pesquisadores direcionam suas pesquisas a partir de fundamentos estabelecidos anteriormente, e outros períodos que ele vem chamar de “revoluções científicas”. Os chamados períodos onde se pratica a “ciência normal” são caracterizados pelos tempos em que “os membros de uma comunidade científica amadurecida trabalham a partir de um único paradigma ou conjunto de paradigmas estreitamente relacionados” (Kuhn, 1998, p. 204). O autor aponta ainda que nos períodos revolucionários, os princípios fundamentais de uma disciplina são questionados, fazendo surgir outros paradigmas, nem sempre inicialmente bem aceitos, até que sejam consolidados e o processo se repita.

Esses diferentes momentos das práticas científicas durante a história podem ainda ser reconhecidos como paradigmas da Ciência. Bourdieu (2004, p. 29), também, fala sobre paradigmas em suas obras e conceitua:

O paradigma é o equivalente de uma linguagem ou de uma cultura: determina as questões que podem ser formuladas e as que são excluídas, o pensável e o impensável; sendo simultaneamente um conhecimento adquirido (*receive achievement*) e um ponto de partida, é um guia para a ação futura, um programa de investigações a empreender, mais do que um sistema de regras e normas.

Assim, de maneira simplista, concebe-se que as transformações paradigmáticas são revoluções científicas, que, ao se estabilizar via formação de consenso em torno do novo

paradigma, inaugura um novo ciclo no desenvolvimento da Ciência. Desta maneira, a Ciência entra em um estágio de normalidade: ciência normal ou amadurecida (Kuhn, 1998). Os estudiosos das obras de Kuhn, como Gattei (2008), esclarecem que Kuhn se refere às revoluções nas práticas de pesquisa de todos os campos disciplinares e que podem ser metodológicas, ontológicas, semânticas, ou mistas.

Sem o objetivo de aprofundar nos estudos de Kuhn, chega-se ao ponto que se pretende, evidenciar o entendimento de que os paradigmas são extremamente importantes para se compreender as revoluções da Ciência, mas, além disso, são fundamentais para a compreensão das mais variadas áreas do saber. Seria ingênuo pensar que para compreender a literatura ou a filosofia, como cita Bourdieu (2003), bastaria ler textos dessas áreas do conhecimento. Pelo contrário, se não houver um estudo prévio dos paradigmas epistemológicos nos quais esses campos foram construídos, não se passará da superficialidade dos entendimentos.

Alguns autores, como Hey, Tansley e Tolle (2011), numa tentativa de sintetizar de maneira mais compreensível, as revoluções específicas das práticas de se fazer ciência, e que podem ser comuns a diferentes áreas do conhecimento, trazem a noção de quatro paradigmas, sendo o último, o contemporâneo.

O primeiro paradigma remonta aos primeiros registros de pesquisas científicas, a milhares de anos, as quais eram baseadas no empirismo e na descrição dos fenômenos naturais. O segundo, apontado como presente nos últimos cem anos, tem como base as pesquisas teóricas e a utilização de modelos e generalizações. O terceiro, observado na segunda metade do século passado, tem na ciência computacional o seu principal foco, utilizando-se excessivamente de simulações. Considerando o início em meados dos anos 2000, vive-se hoje a era da *eScience*, o quarto paradigma, voltada para a computação intensiva e compartilhamento de dados (Hey; Tansley; Tolle, 2011).

Independente do momento da história da Ciência pode-se facilmente identificar o que Bourdieu chamou, em diferentes obras, de campo científico. Em suma, o campo científico se refere ao “universo no qual estão inseridos os agentes e as instituições que produzem, reproduzem ou difundem [...] a ciência” (Bourdieu, 2003, p. 20). O autor esclarece ainda que os campos científicos possuem diferentes graus de autonomia, mas apresentam-se como um mundo social muito peculiar, como um campo de forças ou de lutas pela autoridade (ou competência) científica, ou ainda, o esforço empreendido para obter aprovação e credibilidade entre os pares. Logo, a competência científica está associada ao poder simbólico (capital

científico – conhecimento e reconhecimento), ou pelo lugar que os agentes de um campo social ocupam nas ditas hierarquias instituídas.

Grosso modo, o campo científico pode ser compreendido como sendo as comunidades científicas, nas quais os agentes (cientistas e suas equipes) determinam as autoridades (as que detêm maior capital científico) e as relações que se estabelecem determinam as práticas dos cientistas no interior do campo e, por conseguinte, o nível das forças que se exercem sobre a produção científica (Bourdieu, 2004).

Poder-se-ia estender estas reflexões para as questões do campo como espaço de conflitos, sobre a distribuição desigual do capital científico e tantas outras. Entretanto, chega-se ao ponto pretendido para esse estudo, no qual os princípios gerais do campo e do capital científico permitem elucidar os fenômenos de comunicação e circulação da informação científica. Não deixando escapar, portanto, questionamentos sobre de que “lugar” o cientista fala? onde publica? e quais práticas científicas acaba por disseminar? Na próxima seção há uma tentativa de sanar tais indagações.

2.2 A COMUNICAÇÃO CIENTÍFICA

Anteriormente foi mencionado que a Ciência exige métodos testáveis e verificáveis, além da necessidade de ser analisada e aceita pelos pares para ser legitimada e finalmente comunicada (Meadows, 1999). Nesse sentido, insere-se a informação científica, empregada aqui como sinônimo de conhecimento científico, como “aquilo que se deseja comunicar”.

Antes, porém, de focar na informação classificada como “científica”, é preciso contextualizá-la sob dois campos científicos: a Ciência da Informação e a Comunicação.

Le Coadic (1996) apresenta o conceito de que informação é um conhecimento gravado sob diferentes formas: escrita, oral ou audiovisual, e que o conhecimento é um saber, fruto do ato de conhecer, por meio da simples identificação das coisas (conhecimento comum) ou da compreensão exata e completa dos objetos (conhecimento científico).

Essa noção faz lembrar outro autor, Goulemot (2011), e sua conhecida expressão “biblioteca das memórias”. Fazendo uma analogia com o expresso por tal autor que falava especificamente sobre leitura, o que se pretende dizer é que, para que uma informação seja enquadrada com um conhecimento, não se podem deixar de lado outros conhecimentos que o sujeito já absorveu. Assim, expande-se para a noção “de que todo novo conhecimento advém de conhecimento prévio e é, portanto, uma produção social” (Albagli, 2015, p. 11). Da mesma

forma, o conhecimento científico é inicialmente embasado nas informações anteriores de outros cientistas.

Com o intuito de situar ainda mais a conceituação de informação, Le Coadic (1996) elucida que:

A informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal: impresso, sinal elétrico, onda sonora etc. Essa inscrição é feita graças a um sistema de signos (a linguagem), signo este que é um elemento da linguagem que associa um significante a um significado: signo alfabético, palavra, sinal de pontuação (Le Coadic, 1996, p. 5).

Por si só, essa definição já traz um elemento importante de todo processo comunicacional, a significação envolvida nos processos de troca de informações. Então, embora a Ciência da Informação estude e trabalhe a informação como processo e construção, preocupando-se, sobremaneira, com o conteúdo e com o melhor canal para se ofertar a informação ao receptor e a Comunicação interprete a informação como um processo de mensagens entre um emissor e um receptor (Bazi, 2007), a partir de um contexto e com intencionalidades e significações muito subjetivas para se mensurar, não é possível dissociar o termo informação, tomando por base os dois campos científicos, para se trabalhar, especificamente, a comunicação científica.

Entretanto, se pode aclarar que a tipificação “científica” qualifica a comunicação de um modo peculiar, fazendo emergir as necessidades de se considerar que essa qualificação ocorre num contexto que é próprio dos campos científicos, possui objetivos próprios e deve se preocupar, sobremaneira, com o público-alvo (receptor) dessas informações (Bufrem, 2019).

Desse modo, a informação científica é aquela passível de ser comunicada por sistemas estruturados ou por meios formais de comunicação, sendo caracterizada por diferentes tipos de literatura científica, ou manuscritos científicos (Vanz; Silva Filho, 2019,). Esses manuscritos apresentam-se de variadas maneiras, como artigos originais, teses, dissertações, trabalhos para eventos, entre outros.

Logo, concebe-se que:

[...] comunicar é um processo vital para a ciência, de transferir conhecimento e estimular o fluxo de ideias entre uma fonte geradora (cientista) e um receptor (outros cientistas ou sujeitos que façam uso de informação científica. A informação científica veiculada em um produto informacional assume a condição de fonte de informação que poderá estimular futuras pesquisas e, possivelmente, o processamento de novas informações entre os pesquisadores (Vanz; Silva Filho, 2019, p. 23).

Assim como a história das revoluções científicas, segundo Kuhn (1998), apresentaram tendências metodológicas, ontológicas, semânticas, ou mistas que foram naturalmente

evoluindo, houve, também, alterações nas maneiras de se comunicar as descobertas científicas. A expressão “comunicação científica” aparece pela primeira vez na literatura, em 1939, na obra *The Social Function of Science* do cientista britânico John Bernal (Caribé, 2011).

Meadows argumenta que há pelo menos dois caminhos pelos quais se podem analisar as evoluções da comunicação científica: “a natureza do meio empregado para transmitir informações e as necessidades dos membros da comunidade científica” (Meadows, 1999, p. 1).

Quanto aos meios utilizados pela comunicação científica, em seus primórdios, utilizava-se basicamente a fala e a escrita, ou seja, a comunicação formal (como cartas entre pesquisadores, até então manuscritas) e a comunicação informal (conferências, por exemplo).

O surgimento da imprensa possibilitou o crescimento das publicações circulantes e representou uma grande revolução na comunicação e com isso afetou também o comunicar das informações científicas (Gomes; Santos; Reis, 2020). “A capacidade de multiplicar os exemplares de um livro representou um passo importante rumo a uma difusão melhor e mais rápida das pesquisas” (Meadows, 1999, p. 4). Mas além da impressão de livros, os pesquisadores vislumbraram a oportunidade de imprimir suas cartas, ao invés de manuscritas, com o intuito de alcançar um número maior de pares, sendo esse o início histórico das revistas ou periódicos científicos.

Após a imprensa, o surgimento do computador e, posteriormente, da internet representou um salto nas possibilidades de disseminação das descobertas científicas. Somado a isso, o grande volume de dados gerados cotidianamente e a necessidade de agilidade e transparência, tanto na execução das pesquisas, quanto na comunicação de seus resultados. Nesse cenário, surgiram outros dois meios para se comunicar as informações científicas: os chamados canais semiformais e os super formais.

Nos canais semiformais, a comunicação permite o compartilhamento ainda no processo de produção, ou seja, antes que os manuscritos científicos sejam avaliados pelos pares ou formalmente publicados em um periódico científico, por exemplo, os repositórios de *preprints* (Bufrem, 2019).

Já os canais super formais são aqueles em que a literatura científica já existe, cabendo a eles indicar sua existência e facilitar sua identificação, localização e recuperação para leitura e estudos futuros, como as bases de dados do Portal da Capes ou do Portal SciELO (Bufrem, 2019).

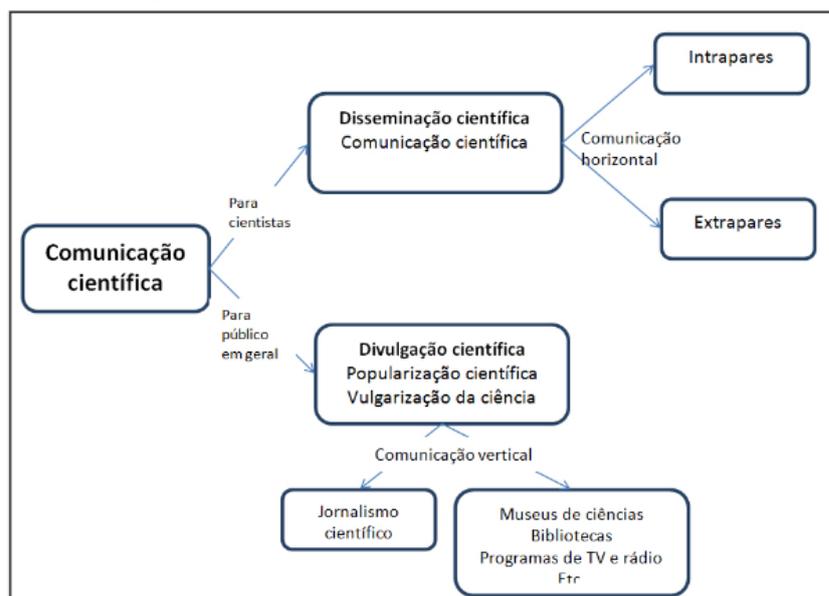
Retomando ao segundo caminho pelo qual Meadows (1999) analisa as transformações da comunicação científica, chega-se ao passo de refletir sobre as necessidades das comunidades científicas, visto que elas também passaram por transformações, como seu crescimento acelerado, sobretudo porque os níveis de escolaridade têm se elevado significativamente, por exemplo, com o aumento do número de pessoas que cursam pós-graduação *stricto sensu*. Logo, têm-se mais pessoas pesquisando, mais resultados de pesquisas sendo comunicados e surge a dificuldade para que os pesquisadores assimilem esse volume crescente de informações.

O caminho encontrado pelas comunidades científicas tem sido as especializações, ou seja, à medida que a Ciência se expande, os pesquisadores focam sua atenção e esforços intelectuais em temáticas selecionadas, desabrochando a figura das sociedades ou associações científicas e dos periódicos também especializados, para atender esse perfil de pesquisador, (Meadows, 1999), mas não deixando de existir os títulos que abarcam produções gerais ou interdisciplinares.

Ao se considerar o conceito apresentado por Bufrem (2019, p. 15) de que “a comunicação científica abrange um ciclo de atividades que incluem a produção, a disseminação e o uso da informação científica”, o pesquisador aparece como protagonista e transita cotidianamente entre práticas e etapas de produção efetiva da pesquisa e a maneira como deve relatar (disseminar) o que fez, de modo que tais resultados sejam aceitos como conhecimento científico, embasando outros estudos, ou ainda encontrem uso na sociedade.

Em suma, um pesquisador (autor) espera que os resultados de seus estudos sejam acessados pelos pares, validado e citado. A partir dessa assertiva é preciso esclarecer que há uma distinção conceitual e de práticas entre comunicação e divulgação científica, visualmente explicitada na Figura 1.

Figura 1 – Modelo de comunicação científica: processos



Fonte: Caribé (2015, p. 101).

Para o estudo aqui pretendido, focou-se especificamente na comunicação científica para cientistas (disseminação científica) intrapares (entre cientistas que estão no mesmo campo de especialização) e extra pares (para outros pesquisadores de campos científicos mais gerais ou interdisciplinares).

Considerando que no decorrer das revoluções científicas e, principalmente, com o desenvolvimento das TICs, os processos de comunicação científica fizeram emergir diferentes canais de informação e comunicação, apresenta-se a Figura 2 na qual o ciclo da comunicação científica pode ser vislumbrado em grande parte de suas possibilidades de efetivação.

Figura 2 – Ciclo da comunicação científica



Fonte: elaborado pela autora com base em Bufrem (2019).

Concebe-se que todo ciclo da comunicação científica origina-se na figura do pesquisador: no contexto das Instituições de Ensino Superior (IES) brasileiras, as universidades, têm-se os docentes que são os pesquisadores que desenvolvem suas pesquisas individualmente ou coletivamente e, no âmbito dessas pesquisas, há também os pesquisadores em formação, que são os discentes que participam de iniciativas de iniciação científica. Têm-se ainda os pós-graduandos que desenvolvem suas pesquisas com a orientação dos docentes. Há, ainda, os pesquisadores independentes e os pesquisadores que estão ligados a instituições ou centros de pesquisa.

Normalmente uma pesquisa se origina do conhecimento prévio, do interesse por temáticas específicas por parte do pesquisador ou de demandas encaminhadas pelas instituições.

A partir de leituras, observações, lacunas encontradas e questionamentos sob diversas perspectivas, o pesquisador chega a uma delimitação de tema, no qual estabelece limites concretos para a pesquisa e que são os recortes de recursos e marco temporal. Com base nessa delimitação, o pesquisador planeja a pesquisa, define etapas e métodos e, posteriormente, a executa. Durante a fase de produção, ou seja, antes que os resultados sejam formalmente

concluídos, o pesquisador pode submetê-los, por exemplo, em repositórios de *preprints* (canais semi formais) para que receba contribuições de outros pesquisadores.

Com a redação do relatório final da pesquisa tem-se, então, um novo conhecimento científico, como os artigos originais, que necessitam ser avaliados pelos pares para ser legitimado. Então, esse manuscrito pode ser comunicado, por exemplo, intrapares ou extra pares em eventos científicos (canal informal) ou submetido para publicação em um periódico científico (canal formal).

Os artigos e periódicos, por sua vez, podem ser indexados em bases de dados ou bibliotecas virtuais e os dados científicos originados ou obtidos nas pesquisas, disponibilizados em repositórios de dados. Os manuscritos são indexados com o objetivo de localizar, recuperar e divulgar o conhecimento científico.

Todo esse ciclo comunicacional pode ser gerido pelas bibliotecas com o uso de sistemas de informação (*softwares*) criados para essa finalidade. Este é o cenário mais abrangente e que compreende as etapas apresentadas na Figura 2. Obviamente nem todas as pesquisas científicas atingirão todos os meios (canais) passíveis de serem utilizados pela comunicação científica.

Gomes, Santos e Reis (2020, p. 103) associam o ciclo da comunicação científica ao ciclo documentário o qual “compreende uma série de procedimentos que engloba sete etapas: produção ou geração do conhecimento, seleção e aquisição, representação, armazenamento, recuperação, disseminação e uso da informação”. Essa caracterização assemelha-se ao apresentado na Figura 2 e, além disso, estes autores apresentam que as publicações derivadas nessas etapas, podem ser classificadas como:

a) Fontes Primárias: “apresentam informação segura e completa sobre determinado assunto e que possibilitam um maior aprofundamento” (Gomes; Santos; Reis, 2020, p. 104), como os livros, os artigos de periódicos, as normas técnicas, entre outros;

b) Fontes Secundárias: “visam facilitar o uso e a consulta de determinada informação” (Gomes; Santos; Reis, 2020, p. 105), como as revisões de literatura, os dicionários, as enciclopédias etc.

c) Fontes Terciárias: “remetem e guiam o usuário para as fontes primárias e secundárias” (Gomes; Santos; Reis, 2020, p. 105), como as bibliografias, os diretórios e as bases de dados.

Assim como outras investigações no campo da comunicação científica, o recorte do estudo aqui apresentado é feito a partir da utilização dos periódicos científicos, em especial os

geridos pelas IES. Por esse motivo e com a intenção de elucidar sua importância, a próxima seção apresentará suas peculiaridades, formatos, funções, desde seu surgimento.

2.3 O PERIÓDICO CIENTÍFICO

Como brevemente mencionado anteriormente sobre as evoluções paradigmáticas de se fazer Ciência, até o século XVI, as práticas científicas eram, em suma, baseadas em argumentação e dedução. No século XVII, num movimento que se iniciou na Europa, a comunidade científica começou a exigir evidências baseadas na observação e na experiência empírica para aceitar um conhecimento como sendo científico. Com isso, novas formas de comunicar os estudos também surgiram. Se antes, os pesquisadores utilizavam encontros pessoais e cartas para discutirem suas pesquisas, a partir do século XVII, a divulgação formal se tornou predominante e iniciou-se a utilização de livros e longos tratados para discorrer acerca do conhecimento acumulado sobre determinado assunto, em determinada localidade (Mueller, 2000b).

Conceitualmente, um periódico científico é caracterizado por manuscritos científicos editados e encadeados sequencial e cronologicamente em números ou fascículos, independente de seu suporte, publicados, preferencialmente, em intervalos regulares, por tempo indeterminado e atendendo as normalizações básicas de controle bibliográfico. Um periódico normalmente é dirigido por um corpo editorial, com a figura de um editor chefe e pode ser de assuntos gerais, interdisciplinares ou especializados (Fachin; Hillesheim, 2006).

Mueller (2000b) relata que um dos primeiros periódicos científicos que se tem notícia é o parisiense *Journal de Sçavans*, sucedido por outros na Europa e, com o passar dos anos, se expandindo para todo o mundo com a publicação crescente de periódicos impressos, seguidos da disponibilização eletrônica, primeiro em formato de CD *Room* e depois, como é mais comum na atualidade, por meio da internet. Já no Brasil, os primeiros periódicos científicos que se tem registro são a “Gazeta Médica do Rio de Janeiro” e a “Gazeta Médica da Bahia” (Yamamoto *et al.*, 2002).

Desde seu surgimento, e até os formatos atuais, os periódicos científicos cumprem algumas funções específicas: como comunicação formal dos resultados das pesquisas aos seus pares e demais interessados; preservação das ideias e reflexões dos cientistas ao longo do tempo em forma de conhecimento registrado; garantia de autoria de uma descoberta ou tese científica (propriedade intelectual); estabelecimento de confiabilidade e qualidade da Ciência, possibilitadas, sobretudo, pela avaliação por pares (Mueller, 2000b).

Desde a popularização dos periódicos no meio científico, ele se tornou o principal canal de comunicação da ciência, uma vez que permite que os pesquisadores saibam o que vem sendo publicado sobre um determinado assunto que está sendo pesquisado e também pela necessidade de apresentar seus resultados para o crivo de seus colegas (pares) (Vanz; Silva Filho, 2019).

Porém, a morosidade do processo de publicação de pesquisas em periódicos científicos impressos caracterizou, por muitos anos, a lentidão da divulgação de pesquisas nacionais e internacionais. Transcorrem-se, em média, dois anos, entre a submissão e a publicação final do artigo (Bomfá; Castro, 2004). Estes autores citam que ainda no início do século XXI, os principais problemas destes tipos de publicações eram: demora da análise do artigo pelo corpo editorial e comitê científico; custo da produção impressa; alto custo para os leitores adquirirem uma coleção completa; e dificuldade de distribuição dos exemplares a todas as bibliotecas das universidades.

Todos esses fatores, somados à popularização da internet alavancaram a criação de periódicos eletrônicos, primeiramente dos que já eram disseminados no formato impresso, que passaram a coexistir por um tempo, e, posteriormente, o surgimento de periódicos que já nasceram em formatos digitais.

Estes primeiros começam a surgir em 1990, sendo disponibilizados em arquivos *American Standard Code for Information Interchange (ASCII)*, o acesso era caro, uma vez que era centrado nas instituições acadêmicas, sendo necessários *hardwares* e *softwares* específicos para lê-los (Fachin; Hillesheim, 2006). Ainda nessa década, os avanços tecnológicos começaram a se desenvolver rapidamente, com isso os arquivos eletrônicos puderam ser disponibilizados completos e *online*, a depender das políticas de cada periódico.

Vanz e Silva Filho (2019) defendem que os formatos eletrônicos permitiram novas possibilidades de divulgação e distribuição dos periódicos científicos, a saber: agilidade na publicação, maior alcance e diversidade de público, interatividade, navegação hipertextual, redução de custos e facilidade de recuperação por meio da internet. Além de que a evolução dos arquivos ASCII para os arquivos *online* abriram portas para o uso de multimídias nos artigos.

Segundo Castro (2006, p. 63), a publicação eletrônica de revistas científicas na internet permite:

[...] que os artigos estejam disponíveis imediatamente após a aprovação pelos editores. Essa modalidade de publicação contribui para aumentar a visibilidade dos resultados de pesquisa e diminuir o tempo entre a aprovação dos trabalhos e sua publicação em formato impresso. O artigo científico passa a ser uma unidade informacional independente, embora reunido posteriormente em fascículos,

enquanto permanecerem vigentes os princípios tradicionais. Informações complementares e versões em outros idiomas podem ser incluídas nos sites das revistas, constituindo valor agregado à forma impressa. [...] A publicação em papel e a organização em fascículos no novo fluxo de comunicação científica passam a ser subprodutos do formato eletrônico.

Com o aumento exponencial do número de periódicos científicos *online*, muitas preocupações também surgiram, como o controle bibliográfico e as questões ligadas à preservação dos registros efetuados em ambiente virtual. No caso dos volumes físicos (impressos, CD-Room, fitas, partituras musicais etc.), a Biblioteca Nacional (BN) deveria receber anualmente o depósito legal da produção intelectual brasileira, entretanto a falta de instrumentos legais e de pessoal faz com que esses recebimentos estejam muito aquém do que é produzido no país (Fachin; Hillesheim, 2006).

Assim como existe o *International Standard Book Number* (ISBN), um sistema internacional que identifica e individualiza os livros por seu país, editora, título e edição, foi criado o *International Standard Serial Number* (ISSN), um identificador internacionalmente aceito para individualizar as publicações seriadas, como os periódicos científicos, tanto físicos, quando eletrônicos.

No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) atua como Centro Brasileiro de ISSN por meio de acordo firmado entre o Centro Internacional do ISSN e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Instituto, 2021).

No sentido do controle das publicações em meio *online*, outro aspecto importante foi a criação do Identificador de Objeto Digital (DOI), uma sequência numérica aceita internacionalmente para identificar conteúdos digitais (documentos, imagens, artigos, *e-books* etc.). O objetivo principal é localizar e acessar mais facilmente conteúdos na internet (Fachin; Hillesheim, 2006).

Se por um lado, têm-se vários órgãos preocupados com essas normalizações e meios para facilitar o acesso aos arquivos disponíveis *online*, por outro, ainda se mantém outras preocupações como, por exemplo, organizar, tratar e preservar as informações científicas comunicadas por meio deles. É nesse contexto que começam a surgir os primeiros portais de periódicos científicos das IES que evoluíram, sobremaneira, com a adoção do *Open Journal System* (OJS), um software *open source* (código aberto) desenvolvido pelo *Public Knowledge Project* (PKP) com o objetivo de gerenciar e publicar periódicos acadêmicos.

O software chegou ao Brasil por meio do IBICT, no ano de 2003, que o traduziu e o personalizou a partir da realidade brasileira, chamando-o de Sistema Eletrônico de Editoração

Eletrônica (SEER). O IBICT, então, realizou testes de implementação com o periódico Ciência da Informação e, com o êxito, passou a divulgar e distribuir o SEER a editores brasileiros interessados em publicar periódicos científicos de acesso aberto por meio da internet. Atualmente o software voltou a ser reconhecido apenas pela sigla OJS, a sigla de seu nome original (Instituto, 2021).

A existência de portais de periódicos nas IES garante que os periódicos de uma mesma instituição estejam organizados de modo a registrar a identidade da produção científica a ela associada e que sejam garantidos princípios de padronização, preservação de dados, tipos de arquivos suportados e segurança (Garrido; Rodrigues, 2010).

Uma característica importante desses portais, e que aparece ao se mencionar a distribuição do OJS, é o acesso aberto. Oliveira (2020) esclarece que o movimento pelo acesso aberto foi uma resposta à crise dos periódicos científicos da década de 1980, devido ao alto preço das assinaturas dos periódicos cobrado pelas editoras da época. Apareceu, assim, a demanda por novos meios de difusão das informações científicas, principalmente “porque diante do corte de financiamentos, a atualização das coleções das bibliotecas ficou comprometida” (Oliveira, 2020, p. 125). Emergindo a necessidade de se acessar e manipular livremente a informação científica.

Desse modo, o movimento pelo acesso aberto prevê a disseminação sem o pagamento de taxas econômicas pelos leitores, e sem quaisquer outros dificultadores para que se acesse a informação científica (Oliveira, 2020).

A ideia do acesso aberto foi um dos frutos do evento *Budapest Open Access Initiative* (BOAI), da qual o Brasil é signatário, realizado em Budapeste em 2002, que, inicialmente, pretendia reunir projetos já existentes em torno desse objetivo comum. Do evento, surgiu a definição que, mesmo nos dias atuais, vinte anos depois do surgimento, ainda é reafirmada pelo movimento:

“Acesso aberto” à literatura científica revisada por pares significa a disponibilidade livre na Internet, permitindo a qualquer usuário ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral desses artigos, recolhê-los para indexação, introduzi-los como dados em software, ou usá-los para outro qualquer fim legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis ao próprio acesso a uma conexão à Internet. As únicas restrições de reprodução ou distribuição e o único papel para o direito autoral neste domínio é dar aos autores o controle sobre a integridade do seu trabalho e o direito de ser devidamente reconhecido e citado (BOAI, 2022).

A declaração do BOAI foi seguida por outras, de diferentes países, que abraçaram a causa do acesso aberto e se comprometeram a criar políticas para viabilizá-lo, como a *Bethesda Statement on Open Access Publishing* (2003); e a *Berlin Declaration on Open*

Access to Knowledge in the Sciences and Humanities (2003). No Brasil, desde 2005, o IBICT vem empenhando-se na execução de muitas ações concretas, como a customização do OJS, citada anteriormente, a criação do Banco de Teses e Dissertações (BDTD), dentre outras (Oliveira, 2020).

Inicialmente, o movimento pelo acesso aberto propunha duas estratégias para alcançar seus objetivos: a via verde e a via dourada (Kuramoto, 2007). Entretanto, com a evolução das práticas de acesso, suas possibilidades vêm crescendo, sobretudo, devido aos modelos econômicos em que operam. Atualmente os modelos se caracterizam em, pelo menos quatro, como apresenta Mendes (2019):

a) Via verde: arquivamento da produção científica em repositórios;

b) Via dourada: o conteúdo é publicado em um periódico de acesso aberto e todos os artigos podem ser acessados livremente pelos leitores diretamente no site do periódico, entretanto podem ser cobradas taxas dos autores, as chamadas taxas de processamento dos artigos (APC);

c) Via diamante: a produção é publicada em periódicos científicos de acesso aberto e não são cobradas taxas nem dos autores e nem dos leitores. Normalmente, estes periódicos são geridos por instituições de ensino e pesquisa;

d) Via híbrida: “o conteúdo é publicado em um periódico de assinatura, mas possuem a opção do autor publicar seu artigo em acesso aberto por meio do pagamento de APC” (Mendes, 2019, p. 62).

Os periódicos geridos pelas IES brasileiras, notadamente as universidades públicas, comumente, já praticam o acesso aberto pela via diamante, para suas publicações científicas. Cabe demarcar que, embora muitos pesquisadores utilizem os termos acesso aberto e acesso livre como sinônimos há diferenças conceituais entre eles. Baptista *et al.* (2007, p. 2) trazem que o acesso livre prevê “disseminação ampla e irrestrita dos resultados de pesquisas financiadas com recursos públicos”. Portanto, alguns condicionantes das vias douradas e híbridas do acesso aberto, inviabilizam o real acesso livre às produções.

Recentemente, emergem outras preocupações, as políticas e infraestruturas necessárias para a abertura dos dados científicos que foram gerados ou coletados durante os estudos que resultaram em tais publicações (Santos; Almeida; Henning, 2017). Nesse particular, quando se trata da abertura destes dados, foco do estudo, há diferentes licenças aplicáveis ao acesso, uso e reutilização de tais dados, por isso optou-se aqui pelo uso predominante da expressão acesso aberto.

Para tecer reflexões sobre a abertura dos dados científicos é preciso, antes, elucidá-la à partir de um movimento maior e mais abrangente: a Ciência Aberta, um tema que será discutido, com o recorte aqui pretendido, na próxima seção.

2.4 A CIÊNCIA ABERTA

Para situar a Ciência Aberta no contexto da comunicação científica é preciso entender que, para além de definições, existem articulações entre governos, comunidades científicas e financiadores das pesquisas que promovem embates sobre a socialização do conhecimento e sua privatização, ou ainda, incomoda as estruturas hierárquicas dos campos científicos e faz repensar as práticas produtivistas, até então dominantes (Albagli, 2015). Assim, é preciso esclarecer que a Ciência Aberta pode ser vislumbrada a partir da ótica de um movimento amplo que sugere discussões sociais, políticas, econômicas etc. e sob o viés de práticas científicas transparentes e colaborativas, que culminam na democratização do conhecimento científico.

No que concerne à comunicação científica, embora o estudo aqui realizado esteja centrado mais nas questões das práticas científicas abertas, Albagli (2015) elucida alguns desafios, que se associam, também, às questões mais amplas do movimento pela Ciência Aberta. Tais desafios, segundo a autora, são caracterizados pelo interesse pela propriedade intelectual, pela dependência das editoras privadas, que efetivavam as publicações científicas e, ainda, pelas limitações ao acesso e ao uso de materiais digitais, promovidas por diferentes licenças. Principalmente por esses motivos, é que é possível afirmar que as primeiras iniciativas do movimento da Ciência Aberta estavam centradas no fortalecimento do acesso aberto às publicações científicas, um movimento anterior às discussões da Ciência Aberta e que foi descrito na seção anterior.

Nesse contexto, a Ciência Aberta emerge em contraposição à mercantilização e privatização do conhecimento científico e, com isso, atender às necessidades de novos modelos de publicação, nos quais se obtenha, além dos resultados finais das pesquisas científicas, por meio de relatórios manuscritos, como os artigos originais, também se possam acessar os modelos que evidenciem como a pesquisa foi desenvolvida em sua totalidade, seus métodos, suas anotações, os dados obtidos nos levantamentos, dentre outros, a fim de que essas informações “possam servir de base para avaliações mais confiáveis e para testes de veracidade e replicação de experimentos, estarem disponíveis para serem reutilizadas mais

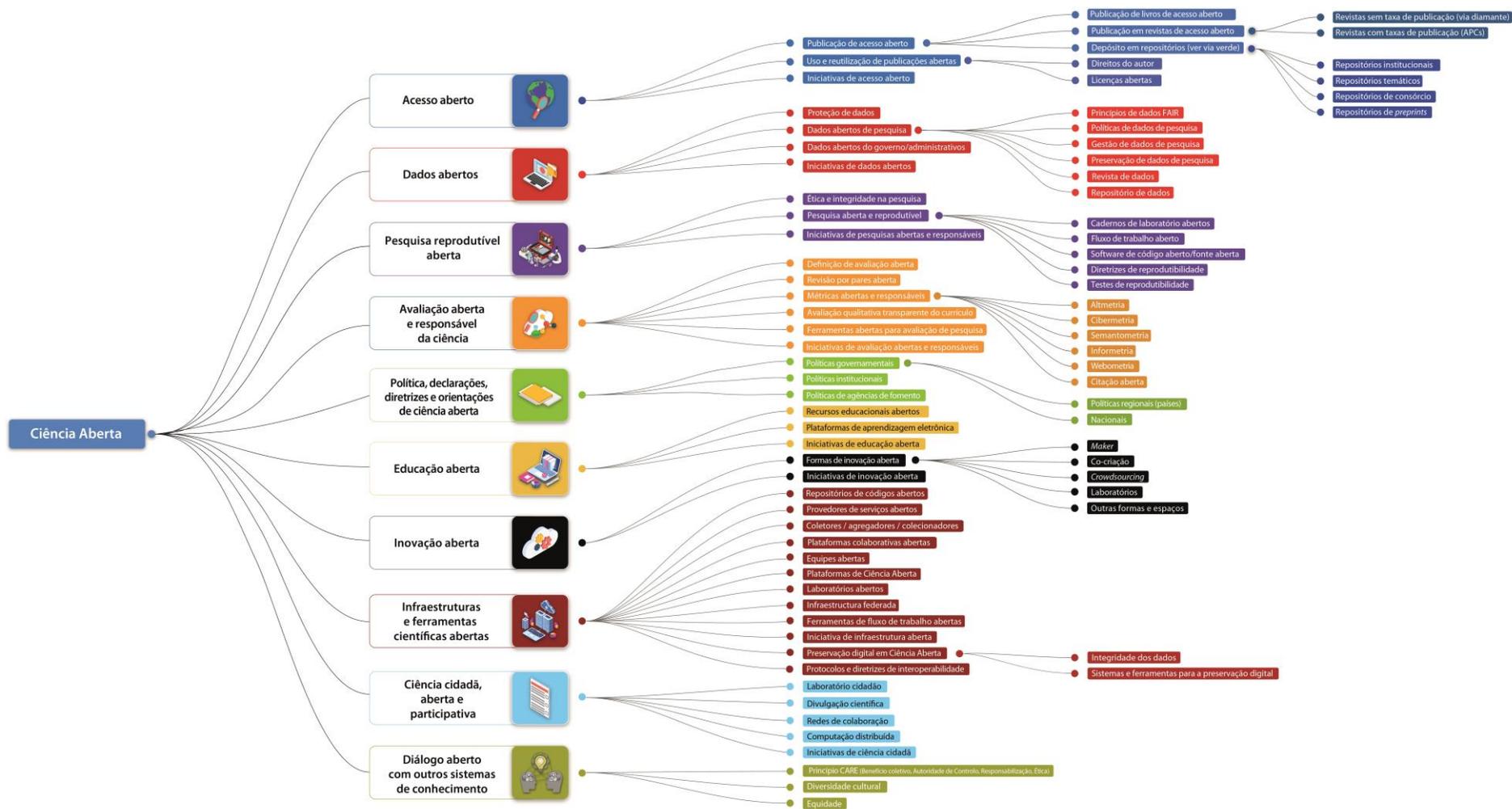
eficazmente por novas pesquisas ou que permitam ao pesquisador interagir com outros pesquisadores” (Shintaku; Sales, 2019, p.11).

Logo, o tema da Ciência Aberta é amplo, permeado por muitas dimensões. Na Figura 3 apresenta-se a Taxonomia de Ciência Aberta elaborada por pesquisadores brasileiros. Trata-se de uma atualização da Taxonomia Foster, que foi inicialmente concebida por um grupo europeu comprometido com o projeto Open Science do Programa Comunitário de Investigação e Inovação, intitulado Horizon 2020.

Cabe citar que a Figura 3 é uma ampliação da versão que havia sido apresentada pelos pesquisadores brasileiros em 2021 (Silveira *et al.*, 2021), naquela ocasião, uma atualização daquela proposta do grupo europeu se fez necessária, uma vez que, “o uso dessa taxonomia em alguns estudos pode ter sido enviesado, além disso, pelo tempo de existência da taxonomia os aprofundamentos da literatura trouxeram novos elementos e relacionamentos” (Silveira *et al.*, 2021, p. 3).

Já em 2023, novos elementos foram incluídos considerando as recomendações da UNESCO sobre Ciência Aberta e o contexto atual da comunicação científica (Silveira *et al.*, 2023). Por meio da classificação hierárquica de conceitos e objetos, “[...] a taxonomia pode ser utilizada para ajudar a entender melhor a Ciência Aberta e suas diferentes facetas, bem como para orientar o desenvolvimento de políticas e programas de treinamento e educação em Ciência Aberta” (Silveira *et al.*, 2023, p. 10).

Figura 3 - Taxonomia de Ciência Aberta³



Fonte: Silveira *et al.* (2023, p. 16).

³ <https://zenodo.org/record/7836884>

Assim, a Taxonomia da Ciência Aberta, apresentada na Figura 3, “remonta à ideia de um ecossistema: é um conjunto de atividades que interagem entre si e impulsionam umas às outras” (Silveira *et al.* 2021, p. 7). Isso quer dizer, também, que à medida que avança, a Ciência Aberta pode incorporar outras frentes além destas representadas na Figura 3. Assim, além do acesso aberto às publicações científicas, práticas científicas abertas incorporam,

[...] aspectos técnicos e tecnológicos (como o desenvolvimento de ferramentas livres, disponibilidade de plataformas computacionais abertas, e infraestrutura tecnológica para compartilhamento de dados), [...] questões de ordem cultural, política e institucional (formais e informais) que interferem no caráter aberto ou proprietário dessas práticas (Albagli, 2015, p. 7).

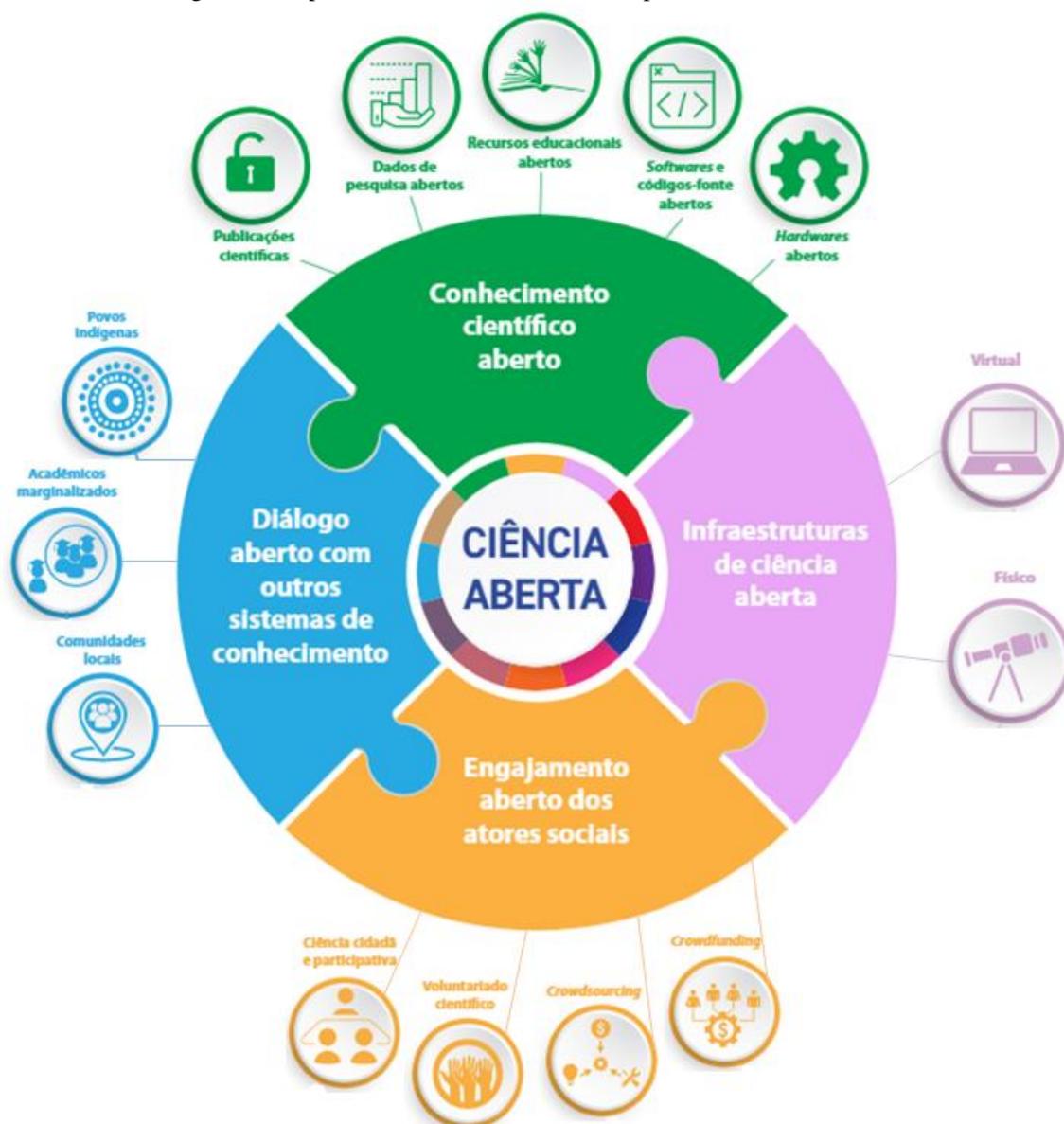
Nesse sentido, Albagli (2015) conclui que a Ciência Aberta mobiliza desde o dinamismo das práticas científicas, de tecnologia e de inovação, até as múltiplas possibilidades de colaboração entre pesquisadores e de participação da sociedade.

A partir dessa noção, a Conferência Geral da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), após inúmeras reuniões para discutir o tema emergente da Ciência Aberta, em sua 41ª sessão, realizada em 2021, elaborou um documento amplo com recomendações que determinam o padrão e organização da ciência, da tecnologia e da inovação, sob o viés da Ciência Aberta. O documento é claro ao considerar que diferentes nações possuem diversidade de leis e culturas e, por isso, cada Estado-membro deve aplicar as disposições constantes na recomendação junto aos seus governantes, a partir de suas realidades (UNESCO, 2022). O documento apresenta o seguinte conceito:

[...] a Ciência Aberta é definida 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. Abrange todas as disciplinas científicas e todos os aspectos das práticas acadêmicas, incluindo ciências básicas e aplicadas, ciências naturais, sociais e humanas, e se baseia nos seguintes pilares-chave: conhecimento científico aberto, infraestrutura científica aberta, comunicação científica, envolvimento aberto dos atores sociais e diálogo aberto com outros sistemas de conhecimento (UNESCO, 2022, p. 7).

Por meio dessa definição, é possível pensar a Ciência Aberta sob quatro principais aspectos: 1 – Conhecimento científico aberto; 2 – Infraestruturas de Ciência Aberta; 3 - Engajamento aberto dos atores sociais; 4 – Diálogo aberto com outros sistemas de conhecimento (Figura 4).

Figura 4 – Aspectos a serem considerados nas práticas de Ciência Aberta



Fonte: UNESCO (2022) adaptado pela autora.

Na Recomendação da UNESCO sobre Ciência Aberta se reconhecem-se as diferenças regionais para efetivação de ações práticas nas perspectivas de Ciência Aberta e uma diversidade de agendas como ciência e gênero, países em desenvolvimento, desigualdades de acesso a recursos digitais e tecnológicos, entre outras (UNESCO, 2022). No que concerne à comunicação científica e aos periódicos científicos, recortes do presente estudo, alguns dos aspectos apresentados na Figura 4 se mostram mais relevantes:

1 – Quanto ao conhecimento científico aberto: o acesso a todos os produtos das pesquisas científicas deve ser o mais aberto possível, a saber, as publicações científicas - como livros, artigos de periódicos, trabalhos apresentados em eventos, entre outros, revisados por pares; dados científicos abertos - todos os dados obtidos ou produzidos durante as

pesquisas, independente do formato, e que sejam apresentados em formatos de fácil interpretação tanto por máquinas, quanto por humanos;

2 – Infraestruturas de Ciência Aberta: notadamente as virtuais, que incluem plataformas e repositórios de acesso aberto, recursos que permitam identificar objetos digitais, em sua unicidade, por identificadores únicos e persistentes, como o Identificador de Objeto Digital (DOI), que pode ser, por conseguinte, atribuído a publicações científicas e a conjuntos de dados científicos abertos;

3 - Engajamento aberto dos atores sociais: iniciativas que envolvam a sociedade na fase de produção das pesquisas científicas o que pode incluir práticas de ciência cidadã, *crowdfunding* (por exemplo, por meio do financiamento de pesquisas que sejam de interesse comum, por grupos de indivíduos) e *crowdsourcing* (grosso modo, construções a partir de conhecimentos coletivos);

4 – Diálogo aberto com outros sistemas de conhecimento: como exemplo, os acadêmicos com menor capital científico, poderiam contribuir com pesquisas cujos resultados estão disponíveis em repositórios de *preprints*. Além disso, há de se considerar a importância do setor privado, principalmente no que concerne ao financiamento das pesquisas. Vale ressaltar que “[...] existem múltiplos atores e partes interessadas nos sistemas de pesquisa e inovação, e cada um deles tem um papel a desempenhar na operacionalização da Ciência Aberta” (UNESCO, 2022, p. 16).

A partir dessa abordagem mais geral sobre o movimento pela Ciência Aberta, parte-se para a apresentação de algumas iniciativas associadas à abertura do fazer científico, no cenário global, nos últimos anos.

2.4.1 Cenário Internacional

As iniciativas em prol da Ciência Aberta em diferentes países apontam como foco emergente do movimento, a abertura de dados científicos. As ações adotadas pelos países possuem muitos pontos em comum como, por exemplo, a criação de políticas públicas e abertura de dados governamentais (Governo Aberto) para a execução de práticas mais transparentes, colaborativas e a participação social em prol do avanço da ciência; agências de fomento à pesquisa e financiadores privados passam a exigir publicação aberta tanto de artigos, quanto de dados; criação de novas métricas e indicadores para avaliar a atividade de pesquisa; entre outras (Santos; Santos, 2022). Seguem as iniciativas de alguns países:

Quadro 1 – Experiências internacionais em prol da Ciência Aberta

Origem	Exemplo de Iniciativa	Principais focos da iniciativa
União Europeia	<i>Horizon Europe</i> (Continuação do <i>Horizon 2020</i>)	Reforma dos sistemas de avaliação da investigação e multilinguismo para aumentar a visibilidade dos resultados da investigação da UE (CONSELHO UE, 2022).
Alemanha	<i>Alliance of Science Organisations in Germany</i>	Preservação e acesso aberto (via dourada), em longo prazo, dos dados científicos financiadas com recursos públicos; Acordos entre cientistas e órgãos públicos em prol do uso responsável dos fundos públicos para ciência; Hospedagem em rede de publicações científicas digitais; Gestão de dados científicos com foco nos requisitos de metadados e interoperabilidade; Plataforma de pesquisa colaborativa; Propostas de alterações em leis que tratem sobre direitos autorais (Santos; Almeida; Henning, 2017).
Austrália	<i>Australian National Data Service</i> (ANDS)	Compartilhamento de dados por meio de bases de dados e repositórios acessíveis; Preocupação em agregar valor aos dados científicos; Estabelecimento de parcerias; Incentivo à elaboração do plano de gestão de dados que permita boas práticas de gestão, armazenamento, disseminação e reutilização dos dados (Santos; Almeida; Henning, 2017).
Canadá	1 - <i>Open Data Institute</i> 2 - <i>Action Plan on Open Government</i>	1 - Interlocações com o governo, com empresas, com organizações filantrópicas com vistas a influenciar decisões políticas e promover infraestruturas de abertura de dados ética e equitativa (ODI, 2022). 2 – Desenvolvimento e acompanhamento de políticas públicas que orientem os canadenses quanto à abertura das pesquisas científicas; Fornecer uma plataforma para os canadenses encontrarem e acessarem publicações de acesso aberto (GOVERNMENT OF CANADA, 2020).
EUA	<i>5th U.S. National Action Plan for Open Government</i> (2022-2024)	Envolve todas as partes interessadas da sociedade civil, agências governamentais e o público americano no desenvolvimento de práticas científicas com as seguintes premissas: transparência, responsabilidade, participação e colaboração (OPEN GOVERNMENT, 2022).
Holanda	<i>Netherlands Organization for Scientific Research</i>	Gerenciamento dos dados científicos; Publicações resultantes de financiamentos da NWO devem estar em acesso aberto imediato (via dourada, verde ou híbrida); Módulo de financiamento de práticas de Ciência Cidadã; Novas formas de avaliar pesquisas e pesquisadores (NWO, 2022).
Portugal	Política Nacional de Ciência Aberta	Transparência em todas as etapas das pesquisas científicas; Disponibilização e reutilização de dados científicos; Utilização de ferramentas de colaboração científica (CIÊNCIA ABERTA, 2022).
Reino Unido	<i>UKRI Open Access Policy</i>	Acesso aberto imediato para artigos de pesquisa, revisados por pares, podendo ser publicado em um periódico, por um editor, ou disponibilizado em um repositório, pelo próprio autor; Atribuição da licença Creative Commons (CC BY) para as publicações científicas; Estabelecimento de normalizações técnicas para periódicos e repositórios para facilitar busca, acesso, interoperabilidade e reutilização de artigos (UKRI, 2022).
França	<i>Plan national pour la science ouverte</i>	Tornar obrigatória a publicação em acesso aberto de toda produção científica (artigos e dados) resultante de financiamento público; Estimular práticas de Ciência Aberta, sobretudo nos programas de pós-graduação (MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE, 2019).

Fonte: elaborado pela autora.

Além dos exemplos apresentados no Quadro 1, Santos e Santos (2022) indicam que agências de fomento de importância global, como *Wellcome Trust*⁴, *Fundação Bill e Melinda Gates*⁵ e *National Institutes of Health (NIH)*⁶ estão exigindo, além da publicação em acesso aberto dos artigos e dados, também, a elaboração de um documento (Plano de Gestão de Dados) que detalhe os tipos de dados que serão produzidos durante a pesquisa, bem como questões referentes à sua preservação e compartilhamento futuros.

Diferente do que ocorre no cenário internacional, onde se percebe a preocupação já consolidada com as questões relativas à Ciência Aberta, no Brasil, as ações ainda estão em seus estágios iniciais. Na próxima seção algumas dessas iniciativas serão apresentadas.

2.4.2 Cenário Brasileiro

No Brasil observam-se iniciativas correlatas à Ciência Aberta promovidas por quatro principais agentes: 1 – Governo; 2 - Instituições científicas; 3 - Agências de fomento; e 4 - Instituições provedoras de produtos e serviços de informação (REZENDE; FALGUERAS, 2020). A seguir, apresentam-se algumas dessas iniciativas:

1 – No âmbito de ações promovidas ou integradas pelo Governo Federal destacam-se, pelo menos, onze marcos (Quadro 2).

Quadro 2 – Marcos na esfera do Governo Federal relacionados com a Ciência Aberta

Marcos Regulatórios Brasileiros (Ano)	Itens relacionados com a Ciência Aberta
Edital de Chamada FINEP/PCAL/XBDB N° 002/2009 (FINEP, 2009).	Edital destinado a apoiar projetos de implantação de repositórios institucionais nas instituições públicas de ensino e pesquisa.
Declaração de Governo Aberto (OGP, 2011).	Declaração de Princípios do Governo comprometendo-se a adotar medidas para o fortalecimento da transparência das informações e atos governamentais.
Decreto de 15 de Setembro de 2011- Institui o Plano de Ação Nacional sobre Governo Aberto e dá outras providências (Brasil, 2011).	Institui o Plano de Ação Nacional sobre Governo Aberto e o Comitê Interministerial de Governo Aberto (CIGA).
Decreto n.º 7.724/2012, de 18 de novembro de 2011 - Regulamentação da Lei de Acesso à Informação (LAI) (Brasil, 2012).	Regulamenta o direito constitucional de acesso dos cidadãos às informações públicas.
Instrução Normativa SLTI/MP (n.º 4/2012) - Institui a Infraestrutura Nacional de Dados Abertos (INDA) (Ministério do Planejamento, Orçamento e Gestão, 2012).	Define uma política na qual conste, desde aspectos de compartilhamento de elementos técnicos, até critérios para geração e gestão de dados públicos.
Decreto de 12 de Março de 2013 - Altera o Decreto de 15 de setembro de 2011, que institui o Plano de Ação Nacional sobre Governo Aberto (Brasil, 2013).	Inclui o Ministério da Justiça no grupo executivo do CIGA, cuja função é prioritariamente: elaborar a proposta do Plano de Ação Nacional sobre Governo Aberto.

⁴ <https://wellcome.org/>

⁵ <https://www.gatesfoundation.org/>

⁶ <https://www.nih.gov/>

Decreto nº 8.777, de 11 de maio de 2016. Institui a Política de Dados Abertos do Poder Executivo federal – 3ª versão do Plano de Ação Nacional sobre Governo Aberto (Brasil, 2016a).	Aprimora a cultura de transparência pública e facilita o intercâmbio de dados entre órgãos e entidades federais. Fomentar a pesquisa científica e o compartilhamento de recursos de tecnologia da informação.
Decreto nº 8.789, de 29 de junho de 2016. Dispõe sobre o compartilhamento de bases de dados na administração pública federal (Brasil, 2016b).	Os órgãos e as entidades da administração pública federal direta, detentoras ou responsáveis pela gestão de bases de dados oficiais, disponibilizarão aos órgãos e às entidades da administração pública federal o acesso aos dados sob a sua gestão.
Lei n.º 13.709, de 14 de agosto de 2018 - Lei Geral de Proteção de Dados Pessoais (LGPD) (Brasil, 2018).	Disciplina a proteção dos dados pessoais e define as situações em que estes podem ser coletados e tratados tanto por empresas quanto pelo Poder Público.
Decreto n.º 10.160, de 9 de dezembro de 2019. Institui a Política Nacional de Governo Aberto e o Comitê Interministerial de Governo Aberto - Publicação do 4º Plano de Ação Nacional sobre Governo Aberto (Brasil, 2019).	Definição dos temas prioritários para o governo: Inovação, Governo Aberto na Ciência e governança de dados científicos para o avanço da Ciência Aberta no Brasil.
Portaria n.º 3.679, de 07 de agosto de 2019. Dispõe sobre a recriação do grupo de trabalho para elaboração de proposta de política nacional para Ciência Aberta. (Ministério da Ciência, Tecnologia Inovações e Comunicações, 2019).	Recriação do grupo de trabalho para elaboração de proposta de política nacional para Ciência Aberta - Grupo formado por representantes de diversos órgãos que compõem o MCTIC e Ministério da Educação, dentre eles IBICT, CNPQ, Capes.

Fonte: Rezende e Falgueras (2020) adaptado pela autora com base em Santos e Santos (2022).

2 – Quanto às iniciativas em prol da Ciência Aberta promovidas por instituições científicas, considerando que, no Brasil, mais de 95% da produção científica advém das universidades públicas (Moura, 2019) e com base no estudo aprofundado das políticas dessas instituições, elaborado por Rezende e Falgueras (2020), infere-se que as ações estão principalmente centradas na criação e manutenção dos repositórios institucionais de produção científica, ou bibliotecas digitais.

3 – Quanto às agências de fomento às pesquisas científicas, no cenário brasileiro, destacam-se a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), ambas as instituições assumem compromissos de promover iniciativas, conjuntamente entre si, e com outras organizações, pelo acesso aberto e governo aberto.

Assim pode-se exemplificar a criação, pela CAPES, do Banco de Teses e Dissertações (BDTD) e do Portal de Periódicos. Além disso, a CAPES integra a série Compromisso pela Ciência Aberta em parceria com a Fundação Oswaldo Cruz (Fiocruz), Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Comissão Nacional de Energia Nuclear (CNEN) e Empresa Brasileira de Pesquisa Agropecuária (Embrapa). Como produtos dessa

parceria, se pode citar uma série de vídeos⁷ que pretende sensibilizar e explicar pontos importantes sobre Ciência Aberta.

Já o CNPq faz parte de um grupo de trabalho do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) que objetiva a criação da política nacional para a Ciência Aberta (CNPQ, 2022).

Nesse sentido, cabe citar a Chamada CNPq/MCTI-FNDCT N° 61/2022, com vistas a apoiar projetos de ciência cidadã e de difusão científica para o desenvolvimento e o aperfeiçoamento do Programa Ciência no Mar e Ciência Antártica do MCTI, entre outras ações relacionadas com o tema (CNPQ, 2022). Além de abordar sobre a ciência cidadã, uma das vertentes da Ciência Aberta, e sobre a divulgação científica, a chamada em questão, e outras disponíveis na página do CNPq na *internet*⁸ apresentam itens importantes quanto à preocupação com a Ciência Aberta:

[...] 5.2.3 - O pagamento de despesas de publicação deverá priorizar o modelo de acesso aberto.

[...] 10.7.2.2 – Em que pese a possibilidade de restrição de acesso, esta não reduz, contudo, a responsabilidade que os pesquisadores, suas equipes e instituições têm, como membros da comunidade de pesquisa, de manter, sempre que possível, os resultados da pesquisa, dados e coleções à disposição de outros pesquisadores para fins acadêmicos.

[...] 14.3 – Os artigos científicos resultantes dos projetos apoiados deverão ser publicados, preferencialmente, em periódicos de acesso público e depositados, em conjunto com os dados científicos e com todo material suplementar relacionado, em repositórios eletrônicos de acesso público (CNPQ, 2022).

No âmbito dos estados da federação, uma agência de fomento que merece destaque é a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). A fundação lançou em 2019 a Portaria CTA n° 01/2019 que institui a política para acesso aberto às publicações resultantes de auxílios e bolsas FAPESP e Portaria CTA n. 37, de 27 de outubro de 2021 que altera o Anexo 1 da primeira portaria (FAPESP, 2019). Em suma, os documentos preveem que as publicações científicas oriundas de pesquisas financiadas pela FAPESP sejam publicadas em acesso aberto, no mínimo, pela via verde. Além disso, os editais de financiamento exigem o envio de um documento chamado Plano de Gestão dos Dados (PGD) no ato das submissões, devendo contemplar informações sobre a coleta, disponibilização, acesso e uso dos dados obtidos ou coletados nas pesquisas financiadas pela FAPESP. Na seção sobre gestão de dados científicos se discorrerá mais detalhadamente sobre esse documento.

⁷ https://www-periodicos-capes-gov-br.ez49.periodicos.capes.gov.br/index.php?option=com_pnews&component=NewsShow&cid=931&mn=0&

⁸ <http://memoria2.cnpq.br/web/guest/chamadas-publicas>

4 – No que concerne às instituições provedoras de produtos e serviços de informação cabe ressaltar os papéis da *Scientific Electronic Library* (SciELO) e do Instituto Brasileiro em Informação, Conhecimento e Tecnologia (IBICT).

A SciELO é uma biblioteca digital de acesso aberto que visa a disseminação e avaliação da produção científica de periódicos científicos. Nesse sentido, a SciELO é pioneira, uma vez que, já em 1998, sob a liderança do Brasil e do Chile, adotou a abertura de suas publicações científicas, ou seja, quatro anos antes da formalização do movimento de acesso aberto pela Declaração de Budapeste (Packer, 2021).

Ainda nesse tocante, o documento SciELO – Linhas prioritárias de ação 2019-2023 (SCIELO, 2021) estabelece ações que se alinhem com a Ciência Aberta no que se refere a pontos específicos: aceitação de manuscritos previamente depositados em um servidor de *preprint*; publicação em fluxo contínuo; fortalecimento da abertura e transparência do processo de avaliação por pares, deixando claro que essa política é de autonomia de cada periódico, entretanto a SciELO se posiciona pela necessidade de que os periódicos explicitem claramente aos autores como esse processo ocorre; gestão de citação e referências dos dados, códigos e materiais. O documento traz como possíveis ações:

[...] disponibilizar o conteúdo da pesquisa que inclui os pré-registros da pesquisa, quando isso se aplica, o depósito dos dados utilizados na pesquisa, os métodos de análise utilizados, incluindo os códigos dos programas de computador. Para cada tipo de conteúdo existem repositórios disponíveis na web, em muitos casos relacionados a disciplinas ou áreas temáticas específicas. A segunda ação é elaborar uma descrição dos dados que foram depositados orientada ao uso dos dados, o que é feito através dos chamados artigos de dados, notas de dados, ou notas de pesquisa. Ao utilizar estas opções de abertura de métodos e materiais, os autores podem comunicar suas pesquisas combinando três modos diferentes e complementares de comunicação: os dados e seus metadados, a descrição dos dados e os artigos propriamente ditos (SCIELO, 2021, p. 5-6).

Complementarmente, a SciELO possui o SciELO *Data* e o SciELO *Preprints*. O primeiro pode ser utilizado para depósito, preservação e disseminação de dados científicos de artigos submetidos ou aprovados para publicação ou já publicados em periódicos da Rede SciELO ou depositados no SciELO *Preprints*, que, por sua vez, funciona como uma coleção de manuscritos não revisados por pares dentro da Rede SciELO.

Em relação ao IBICT, este também possui diferentes frentes para consolidar práticas de Ciência Aberta, como o manifesto brasileiro de apoio ao acesso livre à informação científica, de 2005, que foi o pontapé inicial para a criação dos repositórios institucionais, e o manifesto de acesso aberto a dados da pesquisa brasileira para ciência cidadã, de 2016, que considera a abertura dos dados científicos, um importante passo para a consolidação de práticas de Ciência Aberta.

Além disso, foi de responsabilidade do IBICT, a criação do portal brasileiro de publicações científicas em acesso aberto, o Oasisbr, que se conecta à rede de periódicos científicos de acesso aberto de Portugal.

Uma iniciativa mais recente, de 2022, e fruto da parceria entre IBICT e CNPq é a plataforma Lattes *Data*, um repositório para depósito de dados provenientes de pesquisas que tenham sido fomentadas pelo CNPq e que utiliza o software do repositório *Dataverse*, elaborado e gerido pela Universidade de *Harvard*, dos Estados Unidos da América.

Cabe ressaltar que não se pretendeu listar, de modo exaustivo, todas as iniciativas advindas de órgãos do governo ou de outras instituições, interessadas em promover a Ciência Aberta. Logo, se pretendeu, tão somente, apresentar um panorama geral e nacional dos esforços empreendidos sobre o tema. Infere-se, assim, que, estando o acesso aberto bem estabelecido e reconhecido como ação necessária à garantia do direito à livre circulação do conhecimento científico, as necessidades se voltam agora, principalmente, para as questões sobre a abertura de dados, tanto governamentais (Governo Aberto), quanto de pesquisa.

Concebe-se que essa abertura pode representar um desafio, sobretudo porque graças aos avanços das Tecnologias da Informação e Comunicação (TICs) há de se considerar um vasto volume de dados gerados diariamente: a era da *eScience*, ou o quarto paradigma da evolução científica, baseado no uso intensivo de dados e a defesa de que esses dados necessitam de uma curadoria bem estabelecida a fim de que estejam constantemente acessíveis a qualquer pessoa interessada (Hey; Tansley; Tolle, 2011).

Logo, pensar em abertura de dados é também fazer reflexões sobre a própria conceituação de dados científicos, sobre as políticas e infraestruturas para que esses dados sejam abertos e sobre os requisitos necessários para sua gestão e curadoria, pontos que serão abordados na próxima seção.

2.5 DADOS CIENTÍFICOS

Na primeira seção deste estudo discorreu-se sobre as revoluções (paradigmas) da própria Ciência. Esta seção se inicia com o seguinte questionamento: o que possibilita as revoluções dentro dos campos científicos? Dentre muitas possíveis respostas, concebe-se que o conhecimento acumulado é a base sobre a qual as próximas ligações teóricas e procedimentos experimentais serão construídos (NATIONAL RESEARCH COUNCIL, 1997).

Nesse sentido, os dados provenientes de pesquisas científicas são muito importantes ao permitirem o embasamento de novos estudos, a reprodução de experimentos anteriores, ou ainda, serem utilizados para outras análises. Conceitualmente,

“os dados científicos são as informações registradas ou produzidas através de qualquer forma ou meio durante o decurso de uma pesquisa. Podem ser de tipos numérico, descritivo ou visual e reproduzir-se em formato de papel [...] ou digital. [...] Em suma, [...] são todas as evidências que um investigador necessita para validar suas conclusões após uma pesquisa” (Silva, 2019, p. 21).

Sobre os dados científicos há de se considerar, que existem basicamente duas caracterizações referentes à quantidade de dados gerados, a *Big Science* (grandes projetos colaborativos que, naturalmente, já trazem a noção de altos custos, investimentos e infraestruturas bem definidas, tanto tecnológicas, quanto de pessoal) e a *Small Science* (pequenos laboratórios e pesquisadores individuais ou pequenos grupos, nas universidades e institutos de pesquisa, que desenvolvem um grande número de projetos científicos). A esse conjunto de dados originado na *Small Science*, alguns estudiosos chamam, metaforicamente, de cauda longa da ciência, pois gera uma quantidade expressiva e heterogênea de dados, mais difíceis de compreender, manipular e arquivar, enquanto os dados provenientes da *Big Science* seriam a “cabeça da ciência” (Sales; Sayão, 2019). O Quadro 3 apresenta as principais diferenças entre essas iniciativas.

Quadro 3 - Dados da cabeça e da cauda longa: principais diferenças

		Cabeça	Cauda Longa
Uniformidade	Diversidade	Homogêneos	Heterogêneos
	Geração/Coleta	Instrumentos automatizados	Gerados/coletados manualmente
	Procedimentos	Padronizados	Específicos
Gestão	Curadoria	Centralizada/institucionalizada	Individual
	Repositórios digitais	Disciplinares ou referenciais	Institucionais ou multidisciplinares
	Preservação	Preservados	Não preservados
	Armazenamento	Sistemas de <i>Storage</i>	Computadores pessoais/dispositivos portáteis
	Estruturação	Banco de dados	Planilhas
Compartilhamento	Acesso	Acesso aberto/distribuído	Obscuro ou protegido
	Reuso	Imediato/globalizado	Episódico/ entre a equipe
Institucionalização	Financiamento	Fluxo contínuo/ Apoio institucional	Por Projeto
	Reconhecimento/Recompensa	Sim	Não

Fonte: Sales e Sayão (2019).

Por meio do comparativo apresentado no Quadro 3 fica evidente que a instrumentação padronizada, tanto dos métodos de coleta, quanto da posterior curadoria, facilita o acesso e uso dos dados na chamada “*Big Science*”. Em contraposição, em cada pesquisa da cauda

longa, a depender do campo científico, apresentam-se diferentes métodos de coleta e formato dos dados e, por esse motivo, “os dados não são encaminhados para qualquer infraestrutura formal de arquivamento, por conseguinte, recaem, na maioria dos casos, no desuso e na obscuridade” (Sales; Sayão, 2019, p. 162). Visando demonstrar a grande variedade de tipos de evidências, fontes de informações de coleta e formatos de dados das pesquisas científicas, no Quadro 4 elucidam-se alguns exemplos provenientes de distintas áreas do conhecimento, tanto na cauda longa da ciência, quanto na “*Big Science*”.

Quadro 4 – Exemplos de variações nos tipos de evidências, fontes de coleta e formatos de dados científicos

	Engenharias e Ciências Biológicas	Ciências Sociais	Ciências Humanas
Evidências	Observações; Mensurações.	Observações; Mensurações; Fontes primárias (como periódicos e livros).	Fontes primárias; Indícios registrados.
Fontes de coleta	Instrumentos de coleta; Pesquisa em laboratório; Ensaios.	Instrumentos de coleta; Pesquisa de campo; Entrevistas; Questionários.	Instrumentos de coleta; Artes; Artefatos; Documentos; Manuscritos.
Formatos	Dados binários; Imagens; Anotações de laboratório; Amostras; Espécimes; Dados tabulares.	Registro de áudio; Diários; Livros; Documentos; Imagens; Planilhas; Transcrições.	Registros de áudio; Livros; Documentos; Imagens; Mapas; Arquivos de texto; Arquivos 3D.

Fonte: elaborado pela autora.

Ao observar o Quadro 4 e retomar o conceito de dados científicos de Silva (2019) pode-se afirmar que a abertura de tais dados é benéfica para todas as áreas do conhecimento, não somente para reprodução das pesquisas, mas, sobretudo, para evidenciar sua transparência, validar os resultados apresentados nos manuscritos científicos ou possibilitar outros estudos com integração de dados de diversas fontes.

Exemplos importantes de grandes coleções na área de Ciências Sociais incluem o repositório holandês de dados arqueológicos (com fotos, mapas, vídeos e milhares de relatórios de escavações, que podem ser acessados por meio do centro de dados de pesquisa daquele país⁹) e um diretório de mais de 200 coleções alemãs de dados em artes e ciências humanas e sociais disponibilizado pelo projeto europeu *Digital Research Infrastructure for the Arts and Humanities*¹⁰ (DARIAH) (Laender *et al.*, 2020, p. 8).

Diante disso, sabe-se que os pesquisadores de todas as áreas do conhecimento têm optado pelos periódicos científicos para disseminação dos resultados de suas pesquisas científicas, sobretudo, por conta da credibilidade creditada aos trabalhos, ocasionada pela

⁹ <https://easy.dans.knaw.nl/ui/home>

¹⁰ <https://www.dariah.eu/tools-services/tools-and-services/tools/collection-registry/>

revisão por pares, e a publicação mais acelerada em comparação aos demais itens informacionais impressos (Miranda, Carvalho, Costa, 2018). Por esses motivos, as discussões desta seção estão centradas nos elementos necessários para que os dados científicos sejam, mais do que simplesmente abertos, também localizáveis, acessíveis, interoperáveis e reutilizáveis, ou seja, que haja a gestão desses dados de acordo com os princípios FAIR, em sua língua de origem, *findable, accessible, interoperable, reusable*. A subseção a seguir situa os princípios FAIR no cenário da abertura dos dados científicos, por meio dos repositórios.

2.5.1 Gestão de dados científicos: os princípios FAIR fomentando a sua abertura em repositórios

Como evidenciado na seção 2.4.2 sobre o cenário brasileiro da Ciência Aberta, as primeiras ações do país estavam centradas na abertura dos dados governamentais e das produções acadêmicas dos programas de pós-graduação das universidades (teses e dissertações).

O movimento de abertura de dados científicos começa a surgir proveniente das agências de fomento à pesquisa, tanto públicas, quanto privadas. Como o exemplo da FAPESP, citada como pioneira por exigir, no mínimo, a elaboração de um plano de gestão de dados.

Fato é que, principalmente no Brasil, a cultura de compartilhamento de dados científicos ainda está sendo construída e há, por certo, muitos desafios que envolvem “[...] custos inerentes, como para geração e manutenção de identificadores, aquisição e manutenção de hardware e software, manutenção e atualização de infraestrutura física, e treinamento e formação de pessoal” (Laender *et al.*, 2020, p. 11-12).

Os autores ainda citam barreiras a serem enfrentadas na esfera internacional, como recursos computacionais e protocolos de coordenação para coleta, armazenamento, processamento de grandes volumes de dados e, posteriormente, acessibilidade e interoperabilidade para tais dados. “Soma-se a estes a necessidade crescente do uso de tecnologias e métodos baseados em inteligência artificial para análise dos dados, os quais, cada vez mais, se tornam por demais volumosos para que sejam analisados, inspecionados diretamente ou supervisionados por pessoas” (Laender *et al.*, 2020, p. 12).

Ademais, para a abertura de dados, além das mencionadas barreiras de infraestrutura técnica e tecnológica, as quais as instituições de pesquisa precisam se adequar, há, ainda, dificuldades de ordem da cultura dos cientistas “como, por exemplo, a resistência de

pesquisadores em ceder seus dados por medo de perder a prioridade de descobertas” (Santos; Guanaes, 2018). Por esse motivo é que Silva (2019, p. 45) alenta que, “para qualquer iniciativa em que seja necessário confiar na boa vontade de compartilhar dados, deve ser considerada a sociologia da ciência: os pesquisadores deverão ver as vantagens de compartilhá-lo e necessitarão de incentivos para fazê-lo”.

O auxílio financeiro às suas pesquisas é um incentivo atraente, mas, certamente, ele não é o único. Além disso, Federer *et al.* (2018) admitem que, no caso dos periódicos científicos, as políticas que encorajam (incentivam) ou aquelas que exigem o compartilhamento de dados, têm implicações significativas em relação à quantidade de pesquisadores que realmente disponibilizam seus dados. Os mesmos autores julgam que “é improvável que os pesquisadores compartilhem dados na ausência de requisitos formais ou incentivos específicos” (Federer *et al.*, 2018, p. 2).

Seguindo a linha de vantagens para o compartilhamento de dados, Borgman (2015, p. 8, tradução nossa) apresenta quatro justificativas: “(1) reproduzir pesquisas; (2) disponibilizar ativos públicos; (3) alavancar investimentos em pesquisa; e (4) pesquisa avançada e inovação”. A seguir pretende-se refletir sobre esses quatro motivadores.

(1) Reproduzir pesquisas – ao se mencionar essa expressão, é importante compreender que ela está paralelamente associada a outra marca da “boa ciência”, a replicabilidade. Ambas são mecanismos de validação ou autocorreção da ciência (*National Academies of Sciences, Engineering, and Medicine*, 2019).

Reprodutibilidade e replicabilidade começaram a ganhar destaque quando o uso de ferramentas computacionais começou a se expandir, num contexto de repetibilidade, ou seja, quando um estudo confirma o resultado de outro, entretanto os dois termos não significam a mesma coisa. Por definição, tem-se que:

Reprodutibilidade é obter resultados consistentes usando os mesmos dados de entrada; etapas computacionais, métodos e código; e condições de análise. Essa definição é sinônimo de “reprodutibilidade computacional”. **Replicabilidade** é obter resultados consistentes em estudos destinados a responder à mesma questão científica, cada um dos quais obteve seus próprios dados (*National Academies of Sciences, Engineering, and Medicine*, 2019, p. 46, tradução nossa, grifo nosso).

Desse modo, reproduzir pesquisas através de reprodutibilidade ou replicabilidade depende dos interesses em validar uma nova descoberta, principalmente quando os resultados dos estudos científicos têm impactos para toda a sociedade. Independente do contexto e da necessidade para a reprodução das pesquisas, o compartilhamento dos dados, dos métodos e etapas para se chegar a um determinado resultado, se mostra como ponto crucial (*National Academies of Sciences, Engineering, and Medicine*, 2019).

(2) Disponibilizar ativos públicos - frequentemente, os estudos são financiados por instituições públicas (Silva, 2019), logo todos os produtos resultantes dessas pesquisas, devem também estar disponíveis ao público.

(3) Alavancar investimentos em pesquisa – em relação a esta justificativa, Silva (2019, p. 45) aponta para o fato de que “organismos de financiamento realizam investimentos em pesquisa para se aprofundar em muitas áreas do conhecimento científico”, por conseguinte, o compartilhamento de dados apoia esse propósito.

(4) Pesquisa avançada e inovação – o compartilhamento de dados pode permitir estudos colaborativos entre pesquisadores de diferentes localidades do mundo, ou seja, a pesquisa científica tem potencial inovador, não só por meio da disponibilização dos dados, mas também da conexão entre diferentes grupos de pesquisa, possibilitando enriquecimento nas abordagens, metodologias e análises realizadas.

“Entende-se que a partir do momento em que houver um fluxo contínuo de estruturação, publicação e consumo (reuso) de dados científicos, haverá clareza sobre o processo de aceleração de descobertas [...]” (Santarem Segundo, 2020, p. 6). Nesse sentido, o mesmo autor cita a *Linked Data*, um conjunto de práticas para publicação e interligação de dados na internet, que utiliza os conceitos da *web* semântica, que, por sua vez, prevê a atribuição de vocabulários controlados aos dados e metadados de maneira a permitir que, por meio da *web*, dados sejam “*linkados*”. Tais iniciativas se mostram muito inovadoras e promissoras, entretanto há de se considerar que para sua efetivação necessitam de estruturas técnicas e de pessoal bem definidas, o que pode acarretar na demora para sua efetivação, em âmbito mundial.

Reforçando a premissa de que a preocupação com a gestão dos dados é muito importante para as pesquisas científicas, dentre os principais dos muitos motivos, tem-se:

- Aumenta o impacto e a visibilidade da pesquisa;
- Promove a inovação e novos usos potenciais de dados;
- Conduz a novas colaborações entre usuários de dados e criadores;
- Maximiza a transparência e a responsabilidade;
- Possibilita o escrutínio dos resultados da pesquisa;
- Incentiva a melhoria e a validação dos métodos de pesquisa;
- Reduz o custo para duplicar sua coleta;
- Fornece recursos importantes para educação e treinamento (Córdula; Araújo, 2019, p. 198).

Logo, quando um estudo é iniciado com a premissa da Ciência Aberta, o pesquisador deve estar ciente de que, desde o planejamento de sua pesquisa, os dados devem ser tratados como um ativo a ser gerado e que será, posteriormente, disponibilizado em acesso aberto, preferencialmente em um repositório, salvo as restrições legais que protegem alguns tipos de

dados. Como exemplo de restrições legais cabe-se citar: dados sensíveis dos participantes da pesquisa (pessoas ou instituições); dados de terceiros, os quais não se possui o direito de compartilhamento; entre outros.

Assim, a gestão dos dados científicos é uma tarefa que se inicia antes da coleta e tem nos princípios FAIR, elementos norteadores para descoberta, compartilhamento e reutilização de tais dados (Henning; Moreira, 2020), com vistas a permitir que os dados compartilhados sejam encontrados mais facilmente.

Os princípios FAIR surgiram em 2014 em uma reunião em Leiden, Países Baixos, e, em 2016, após publicação no periódico *Nature – Scientific Data* começaram a ser utilizados mundialmente como norteadores de boas práticas de gestão de dados (Henning; Moreira, 2020), inclusive sendo requeridos pelos repositórios. O Quadro 5 sintetiza as premissas FAIR.

Quadro 5 – Os princípios orientadores FAIR

F - Ser encontrável (Findable)
E1. Os (meta)dados são atribuídos a um identificador persistente, único e global.
E2. Os dados são descritos com metadados ricos (definidos por R1 a seguir).
E3. Os metadados incluem, de forma clara e explícita, o identificador dos dados que descrevem.
E4. Os (meta)dados são registrados ou indexados em um recurso pesquisável.
A - Ser acessível (Accessible)
A1. Os (meta)dados são recuperáveis por seu identificador, usando-se um protocolo de comunicação padronizado.
A1.1 O protocolo é aberto, gratuito e universalmente implementável.
A1.2 O protocolo possibilita um procedimento de autenticação e autorização, quando necessário.
A2. Os metadados são acessíveis, mesmo quando os dados não estão mais disponíveis.
I - Ser interoperável (Interoperable)
I1. Os (meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representar o conhecimento.
I2. Os (meta)dados usam vocabulários que seguem os Princípios FAIR.
I3. Os (meta)dados incluem referências qualificadas para outros (meta)dados.
R - Ser reutilizável (Reusable):
R1. Os meta(dados) são ricamente descritos com uma pluralidade de atributos precisos e relevantes.
R1.1. Os (meta)dados são disponibilizados com uma licença de uso de dados clara e acessível.
R1.2. Os (meta)dados estão associados a uma proveniência detalhada.
R1.3. Os (meta)dados estão de acordo como padrões comunitários relevantes para o domínio.

Fonte: Dias; Anjos; Rodrigues (2019).

Com base no Quadro 5 faz-se necessário explicar alguns aspectos. Primeiramente, em muitos pontos utiliza-se o termo “metadados”. Estes são elementos descritivos que detalham a origem dos dados, âmbito, autoria, propriedade e condições de reutilização (Silva, 2019).

No que concerne à etapa de planejamento da pesquisa, o pesquisador deve se preocupar com questões referentes aos dados a serem coletados ou gerados. É preciso documentar estas informações visando ter o registro das ações referentes aos dados, possíveis de serem atualizadas, para isso existe o Plano de Gestão de Dados (PGD). Cabe esclarecer que o PGD ainda não é uma exigência generalizada, assim como os metadados, para depósito em

repositórios de dados, no entanto destaca-se sua importância por abordar, desde o início da pesquisa, como os dados serão tratados.

Existem diversas ferramentas, inclusive em formatos *online*, que facilitam a tarefa de redigir um PDG, como *DMP Tool*¹¹, *DMPonline*¹² e *PGDOnline*¹³. Entretanto, mesmo que um autor as utilize, é importante estar atento a elementos comuns sobre a gestão dos dados. Veiga *et al.* (2019) apresentam uma proposta de *template* que contém os principais componentes de um PDG, alinhados aos princípios FAIR, conforme Quadro 6 a seguir.

Quadro 6 - Proposta de template de Plano de Gestão de Dados do ICICT/Fiocruz

Seção A: Informação Administrativa	
A1.	Informe o nome completo do pesquisador principal
A2.	Informe o <i>e-mail</i> de Contato
A3.	Qual o seu telefone?
A4.	Informe o <i>link</i> do Currículo Lattes do Pesquisador Principal
A5.	Informe o <i>link</i> do ORCID do Pesquisador Principal.
A6.	Informe o título do seu projeto
A7.	Descreva o resumo do projeto
A8.	A qual Unidade da Fiocruz o seu projeto está vinculado?
A9.	Qual a data de início do projeto?
A10.	Indique a previsão de término do projeto
A11.	O projeto tem financiamento?
A12.	Especifique o financiamento
A13.	Qual a versão deste Plano de Gestão de Dados?
A14.	Quem será o responsável pelo preenchimento e atualização do Plano de Gestão de Dados?
Seção B: Descrição dos dados coletados ou reuso de dados existentes	
B1.	Como os dados serão coletados, produzidos ou como os dados existentes serão reutilizados?
B2.	Quais os tipos de dados que serão coletados ou produzidos?
B3.	Quais os formatos de dados que serão coletados ou produzidos?
B4.	Qual o volume aproximado dos dados coletados ou produzidos?
Seção C: Documentação e Qualidade dos Dados	
C1.	Indique os metadados adotados
C2.	Indique o padrão de metadados adotado

¹¹ <https://dmptool.org/>

¹² <https://dmponline.dcc.ac.uk/>

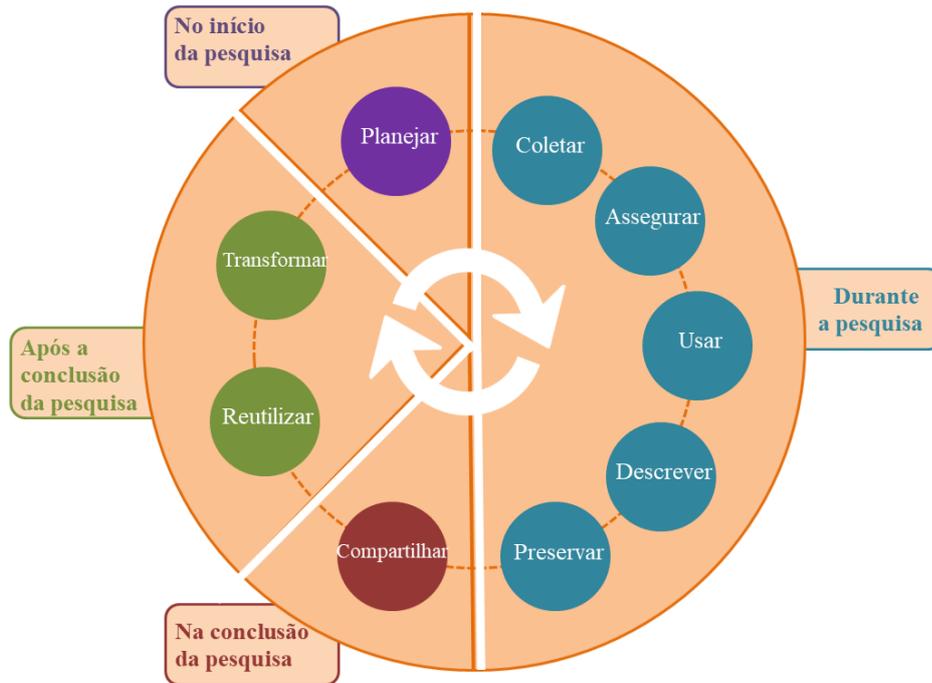
¹³ <https://dmp.consociomadrono.es/>

C3.	Indique os documentos que acompanharão os dados
C4.	Indique as medidas adotadas para a organização e controle de qualidade dos dados
Seção D: Armazenamento e Backup durante o processo de pesquisa	
D1.	Como os dados serão armazenados e como serão feitas as cópias de segurança durante a pesquisa?
D2.	Como a segurança dos dados e a proteção dos dados sensíveis serão tratadas durante a pesquisa?
Seção E: Requisitos Legais, Éticos e de Códigos de Conduta	
E1.	Se os dados pessoais forem tratados, como será assegurado o cumprimento da legislação relativa aos dados pessoais e à proteção dos dados?
E2.	Como as questões legais referentes aos direitos de propriedade intelectual serão gerenciadas? Qual legislação se aplica?
E3.	Indique a licença que será aplicada.
E4.	Como as possíveis questões éticas e os códigos de conduta serão levados em consideração?
Seção F: Compartilhamento de Dados e Preservação a longo prazo	
F1.	Como e quando os dados serão compartilhados? Existem possíveis restrições ao compartilhamento e/ou à abertura de dados ou motivos para embargo?
F2.	Como os dados para preservação serão selecionados e onde os dados serão preservados a longo prazo (por exemplo, em um repositório de dados)?
F3.	Quais métodos ou ferramentas de software serão necessários para acessar e usar os dados?
F4.	Como será assegurado o registro de um identificador único e persistente (como um DOI - Digital Object Identifier) para cada conjunto de dados?
Seção G: Responsabilidades e Recursos Financeiros para a Gestão de Dados Científicos	
G1.	Quem será o responsável pela gestão dos dados?
G2.	Quais recursos (ex.: financeiros, tempo) serão dedicados à gestão de dados e à garantia que eles sejam FAIR (localizável, acessível, interoperável, reutilizável)?
G3.	Para projetos colaborativos, explique como será a coordenação da gestão de dados e as responsabilidades de cada um dos parceiros.

Fonte: Veiga *et al.* (2019).

Com base nesses componentes (Quadro 6), concebe-se que no contexto de uma gestão eficiente dos dados científicos, o PGD é um documento que traz informações que perpassam por todas as etapas do ciclo de vida dos dados, conforme a Figura 5.

Figura 5 – Plano de Gestão de Dados integrado ao ciclo de vida dos dados científicos



Fonte: elaborado pela autora.

No início da pesquisa é preciso planejar, ou seja, definir quais dados serão coletados, quando e como ocorrerá a coleta e, previamente, estabelecer como serão organizados (formatos e estrutura).

Durante a pesquisa ocorrem as atividades de coleta dos dados, onde é preciso também assegurar efetivamente a qualidade dos dados (se os formatos, tipos e cruzamentos escolhidos são adequados e, também, é o momento onde se estabelecem códigos, unidades de medida, marcadores, dentre outros recursos que facilitem a realização de resumos estatísticos ou outros tipos de análises). Além do uso dos dados coletados para redação dos resultados encontrados no estudo, faz-se necessário, também, descrevê-los (definição dos metadados) e determinar como serão preservados, estando disponíveis futuramente (preservação).

Na conclusão da pesquisa, ou quando as etapas anteriores do ciclo de vida dos dados são concluídas, com vistas ao armazenamento em longo prazo, faz-se necessário que os dados sejam passíveis de serem compartilhados. O compartilhamento pode ocorrer de várias maneiras, desde as mais simples, como envio direto a pessoas interessadas, até o depósito em um repositório, considerada a maneira mais eficiente, uma vez que, amplia-se o alcance de acesso, principalmente por causa da normalização exigida na descrição dos dados, possibilidades de interoperabilidade com outros sistemas e as políticas que os depositantes

devem se atentar e seguir. Além disso, o compartilhamento por meio de repositórios possibilita a reutilização e transformação dos dados para outras pesquisas.

A expressão ‘repositórios digitais’, no contexto do acesso aberto, é empregada para denominar os vários tipos de aplicações de provedores de dados que são destinados ao gerenciamento de informação científica, constituindo-se, necessariamente, em vias alternativas de comunicação científica. Cada um dos tipos de repositórios digitais possui funções específicas e aplicações voltadas para o ambiente no qual será utilizado (Leite, 2009, p. 19).

Obrigatoriamente, todo conjunto de dados depositado em um repositório é acompanhado dos seus metadados, são eles que permitem a interoperabilidade técnica e semântica (Souza; Visoli; Torres, 2020), ou seja, conectam dados entre diferentes repositórios e plataformas de dados, um dos benefícios do compartilhamento de dados.

Embora alguns autores apresentem algumas variações, concebe-se que os repositórios existentes podem ser enquadrados em uma das três categorias:

1 – Repositórios institucionais: “voltados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa” (Leite, 2009, p. 22). Nesse caso poderiam ser incluídos os bancos de teses e dissertações das universidades;

2 – Repositórios temáticos: “tratam da produção intelectual de áreas do conhecimento em particular” (Leite, 2009, p. 22), como, por exemplo, o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBr) que é uma infraestrutura de software responsável pela “organização, indexação, armazenamento e disponibilização de dados e informações sobre a biodiversidade e os ecossistemas brasileiros” (SIBBR, 2022);

3 – Repositórios de propósito geral: podem ser utilizados por qualquer pesquisador “independentemente de sua filiação institucional, para preservar qualquer tipo de produção acadêmica” (Silva, 2019, p. 97).

Os repositório de dados científicos, assim como descrito anteriormente, podem estar associados a alguma instituição, aceitando o depósito exclusivo de dados de seus pesquisadores (repositório institucional), ou podem ser abertos ao público externo, de diferentes afiliações, nesse caso passa a ser um repositório de propósito geral.

A infraestrutura de software dos repositórios de dados científicos é configurada de modo a gerenciar, compartilhar, acessar e arquivar dados bem descritos e bem documentados, para tanto, deve seguir padrões internacionais relacionados a aspectos técnicos, metadados e termos de acordo com requisitos legais, visando basicamente garantir armazenamento, compartilhamento, facilidade de localização, interoperabilidade, proteção de dados, uso e reutilização (Rezende; Barbosa, 2021). A Figura 6 apresenta um ciclo de vida de um

repositório de dados científicos de forma a garantir tais aspectos de segurança, conformidade legal e qualidade.

Figura 6 - The research data repository lifecycle



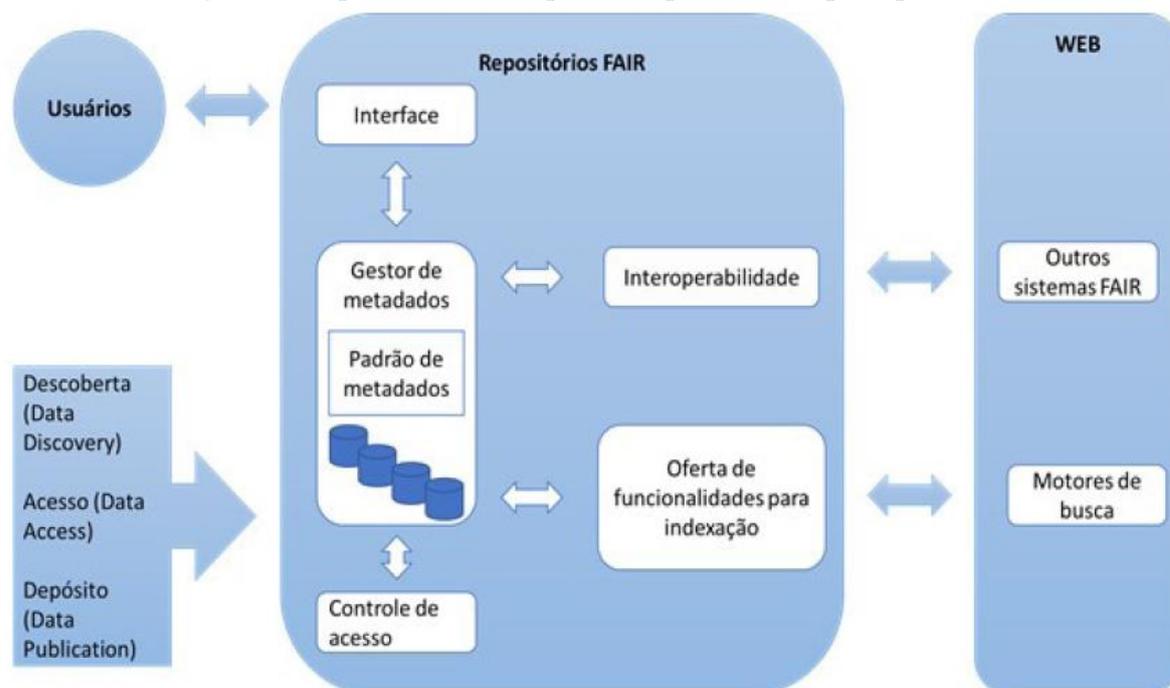
Fonte: Uzwyshyn (2016), traduzido pela autora.

A partir do explanado, Shintaku, Appel e Oliveira (2021) sintetizam que:

[...] a arquitetura de um repositório que atenda aos princípios FAIR atende a usuários humanos por meio de uma interface Web, oferecendo interoperabilidade para outros sistemas e permitindo que motores de busca indexem os seus metadados. O acesso aos metadados e as bases de dados deve ser controlado a fim de possibilitar disseminação escalonada devido a diferentes níveis de sensibilidade dos dados. Os metadados seguem padrões que facilitam a integração, descoberta e descrição das bases de dados (Shintaku; Appel; Oliveira, 2021, p. 148).

Na Figura 7 os autores representam a arquitetura pretendida, de acordo com os princípios FAIR. Neste contexto, pode-se citar o Repositório de Dados da Universidade de Harvard, que será foco da etapa de coleta documental do estudo aqui proposto. Cabe frisar que os repositórios de propósito geral, e até mesmo os temáticos, são uma boa alternativa para o depósito dos dados, associados aos artigos científicos publicados em periódicos científicos, mas antes, faz-se necessário a elaboração de políticas editoriais que orientem os autores a respeito dos dados científicos e que estejam alinhadas com as próprias condições de infraestrutura (Pavão, Silva, Silveira, 2020).

Figura 7 - Arquitetura de um repositório que atenda os princípios FAIR



Fonte: (Shintaku; Appel; Oliveira, 2021, p. 148).

A partir da Figura 7 se faz necessário elucidar sobre as questões envolvidas nas licenças de uso de dados científicos, especialmente para que se tornem reutilizáveis. Para tanto, na próxima seção se discorrerá sobre propriedade intelectual, direitos autorais e licenças.

2.5.2 Direitos autorais sobre dados científicos

A propriedade intelectual possui uma definição ampla e que contempla diferentes formas de proteção das criações humanas, como os direitos autorais, os segredos de negócios, as patentes, entre outros. Por definição, tem-se que “a propriedade intelectual (PI) refere-se a criações da mente, como invenções; obras literárias e artísticas; projetos; e símbolos, nomes e imagens usados no comércio” (WIPO, 2022a).

Os direitos autorais apresentam-se como parte dos direitos de propriedade intelectual que especifica os direitos “que os criadores têm sobre suas obras literárias e artísticas. Os trabalhos cobertos por direitos autorais vão desde livros, músicas, pinturas, esculturas e filmes, até programas de computador, bancos de dados, anúncios, mapas e desenhos técnicos” (WIPO, 2022b). A informação científica se enquadra nessa definição.

No Brasil, a Lei n.º 9610 de 19 de fevereiro de 1998 (Lei dos Direitos Autorais) (Brasil, 1998) protege as obras intelectuais, incluindo textos, composições musicais,

fotografias, programas de computadores, entre outras, nas quais é possível incluir também os artigos científicos e os conjuntos de dados provenientes dos estudos científicos. Na referida lei, Art. 29, fica claro que depende de autorização prévia e expressa do autor, a utilização da obra, por quaisquer modalidades.

No contexto dos trabalhos científicos, por exemplo, artigos e conjunto de dados, os periódicos e repositórios de dados precisam informar claramente os termos para uso e reuso das obras. Neste cenário, os repositórios de dados (assim como os periódicos científicos) têm adotado licenças que garantem os direitos dos autores, principalmente de serem citados no caso de reutilização. As mais utilizadas são as licenças *Creative Commons* (da organização global homônima que as fornece padronizada e gratuitamente) e possui como elemento obrigatório a necessidade do usuário dar o crédito pela criação original (Guanaes, 2020). O Quadro 7 apresenta algumas dessas licenças.

Quadro 7 – Algumas licenças para permissão de reuso de dados científicos em bases legais

Licença (Creative Commons)	Sigla	Características
Atribuição CC BY	CC BY	Distribuição, remixagem, adaptação e criação a partir do trabalho original mesmo para fins comerciais.
Atribuição Compartilhamento Igual	CC BY-SA	Remixagem, adaptação e criação a partir do trabalho original, mesmo para fins comerciais; Exige atribuição de crédito à obra original; Novas criações devem vir sob termos idênticos.
Atribuição Sem Derivações	CC BY-ND	Reuso para qualquer finalidade, inclusive comercial; não permite compartilhamento e adaptações; Exige atribuição de crédito à obra original.
Atribuição Não Comercial	CC BY-NC	Remixagem, adaptação e criação a partir do trabalho original para fins não comerciais; Exige atribuição de crédito à obra original; Novas criações não precisam vir sob termos idênticos.
Atribuição Não Comercial Compartilhamento Igual	CC BY-NC-SA	Remixagem, adaptação e criação a partir do trabalho original para fins não comerciais; Exige atribuição de crédito à obra original; Novas criações devem vir sob termos idênticos.
Atribuição Não Comercial Sem Derivações	CC BY-NCND	É a mais restritiva das seis licenças. Só permite downloads do trabalho original, o qual não pode ser alterado ou usado para fins comerciais. Deve-se atribuir o crédito pela criação original.
CC0 1 Universal Dedicção ao Domínio Público ¹⁴	CC0 1.0	Permite cópia, modificações, distribuição e execução do trabalho, mesmo para fins comerciais, sem necessidade de solicitar autorização prévia. Domínio Público.

Fonte: Guanaes (2020, p. 81-82).

¹⁴ Não são consideradas licenças, mas instrumentos jurídicos para informar a opção do detentor de direitos autorais pelo domínio público (Guanaes, 2020, p. 82).

Questões referentes aos direitos autorais podem vir explicitamente formalizadas nos repositórios onde se deseja compartilhar os dados científicos, cabendo aos autores, detentores desses direitos, assinalar as permissões que devem ser associadas a eles, ou ainda, podem integrar documentos formais de gestão de dados, como os planos de gestão de dados (PGD).

Todas as reflexões empreendidas até aqui fizeram refletir sobre o acesso aberto e público aos dados científicos, vistos como bens comuns. Para uma gestão efetiva é necessário levar em consideração que existem dados de diferentes tipos e naturezas (por exemplo, governamentais, científicos, entre outros) e que é preciso se preocupar com a proteção dos direitos dos autores. Essas questões devem estar claras nas políticas dos periódicos científicos que desejam requerer o compartilhamento dos dados que culminaram nos artigos que publicam.

Por fim, acredita-se que tais reflexões apresentaram-se como subsídios para o atendimento aos objetivos deste estudo, cujas etapas e resultados estão apresentados nas próximas seções.

3 CAMINHOS METODOLÓGICOS

O presente estudo pode ser classificado como uma pesquisa social aplicada, de abordagem qualitativa. Quanto aos objetivos, a pesquisa é exploratória e quanto aos procedimentos, documental e com dimensões da pesquisa-ação.

A partir da conceituação de Gil (2008) sobre a pesquisa social aplicada, com os métodos utilizados no seu desenvolvimento e os resultados alcançados espera-se contribuir com a realidade da UFG no que concerne à comunicação científica praticada por esta instituição através dos periódicos científicos sob sua gestão. Esse tipo de pesquisa “tem como característica fundamental o interesse na aplicação, utilização e consequências práticas dos conhecimentos” (Gil, 2008, p. 27).

O quadro a seguir apresenta um esquema de detalhamento dos objetivos específicos alinhados aos procedimentos e etapas da pesquisa:

Quadro 8 - Caracterização da pesquisa

Objetivos	Procedimentos e fontes de coleta	Etapas da pesquisa
Delinear um panorama teórico sobre Ciência Aberta no contexto da comunicação científica, em especial, a praticada por meio dos periódicos.	Pesquisa exploratória: Levantamento bibliográfico	Pesquisa bibliográfica
Diagnosticar periódicos científicos que já possuem coleções de dados em um repositório de dados, em âmbito nacional e internacional.	Pesquisa exploratória: Levantamento documental	Pesquisa documental em um repositório de dados aberto e público: o <i>Dataverse</i> da Universidade de <i>Harvard</i> . Localizaram-se periódicos científicos (nacionais e internacionais) com algum tipo de vínculo a Instituições de Ensino Superior, que possuem coleções de dados dentro do referido repositório.
Levantar e apresentar as normativas de abertura de dados científicos de periódicos nacionais e internacionais com base na certificação <i>Core Trust Seal</i> .	Pesquisa exploratória: Levantamento documental	Pesquisa documental nas páginas <i>web</i> dos periódicos científicos (brasileiros e internacionais) que já possuem coleções de dados no repositório de dados <i>Dataverse</i> da Universidade de <i>Harvard</i> , a fim de identificar e caracterizar suas diferentes normativas de abertura de dados científicos.
Idealizar, juntamente com a equipe editorial do periódico da UFG, uma proposta de caracterização, normativas e implantação da sua futura coleção de dados científicos no repositório da UFG;	Pesquisa-ação	Elaboração, junto à equipe editorial do periódico científico da UFG selecionado como piloto, o Boletim Goiano de Geografia, das normativas de abertura dos dados científicos associados aos artigos que publicam, bem como definições de acesso e uso à

		<p>coleção de dados dentro do repositório da UFG.</p> <p>Definições de questões de preservação e curadoria dos dados armazenados na referida coleção;</p>
<p>Implantar uma coleção de dados no repositório de dados da UFG para um periódico científico do Portal de Periódicos da UFG</p>	<p>Pesquisa-ação</p>	<p>Criação da coleção de dados do Boletim Goiano de Geografia dentro do repositório de dados abertos da UFG;</p> <p>Definição, junto aos editores do referido periódico, dos metadados a serem requeridos aos autores e as licenças aplicáveis à coleção;</p> <p>Elaboração de um guia de depósito de dados na coleção, a ser utilizado pela equipe editorial, após os autores enviarem, por outro canal, os conjuntos de dados a serem depositados.</p>

Fonte: elaborado pela autora.

Sobre os métodos de pesquisa, escolhidos para a presente investigação, é útil justificar as escolhas. Segundo Gil (2008, p. 27) “pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar visão geral, de tipo aproximativo, acerca de determinado fato”. Essa conceituação ratifica a escolha por esse tipo de pesquisa, uma vez que na primeira fase do estudo buscou-se traçar um panorama da Ciência Aberta no Brasil e em outros países do mundo, no contexto dos periódicos científicos geridos por IES, ademais, a partir dos dados obtidos por meio do levantamento documental, foi possível apresentar as especificidades de diferentes normativas de abertura de dados de periódicos científicos.

No tocante à pesquisa-ação foi necessário selecionar o periódico científico da UFG a ser utilizado como piloto dessa instituição. Primeiramente, verificou-se por meio de consulta ao Portal de Periódicos da UFG¹⁵ que o Instituto de Estudos Socioambientais (IESA) é a unidade da UFG que possui a maior quantidade de periódicos científicos, somando-se quatro, a saber: Boletim Goiano de Geografia, Ateliê Geográfico, Revista Terceiro Incluído e Revista Signos Geográficos.

Considerando que o objetivo do estudo é traçar um conjunto de orientações que sirvam para a criação de coleções de dados abertos que poderão servir de iniciativa piloto para a UFG e outras instituições, espera-se que a abertura de dados científicos, oriundos de pesquisas publicadas em periódicos, estenda-se, primeiramente, dentro da própria unidade e depois para as demais.

¹⁵ www.revistas.ufg.br

Desse modo, dos quatro periódicos do IESA, selecionou-se o Boletim Goiano de Geografia (BGG), uma vez que é o periódico mais antigo da unidade, o mais bem avaliado no último Qualis CAPES (2017-2020), com avaliação A2, e publica uma diversidade de resultados de pesquisas científicas, que geram tanto dados quantitativos, quanto qualitativos, o que abre a possibilidade de discussões, sobretudo, sobre o povoamento dos repositórios de dados científicos.

Além disso, a autora, como servidora da instituição, está lotada no IESA e, a realização desse estudo, gera qualificação administrativa para essa unidade acadêmica, favorecendo o cumprimento do objetivo do Programa Qualificar.

Assim, estão previstas ações participativas junto à equipe editorial do Boletim Goiano de Geografia (BGG), caracterizadas pelo método da pesquisa-ação. Pretende-se, num primeiro momento, realizar uma reunião de diagnóstico, com vistas a entender seus posicionamentos e entendimentos sobre as vertentes da Ciência Aberta correlatas à comunicação científica e esclarecer sobre aspectos importantes que minimizem as resistências, como a questão da autoria sobre os dados.

Após essa investigação inicial, conjuntamente à equipe editorial, pretende-se partir para a elaboração das normativas de abertura de dados e criação da coleção dentro do repositório da UFG e, posteriormente, propor a elaboração de peças de divulgação sobre Ciência Aberta e abertura de dados no contexto do periódico, já que, inicialmente o depósito dos dados oriundos dos artigos científicos aceitos para publicação no BGG não será obrigatório, somente incentivado.

Gil (2008) afirma que quando pesquisadores e pesquisados estão envolvidos em todo processo da pesquisa, têm-se como possibilidades a pesquisa-ação e a pesquisa participante. Convém explicitar os motivos que levaram a escolha de uma em detrimento da outra. Thiollent (1986) conceitua pesquisa-ação como:

[...] um tipo de pesquisa social com base empírica que é concebida e realizada em estreita associação com uma ação ou com a resolução de um problema coletivo e no qual os pesquisadores e os participantes representativos da situação ou do problema estão envolvidos de modo cooperativo ou participativo (Thiollent, 1986, p. 14).

Em comparação com a pesquisa-ação, na pesquisa participante, pesquisadores e pesquisados estão intimamente ligados para transformar algum aspecto social daquele grupo ou comunidade, como aponta Brandão (2006):

[...] na pesquisa participante pretende-se criar alternativas populares de transformação das estruturas sociais que tornam tal “vida” exigente de ser sempre “melhorada” [...] Reconhece-se a pesquisa participante como uma alternativa confiável de abertura da investigação científica à participação popular, na criação e

no desenvolvimento de programas de ação social e educacional (Brandão, 2006, p. 11-12).

Brandão (2006) indica ainda que as experiências de pesquisa participante têm se dado preferencialmente junto a grupos ou movimentos sociais/populares. Além disso, Le Boterf (1984) *apud* Felcher, Ferreira, Folmer (2017) afirma que na pesquisa participante a população está envolvida desde a identificação dos problemas, de sua análise até a busca de soluções adequadas e apresentam um quadro comparativo dos dois tipos de pesquisa.

Quadro 9 - Diferenças entre pesquisa-ação e pesquisa participante

Pesquisa-ação	Pesquisa participante
Toda pesquisa-ação é do tipo participativa;	Tudo que é chamado de pesquisa participante não é pesquisa-ação;
O pesquisador não é pesquisado;	Cada um dos envolvidos é pesquisador e pesquisado ao mesmo tempo; Aspira-se a uma comunicação o mais possível horizontal entre todos os participantes;
Supõe uma forma de ação planejada de caráter social, educacional, técnico ou outro;	Ações planejadas nem sempre se encontra em propostas de pesquisa participante;
O pesquisador TEM uma ação destinada a resolver o problema em questão;	O pesquisador NÃO tem uma ação destinada a resolver um problema;
O pesquisador é que se apropria mais intensamente dos dados;	Utiliza o diálogo como meio de comunicação mais importante no processo conjunto de estudo e coleta de informação;
Pesquisador deve ter um alto grau de análise, de moderação, de interpretação e de animação, dominar técnicas de dinâmicas de grupo.	As metas e o desenvolvimento do projeto não são previamente determinados, mas que se elaborem com a intervenção de todos os participantes.

Fonte: (FELCHER; FERREIRA; FOLMER, 2017, p. 7).

Com base no Quadro 9 é possível inferir que toda pesquisa-ação é também participante, mas o inverso não se verifica. Ademais, o pesquisador embora em contínuo contato e colaboração com os sujeitos da pesquisa, não compõe esse último grupo, ao passo que é responsável por planejar as ações a serem executadas conjuntamente com os sujeitos da pesquisa visando solucionar o problema posto inicialmente.

Feita essa caracterização metodológica, será apresentado a seguir o detalhamento dos procedimentos usados no levantamento dos dados.

3.1 COLETA DOS DADOS

Para melhor esclarecer como os dados para esse estudo foram obtidos, retornar-se-á às etapas da pesquisa já listadas no Quadro 8.

1 - Pesquisa bibliográfica sobre as principais temáticas correlatas ao estudo: comunicação científica, periódico científico, Ciência Aberta, dados científicos e repositório de

dados científicos. Esta etapa permitiu mapear a bibliografia mais recente e relevante sobre as temáticas correlatas ao assunto e que foi utilizada para a construção do referencial teórico.

2 - Pesquisa documental em páginas *web* de periódicos científicos, brasileiros e internacionais, que já possuem coleções de dados científicos, a fim de identificar e caracterizar suas diferentes normativas de abertura de dados. Os periódicos científicos que já orientam os autores a disponibilizarem, além do artigo final com os resultados das pesquisas, também, os dados científicos utilizados para se alcançarem tais resultados, geralmente, recorrem a repositórios institucionais ou repositórios multidisciplinares, porém em seus websites já informam aos potenciais autores sobre tais políticas de disponibilização dos dados científicos.

Para fins deste estudo, foram analisadas as normativas de abertura de dados dos periódicos científicos, com algum tipo de vínculo com IES, e que já utilizam o repositório *Dataverse* da Universidade de *Harvard*. O recorte do estudo, que considera este repositório, em especial, justifica-se pelo fato de que o repositório de dados abertos da UFG foi implantado utilizando a mesma infraestrutura de *software* do *Dataverse* de *Harvard*, facilitando as análises do ponto de vista funcional da solução tecnológica.

3 – Etapas de pesquisa-ação junto aos editores do BGG com o subsídio teórico de Thiollent (1986) que chama de “Seminário” o processo de interação entre pesquisadores e pesquisados na busca pela solução do problema encontrado. Considerando que o problema aqui posto é a criação da coleção de dados do referido periódico, contemplando as definições das normativas de abertura de dados, povoamento e uso da coleção, e de acordo com as principais tarefas a serem realizadas pelo Seminário (Thiollent, 1986) convém explicá-las à luz do que foi realizado durante o estudo.

1 – Definir o tema: nesse caso, o recorte do estudo foi definido anteriormente à formação do Seminário, cujas justificativas foram apresentadas na seção da Introdução;

2 – Elaborar o problema e hipóteses: sob o viés da Ciência Aberta, quais as normativas necessárias para a abertura de dados por parte dos periódicos científicos da UFG?

3 – Constituir os membros do Seminário: sendo formado pela autora, sua orientadora e membros da equipe editorial do BGG.

4 – Centralizar as informações: possível a partir do diagnóstico de periódicos científicos de diferentes nacionalidades que já possuem coleções de dados no repositório *Dataverse* da Universidade de *Harvard* e, por conseguinte, apresentação de suas respectivas normativas de abertura de dados.

5 – Elaborar as interpretações: apresenta-se a criação da coleção do BGG no repositório de dados da UFG, com definições de povoamento, curadoria e uso dos dados científicos, seguindo as orientações da certificação *Core Trust Seal*.

6 – Buscar soluções e definir diretrizes de ação: conjuntamente aos membros da equipe editorial do BGG elaboraram-se modelos de textos sobre a abertura de dados para a página *web* do BGG e para a apresentação da coleção no repositório. Além disso, foram criados modelos de declaração de disponibilidade de dados e formulário de metadados a serem preenchidos pelos autores; orientações de preparo e envio de dados, também direcionada aos autores; e guia de depósito de dados para orientar a equipe editorial no povoamento da coleção, já que inicialmente a coleção não será auto depósito, ou seja, não serão os próprios autores que depositarão os dados.

7 – Divulgar os resultados: também com a aprovação dos membros da equipe editorial do BGG foram elaborados alguns cartazes para página *web* do BGG e redes sociais para formação dos autores e leitores sobre os temas Ciência Aberta e abertura de dados científicos e para divulgar a existência da coleção e estimular uma cultura de disponibilização dos dados associados aos artigos científicos.

Cabe ressaltar que as tarefas de acompanhar e avaliar as ações não foram possíveis de serem realizadas durante o estudo. Primeiramente porque a política do repositório de dados científicos da UFG ainda está em fase de elaboração, portanto o estudo apresentado é uma proposta piloto e não pôde ser validada. Além disso, necessita-se de um tempo maior do que o previsto para uma pesquisa de mestrado para uso contínuo da coleção de dados, seu monitoramento, avaliação e validação. Espera-se que com a criação do Centro de Ciência Aberta da UFG, o estudo seja continuado e fortalecido.

Logo, os principais dados obtidos para a construção deste estudo foram bibliográficos, documentais e constructos advindos da pesquisa-ação. Na subseção a seguir será detalhada a forma como foram analisados com vistas a atender aos objetivos do estudo.

3.2 ANÁLISE DOS DADOS

Primeiramente, por meio da pesquisa bibliográfica, pela qual se traçou um panorama das publicações mais recentes sobre as temáticas correlatas ao estudo, procedeu-se com análise de conteúdo das informações contidas nos títulos, resumos, palavras-chaves e referências das publicações encontradas.

Segundo Chizzotti (2006), a análise de conteúdo é uma técnica capaz de extrair o significado de qualquer comunicação contida num texto. O autor afirma ainda que:

[...] consiste em relacionar a frequência da citação de alguns temas, palavras ou ideias em um texto para medir o peso relativo atribuído a um determinado assunto pelo seu autor [...] socorrendo-se da quantificação das unidades do texto claramente definidas, para gerar resultados quantificáveis ou estabelecer a frequência estatística das unidades de significado (Chizzotti, 2006, p. 114).

Nesse contexto, poderia se questionar se a análise pautada na frequência em que ocorrem determinados termos no manuscrito analisado, não incorreria em análises meramente quantitativas. Contrária a esse argumento, Bardin (2011) defende que a função ou objetivo da análise de conteúdo é a inferência. Assim, “a partir dos resultados da análise, se pode regressar às causas, ou até descer aos efeitos das características das comunicações” (Bardin, 2011, p. 27).

Assim, os dados obtidos através dos títulos, resumos, palavras-chaves e referências das produções bibliográficas selecionadas e que versem sobre Ciência Aberta, comunicação científica e dados abertos foram interpretados para além da quantificação dos termos, mas sim, analisados, sobretudo, com base no contexto, finalidade, abordagem em que essas pesquisas foram produzidas e se foi considerada a questão da comunicação científica por meio dos periódicos científicos. Por esse motivo, foi importante a leitura dos resumos desses manuscritos científicos. O resultado desse levantamento compôs o item de justificativa na seção da Introdução e contribuiu, também, para a elaboração da fundamentação teórica (seção 2 deste trabalho), principalmente os referenciais indicados nos trabalhos.

Sobre o levantamento das características das normativas de abertura de dados científicos de periódicos científicos que possuem coleção de dados no repositório *Dataverse* da Universidade de *Harvard* teve-se o intuito de inferir, ou ainda, propor pontos imprescindíveis a serem considerados nas normativas do BGG, com base nos requisitos de repositórios de dados confiáveis e sustentáveis, certificação *Core Trust Seal*. A *Core Trust Seal* (CTS) é uma organização internacional, voltada para as comunidades em geral, não governamental e sem fins lucrativos que promove infraestruturas de dados sustentáveis e oferece uma certificação de nível básico, conforme o que consideram requisitos necessários para que um repositório de dados seja tido como confiável (*Core Trust Seal*, 2022a).

Como se trata de uma autoavaliação, realizada pela equipe gestora dos repositórios a serem certificados, baseada na norma TRAC - ISO 16363¹⁶, tal certificação é capaz de

¹⁶ <https://public.ccsds.org/pubs/652x0m1.pdf>

orientar a elaboração das normativas de acesso, uso e curadoria de repositório de dados científicos, assim como das normativas de abertura de dados do BGG.

Portanto, foram consideradas as orientações constantes no documento Requisitos de Repositórios de Dados Confiáveis *Core Trust Seal: Orientação Estendida 2023–2025 (Core Trust Seal, 2022b)*. A certificação avalia o repositório com base em requisitos que são agrupados em três vertentes: Infraestrutura Organizacional, Gerenciamento de Objetos Digitais e Informações de Segurança e Tecnologia (Quadro 10).

Na autoavaliação para obter a certificação, atribui-se um dos seguintes pesos, para cada um dos requisitos *Core Trust Seal*:

- 0 – Não se aplica;
- 1 – O repositório ainda não considerou este requisito;
- 2 – O repositório possui um conceito teórico sobre este requisito;
- 3 – O repositório está em fase de implementação deste requisito;
- 4 – Este requisito foi totalmente implementado no repositório.

Segundo o documento, repositórios que possuem níveis de conformidade 1 ou 2 se mostram com implementação insuficiente para a certificação de qualidade referida. No entanto, é possível concedê-la caso o repositório apresente alguns requisitos qualificados como estando em fase de implementação (3).

O Quadro 10 apresenta os requisitos requeridos pela certificação, com base nas três vertentes.

Quadro 10 – Requisitos *Core Trust Seal*

Vertente 1: Infraestrutura Organizacional		
R0.	Informações básicas e contexto	O repositório apresenta sua breve descrição com as seguintes informações: política institucional (missão, escopo, relação com outros repositórios, licenças utilizadas), tipo do repositório, comunidade alvo (e o papel que o repositório exerce nessa comunidade), nível de curadoria, parceiros internos e externos, estatísticas de citação dos <i>datasets</i>
R1.	Missão/Escopo	O repositório tem a missão explícita de fornecer acesso e preservar os dados em seu domínio. Para tanto, deve incluir links para declarações desta missão.
R2.	Gerenciamento de direitos	O repositório mantém todas as licenças aplicáveis que abrangem o acesso e uso de dados e monitora a conformidade, sendo que os depositantes tem liberdade de escolha dentre as licenças disponíveis.
R3.	Continuidade de Acesso	O repositório tem um plano de continuidade para garantir o acesso contínuo e a preservação do seu acervo (Governança), o que deve incluir um orçamento pré-definido, os níveis de responsabilidade da equipe gestora e a gestão de riscos, como possíveis respostas a eventuais mudanças bruscas (perda de financiamento ou mudança nos interesses institucionais).
R4.	Confidencialidade/Ética	O repositório garante, na medida do possível, que os dados sejam criados, curados, acessados e usados em conformidade com as normas disciplinares e éticas. Devem ser explicitado se o repositório solicita a confirmação de que a coleta ou criação de dados foi realizada de acordo com os critérios legais e éticos e quais os procedimentos para gerenciar

		dados sensíveis. Observação: requisito deve estar alinhado com R2.
R5.	Governança e recursos	O repositório tem financiamento adequado, um número suficiente de pessoal qualificado para o tratamento e curadoria dos dados e clareza quanto à governança, responsabilidades e tomada de decisões.
R6.	Orientação de especialistas	O repositório adota mecanismo(s) para garantir orientação e feedback contínuos de consultores internos ou um comitê consultivo externo que pode ser preenchido com especialistas técnicos, de curadoria, ciência de dados e disciplinares, bem como, como ocorre a comunicação com tais especialistas.
Vertente 2: Gerenciamento de objetos digitais		
R7.	Proveniência e autenticidade	O repositório garante a integridade e autenticidade dos dados. Para tanto, deve garantir que quaisquer alterações intencionais nos dados e metadados devem ser documentadas, incluindo a justificativa e o originador da alteração (controle de versões). Devem ser tomadas medidas para garantir que alterações não intencionais ou não autorizadas possam ser detectadas e que versões corretas de dados e metadados sejam recuperadas.
R8.	Depósito e Avaliação	O repositório aceita dados e metadados com base em critérios definidos para garantir relevância e compreensão para os usuários dos dados, devendo ser explicitados se existem procedimentos em vigor que garantam que apenas os dados apropriados à política de coleta sejam aceitos. A equipe do repositório deve ter todas as informações, procedimentos e conhecimentos especializados, necessários para garantir a preservação e o uso em longo prazo.
R9.	Plano de preservação	O repositório define as responsabilidades pela preservação em longo prazo (repositório, depositantes e usuários) e gere esta função de forma planejada e documentada. É preciso definir formatos de arquivos e esquemas de metadados para tal preservação. Há também necessidade de planejar migrações futuras e monitorar a obsolescência e deterioração das mídias de armazenamento dos dados.
R10.	Garantia de qualidade	O repositório trata adequadamente a qualidade de dados técnicos e metadados e garante que informações suficientes estejam disponíveis para que os usuários finais façam avaliações relacionadas à qualidade, ou seja, devem haver documentos que esclareçam sobre o workflow de curadoria de dados, níveis de segurança, integridade e inteligibilidade dos dados. Esse requisito pode ser atendido por meio da revisão dos <i>datasets</i> antes e depois da publicação, da exigência do preenchimento de campos de metadados, a partir de padrões de metadados que podem ser adotados, ou ainda, pela conversão de formatos de <i>datasets</i> .
R11.	Workflows	O gerenciamento de objetos digitais ocorre de acordo com fluxos de trabalho definidos (e seus níveis de segurança), desde a ingestão até a disseminação por meio de documentação e comunicação clara aos depositantes e usuários sobre o manuseio (tratamento) dos dados e eventuais mudanças de fluxos de trabalho.
R12.	Descoberta e identificação de dados	O repositório permite que os usuários descubram os dados e os consultem de forma persistente por meio da citação adequada. Os <i>datasets</i> devem possuir identificadores de objetos digitais, como o DOI. E os repositórios devem estar indexados em diretórios, como o Re3data ¹⁷ .
R13.	Reutilização dos dados	O repositório permite a reutilização dos dados ao longo do tempo, garantindo que os metadados apropriados estejam disponíveis para apoiar a compreensão e o uso dos dados, sendo necessário um conhecimento profundo dos cenários de reutilização e das necessidades da comunidade alvo, em termos de suas práticas, ambiente técnico e (adesão) aos padrões aplicáveis. Avaliação e reavaliação dos formatos dos dados também se faz necessário, além do estabelecimento de padrões de metadados.
Vertente 3: Tecnologia		
R14.	Armazenamento e integridade	O repositório aplica processos documentados para garantir o armazenamento de cópias arquivísticas, dados e metadados. O que inclui checagens dos locais de armazenamento de dados e estratégias para evitar

¹⁷ <https://www.re3data.org/>

		as redundâncias, ou para que haja consistência entre as cópias existentes. Para cada local de armazenamento devem ser implementadas medidas para garantir que mudanças não autorizadas possam ser detectadas e versões corretas dos dados e metadados possam ser recuperadas.
R15.	Infraestrutura técnica	O repositório está hospedado em uma infraestrutura adequada de hardware e software com disponibilidade, largura de banda e conectividade que estejam em consonância com a comunidade alvo. O software está bem documentado e há um plano de continuidade de negócios.
R16.	Segurança	A infraestrutura técnica do repositório fornece proteção da instalação e seus dados, produtos, serviços e usuários. Existem políticas de segurança em vigor para controlar a segurança de todos os sistemas, incluindo segurança de rede, verificações de intrusão, segurança de instalações físicas e procedimentos de autenticação e autorização para gerenciar o acesso ao sistema.

Fonte: *Core Trust Seal* (2022b) organizado pela autora (tradução nossa).

A partir da análise dos dados coletados durante as pesquisas bibliográfica, documental e pesquisa-ação, apresentadas anteriormente, foi possível desenvolver o referencial teórico do estudo; elaborar normativas de acesso e uso da coleção de dados do BGG; definir questões de preservação e curadoria dos dados armazenados na referida coleção; implantar a coleção; criar peças de divulgação sobre os temas tratados no estudo e encorajar a abertura de dados científicos associados a artigos. Tais resultados encontram-se detalhados na próxima seção.

4 RESULTADOS

Os resultados apresentados nesta seção referem-se à pesquisa documental em páginas *web* de periódicos científicos, brasileiros e internacionais, que já possuem coleções em um repositório de dados científicos, a fim de identificar e analisar suas diferentes políticas e fluxos de abertura de dados científicos. Em um segundo momento apresentam-se as orientações, elaboradas junto aos editores do BGG, para criação de coleções de dados abertos, caracterizando as normativas de abertura de dados, povoamento e uso.

A pesquisa documental foi realizada a partir de levantamento de periódicos vinculados a IES que utilizam o repositório de dados *Dataverse* da Universidade de *Harvard*¹⁸. A seleção deste repositório se justifica pelo fato de ser gratuito, aberto a qualquer pesquisador, interno ou externo à Universidade de *Harvard*. Pesquisadores individuais, periódicos ou organizações podem criar coleções personalizáveis dentro dele para organizar, gerenciar e exibir conjuntos de dados (aqui, também chamados de *datasets*). Além disso, diferentes instituições mundiais, assim como optado pela UFG, podem utilizar o *software* livre do Projeto *Dataverse* de *Harvard* para criarem seus repositórios, utilizando os próprios recursos técnicos e tecnológicos para essa instalação. Portanto, a proposta de coleção de dados apresentada neste estudo será elaborada com base nos recursos e funcionalidades do levantamento documental.

Ademais, Rodrigues, Dias e Lourenço (2022) conduziram um estudo sobre a gestão e curadoria em repositórios de dados científicos da América do Sul apresentando os softwares utilizados: *Morpho*, *DSpace*, ou *Dataverse*. A partir das análises pretendidas no estudo chegaram à conclusão de que “os repositórios em maior conformidade com os princípios FAIR foram aqueles estabelecidos mediante o uso do *Dataverse*” (Rodrigues; Dias; Lourenço, 2022).

Portanto, na subseção a seguir teve-se como principal objetivo apresentar as principais características das normativas e definições de abertura de dados dos periódicos do levantamento documental.

¹⁸ <https://dataverse.harvard.edu/>.

4.1 ANÁLISE DESCRITIVA DOS PERIÓDICOS CIENTÍFICOS QUE POSSUEM COLEÇÕES DE DADOS NO REPOSITÓRIO DATAVERSE DA UNIVERSIDADE DE HARVARD

Nesta subseção serão apresentadas as principais características das coleções de dados dos periódicos científicos de diferentes nacionalidades armazenadas no repositório de dados *Dataverse* da Universidade de *Harvard*.

Ao compartilhar um *dataset* no *Dataverse de Harvard* é possível definir as licenças para uso e reutilização, alinhadas ao processo de abertura do fazer científico, conforme visto na seção 2.5.2 deste estudo. Quando se utiliza o repositório *Dataverse*, os *datasets* obtêm automaticamente uma citação padrão com um DOI, ao serem publicados, e seus metadados se tornam disponíveis podendo ser encontrados por meio de mecanismos de pesquisa, mesmo quando os dados são restritos (*Harvard Dataverse, 2022*).

O *Dataverse* de *Harvard* oferece orientações específicas para pesquisadores, periódicos e organizações. Nesse estudo ater-se-á aos elementos conceituais e tecnológicos necessários para sua utilização por periódicos científicos de maneira geral. Nesse escopo, é possível estabelecer vínculo por meio de identificadores permanentes, gerenciar a submissão, revisão e publicação de dados associados aos estudos publicados nos periódicos.

Além disso, ao criar uma coleção dentro do *Dataverse de Harvard* o periódico científico pode (e deveria) associá-la à sua página *web*, é o que se chama bidirecionalidade de acesso (*Harvard Dataverse, 2022*).

Segundo as orientações do repositório, existem, basicamente, quatro maneiras pelas quais os periódicos podem utilizar a infraestrutura do *Dataverse* de *Harvard*:

1 – O periódico cria uma coleção dentro do repositório e, para cada artigo aceito, os editores instruem os autores para que enviem seus dados para essa coleção no *Dataverse*, ou para o próprio editor, que faz o depósito. Quando é o autor quem realiza o depósito dos dados, o editor do periódico é informado por *e-mail* para que possa revisar e publicar os *datasets*. Portanto, nesse cenário, a curadoria é de responsabilidade dos editores do periódico ou uma equipe a ele associada;

2 – O periódico cria uma coleção dentro do repositório e contrata o serviço de curadoria de terceiros. Destaca-se que a equipe de curadoria da Universidade de *Harvard* oferece este tipo de serviço, mediante pagamento de taxas;

3 – O periódico que utiliza o *Open Journal Systems* (OJS) para gerenciamento do seu fluxo editorial, pode integrar seu módulo de submissões ao *Dataverse*, sendo necessário criar uma coleção dentro do *Dataverse* e, posteriormente, no OJS, instalar um *plugin* para

estabelecer uma conexão com o repositório. Um *plugin* é um módulo de extensão a ser instalado dentro de um *software* para adicionar funções extras, que o *software* original não possuía. Com a instalação deste *plugin* é possível requerer o *upload* dos dados científicos para a coleção do *Dataverse* conjuntamente à submissão do artigo;

4 – O periódico pode recomendar algum repositório de dados científicos aos autores, em seu próprio *site*, incluindo o *Dataverse* de *Harvard*. Assim, os autores depositam seus *datasets* e incluem a citação do DOI do conjunto de dados na submissão do seu artigo ou após o aceite, de acordo com a política do periódico.

Para este estudo, realizou-se um levantamento dos periódicos científicos que utilizam o *Dataverse* de *Harvard* para hospedar suas coleções.

Desse modo, para identificar as principais características das políticas adotadas por esses periódicos, se fez necessário estabelecer parâmetros de melhores práticas dentro de um recorte de normativas a serem observadas. Assim, foram considerados os requisitos de repositórios de dados confiáveis da certificação ofertada pelo selo *Core Trust Seal*.

Com base nesses requisitos foi elaborada a Tabela 2 para analisar as políticas de abertura de dados dos periódicos científicos com coleções dentro do *Dataverse* de *Harvard*. Esclarece-se que foi necessário elaborar tal tabela porque não é possível analisar todos os requisitos do Quadro 10 (p. 78) somente com base no levantamento, uma vez que alguns se referem a decisões internas do periódico ou da ferramenta tecnológica do *Dataverse*.

Detalhadamente, o levantamento ocorreu da seguinte maneira: no dia 01 de setembro de 2022, acessou-se a página *web* do repositório¹⁹ e, na opção “*Dataverse Category*”, selecionou-se “*Journals*”, expressão em língua inglesa (língua oficial do repositório) para a palavra “Periódicos”. Foram encontradas 124 coleções. Destas, 23 não continham nenhum conjunto de dados depositado e não foram consideradas no estudo.

Das 101 restantes, foram acessadas as páginas das coleções e dos periódicos associados numa tentativa de identificar se estavam vinculados a alguma IES. Cabe ressaltar que essa informação nem sempre aparece de maneira muito clara na página da coleção, nem tampouco na página *web* do periódico. Sendo assim, os periódicos selecionados para análise são aqueles que fazem menção a parcerias com IES, são sediadas ou hospedadas em portais dessas instituições, ou ainda, são de editoras universitárias, como a *Cambridge University Press*, dos Estados Unidos da América, que hospeda 14 periódicos que compõem a amostra da análise.

¹⁹ <https://dataverse.harvard.edu/>

Foram selecionados, conforme estes critérios, 65 periódicos. Os outros 36 são de editoras privadas, associações ou institutos de pesquisa independentes. Passou-se então para análise individual das 65 coleções, bem como dos sites dos respectivos periódicos.

Com essa análise algumas coleções foram retiradas da amostra, a saber:

- Dois periódicos possuíam duas coleções cada, sendo que, para a análise, considerou-se apenas uma coleção de cada periódico;
- Uma coleção não foi considerada na análise porque o respectivo periódico mudou de nome. Com a nova nomenclatura, foi criada outra coleção, e esta sim compõe a amostra da análise deste estudo;
- Outras três coleções foram, equivocadamente, categorizadas como sendo de “*Journals*”, porém são coleções de dados provenientes de pesquisas científicas específicas;
- Um pesquisador indiano criou uma coleção como sendo *Journal*, no entanto são conjuntos de dados associados às suas produções pessoais;
- A coleção de um periódico chinês, com dois *datasets*, não apresenta um texto indicativo de sua finalidade, nem link para o periódico associado e, também, não foi possível localizá-lo por meio de buscas na *internet*, sendo descartada da amostra.

Portanto, a amostra correspondeu a 57 periódicos científicos, com coleções de dados criadas dentro do *Dataverse* de *Harvard* e com as seguintes nacionalidades:

- 39 dos EUA;
- 04 do Brasil;
- 03 da Holanda;
- 02 da Coreia do Sul;
- 02 do Canadá;
- 01 da Alemanha;
- 01 da França;
- 01 da Itália;
- 01 do Japão;
- 01 da China;
- 01 do Reino Unido; e
- 01 da Turquia.

As análises realizadas buscaram identificar as informações e características descritas na Tabela 2 e que constavam nas descrições das coleções no *Dataverse* de *Harvard*, nas páginas *web* dos periódicos, ou em ambas. Como mencionado, os elementos da tabela foram elaborados com base nos requisitos do selo *Core Trust Seal*, apresentados no Quadro 8, com adaptações às especificidades dos periódicos científicos.

Tabela 2 – Elementos observados nas políticas dos periódicos científicos selecionados

Requisito associado	Requisito analisado	Possui	Não Possui
R0	Finalidade/Apresentação da coleção de dados no Dataverse	51	6
R0	Comunidade atendida/Parceiros	57	-
R0	Vínculo com alguma rede, consórcio, comunidade de prática, etc.	57	-
R0	Recursos Adicionais (estatísticas / fator de impacto, glossário etc.)	57	-
R1, R2 e R8	Política de acesso e uso	57	-
R5	Governança da coleção	57	-
R9 e R14	Solicitação de envio do termo de depósito de dados	19	38
R10 e R12	Bidirecionalidade: Coleção/Periódico	43	14
R11	Guia de preparação de depósito de dados	34	23
R11	Informações sobre quem realiza o depósito dos dados	49	8
R11	Auto depósito	45	12
R11	Workflow	41	16
R12	Orientações sobre citação do <i>dataset</i>	57	-
R13	Informações sobre reprodutibilidade e replicabilidade	57	-

Fonte: elaborado pela autora com base nos Requisitos *Core Trust Seal* (2022b).

A partir da análise quantitativa apresentada na Tabela 2, é possível fazer alguns apontamentos acerca das características identificadas:

a) A finalidade da coleção de seis periódicos não foi apresentada no *Dataverse*, sendo necessário recorrer à página *web* do periódico, onde foi possível compreender a Missão/Esopo do periódico e, por conseguinte, dos *datasets* depositados na coleção do repositório;

b) As informações necessárias para a análise foram encontradas nos seguintes menus (ou submenus) das páginas *web* dos periódicos: “Sobre”; “Instruções aos autores”; “Requisitos para submissão”; “Transparência de Dados”; ou “Política de compartilhamento de Dados”;

c) Quatorze repositórios/coleções não apresentam *link* bidirecional, o que demandou localizar as páginas *web* dos periódicos por meio de *sites* de busca;

d) Sobre a comunidade atendida, geralmente aparece o campo científico atendido e, por conseguinte, os pesquisadores, estudantes e demais interessados nesse campo;

e) Considerou-se que todos os periódicos trazem informações sobre governança, já que minimamente, informam a equipe responsável pelo periódico, as universidades ou outras instituições associadas e, em alguns casos, a proveniência orçamentária;

f) Todos os periódicos automaticamente integram a Rede *Dataverse*, por esse motivo é possível considerar que todos possuem vínculo com alguma rede, consórcio, comunidade de

prática. No entanto, os periódicos também trazem no menu “Sobre” de suas páginas *web*, suas afiliações com associações de pesquisa acadêmicas ou de campos científicos específicos;

g) Em relação aos recursos adicionais utilizados, o próprio repositório *Dataverse*, para cada *dataset* depositado, apresenta o recurso denominado “*Replication data for + ‘o nome do artigo ou pesquisa associado a ele’*”, além disso, informa as métricas de acesso (como número de *downloads* do conjunto de dados por outros pesquisadores) e gera, com base nos metadados informados, uma referência bibliográfica (normalizada pelo formato da *American Psychological Association* (APA)), útil para que se entenda como citar o conjunto de dados, no caso de reutilização. De modo semelhante, os periódicos, em suas páginas *web*, trazem informações estatísticas como número de visualizações, *downloads*, fator de impacto, dentre outras, referentes aos artigos científicos;

h) Para este estudo, considerou-se como “guia de depósito de dados” as informações que orientam minimamente os autores sobre o envio dos conjuntos de dados;

i) A expressão “Termo de Depósito de dados” não apareceu em nenhuma política analisada, em contrapartida, 19 periódicos solicitam uma Declaração de Disponibilidade de Dados que deve ser incluída no início ou no final do artigo e que expressa a concordância com a disponibilização dos *datasets* no *Dataverse de Harvard* para fins de replicação. Alguns periódicos exigem o depósito da declaração no *Dataverse* junto com os dados, sendo este documento um descritivo detalhado sobre o uso correto dos dados do ponto de vista técnico, tecnológico e metodológico, o que favorece a replicabilidade e reprodutibilidade. Desse modo, considerou-se Termo de Depósito de Dados e Declaração de Disponibilidade de Dados, como se referindo ao mesmo documento.

j) A respeito do depósito dos dados, a maioria (45 periódicos) opera com auto depósito, ou seja, o próprio pesquisador realiza o depósito de seus dados na coleção. Nesse caso há instruções (guia) para que os autores depositem seus *datasets* dentro da coleção de dados do periódico, no *Dataverse*. Quando ocorre um depósito, o editor é informado, para que faça a conferência e só então, a publicação. Outros oito periódicos não apresentam informações muito precisas sobre este assunto, apenas afirmam que após a aceitação do artigo, os autores serão orientados sobre a disponibilização dos dados.

k) As políticas de acesso e uso também são um recurso do *software* do repositório *Dataverse de Harvard*. Quando uma pessoa cria uma coleção é possível configurar as permissões tanto para que outros possam incluir *datasets*, quanto para uso dos conjuntos de dados disponibilizados;

l) Quanto ao *Workflow*, ou seja, o fluxo de trabalho tanto do artigo científico, quanto dos conjuntos de dados, buscou-se identificar a existência de informações sobre as ações envolvidas e em que momento os dados científicos devem ser depositados: no momento da submissão do artigo ou após a sua aceitação pelos pareceristas. Observou-se que, em geral, os depósitos são requeridos após a aceitação do artigo, e há informações de que orientações adicionais sobre essa demanda são encaminhadas diretamente aos autores, posteriormente à aceitação;

m) Por fim, algumas políticas de compartilhamento de dados são padronizadas pelas editoras, às quais os periódicos estão associados, como a editora *Cambridge University Press* (EUA).

Reconhece-se que o compartilhamento dos dados científicos, associados aos artigos, ainda é uma prática emergente no Brasil, mas também em outros países do mundo, com exceção dos EUA, que parecem estar mais adiantados, considerando apenas o levantamento realizado no *Dataverse de Harvard*, no qual se percebeu políticas mais definidas e claras nas páginas *web* dos periódicos. Obviamente, o referido repositório é proveniente de uma universidade norte-americana, mas considerando que é gratuito e aberto também à comunidade externa, ainda se observa a presença de poucos periódicos de outras nacionalidades.

De todo modo, esse levantamento se mostrou muito importante para embasar a criação das normativas de abertura de dados por periódicos na UFG. Na próxima subseção apresenta-se o delineamento elaborado a partir da pesquisa-ação junto aos editores do BGG.

4.2 PROPOSTA DE IMPLANTAÇÃO DA COLEÇÃO DE DADOS DO BOLETIM GOIANO DE GEOGRAFIA

A UFG encontra-se numa fase de definições e normativas tanto sobre a criação do Centro de Ciência Aberta (CCA), quanto da utilização do repositório de dados científicos. À frente desta demanda tem-se um Grupo de Trabalho (GT) formado por professores da Faculdade de Informação e Comunicação (FIC), servidores da Pró Reitoria de Pesquisa e Inovação (PRPI), do Sistema Integrado de Bibliotecas (SIBI), do Centro de Recursos Computacionais (Cercomp) e do Centro de Informação, Documentação e Arquivo (Cidarq).

As minutas de criação do CCA preveem objetivos que se alinham aos recomendados pela UNESCO (2022), como a alfabetização da comunidade acadêmica sobre o tema da Ciência Aberta e, posteriormente, a promoção de ações nos diferentes estágios do processo científico.

Na subseção anterior foram apresentadas as características das normativas de abertura de dados científicos de periódicos que utilizam o repositório *Dataverse* da Universidade de *Harvard*. Nesse ínterim, cabe esclarecer que a UFG, instalando localmente o mesmo *software* do *Dataverse* e procedendo com as personalizações possíveis e desejáveis, criou seu próprio repositório de dados²⁰.

Nesta seção o estudo focou na utilização do repositório de dados da UFG por um periódico científico desta instituição, o Boletim Goiano de Geografia (BGG). O BGG é o quinto periódico mais antigo da UFG, tendo lançado seu primeiro volume impresso em 1981. Em 2012 passou a publicar seus volumes bianuais somente no formato digital e em 2013 passou para a periodicidade quadrimestral. Já em 2019 adotou-se o fluxo contínuo. Seu foco são artigos que abrangem a Geografia e áreas afins, frutos de pesquisas tanto quantitativas, quanto qualitativas, o que permitiu refletir sobre abertura de dados científicos nesses dois contextos.

Para tanto, fez-se necessário definir as normativas e configurações do ponto de vista tecnológico que irão pautar o processo de abertura dos dados científicos associados aos artigos publicados no BGG, além de elencar questões de preservação e curadoria dos dados armazenados.

Tais definições foram norteadas pelos requisitos *Core Trust Seal* (2022b) descritos no Quadro 10 (p. 78) e pelas orientações do GT de Ciência Aberta da UFG que prevê que a implantação das coleções de dados do repositório devem considerar as seguintes vertentes:

- 1 - Estrutura organizacional e regulamentar;
- 2 - Estrutura técnica e tecnológica;
- 3 - Estrutura formativa e de divulgação.

A partir de tais vertentes estabeleceram-se ações junto à equipe editorial do BGG, Quadro 11.

²⁰ <https://dadosdepesquisa.ufg.br/>

Quadro 11 – Ações para implantação de uma coleção de dados de periódico científico no repositório da UFG

	AÇÕES / DEFINIÇÕES PRELIMINARES		AÇÕES DE FUNCIONAMENTO	
	INSTITUCIONAIS (REPOSITÓRIO)	PERIÓDICO	INSTITUCIONAIS (REPOSITÓRIO)	PERIÓDICO
Estrutura Organizacional e Regulamentar	<ol style="list-style-type: none"> 1. Aprovação da criação da coleção da revista no repositório; 2. Definição das licenças aplicáveis aos dados no repositório; 3. Definição dos tipos de dados a serem depositados no repositório (formatos); 4. Elaboração do termo de acesso e uso dos dados no repositório; 5. Exigência do depósito dos termos de aprovação em comitês de ética, quando aplicável. 	<ol style="list-style-type: none"> 1. Definição da missão da coleção do periódico; 2. Elaboração do modelo de Declaração de Disponibilidade de Dados em conformidade com o termo de acesso e uso do repositório, regulações gerais referentes às licenças de uso e ética em pesquisa (para os autores). 	<ol style="list-style-type: none"> 1. Acompanhamento do acesso e uso do repositório e cumprimento das políticas e definições previamente estabelecidas; 	<ol style="list-style-type: none"> 1. Acompanhamento do acesso e uso da coleção (estatísticas);
Estrutura Técnica e Tecnológica	<ol style="list-style-type: none"> 1. Criação do repositório institucional de dados científicos; 2. Elaboração de um formulário padrão (template) de preenchimento de metadados bibliográficos do repositório. 	<ol style="list-style-type: none"> 1. Criação e configuração da coleção de dados no repositório de dados científicos da UFG conforme recursos disponíveis; 2. Definição dos membros da equipe editorial responsáveis pela coleção; 3. Criação de um fluxo de trabalho para a preparação e envio dos dados por parte dos autores; 4. Criação de um fluxo de trabalho para o depósito dos dados na coleção por parte da equipe; 5. Elaboração de um formulário padrão (template) de metadados a ser preenchido pelos autores, complementando o template da Universidade. 	<ol style="list-style-type: none"> 1. Suporte técnico e tecnológico referente ao funcionamento do repositório. 	<ol style="list-style-type: none"> 1. Recebimento e conferência dos conjuntos de dados e respectivos metadados enviados pelos autores; 2. Depósitos dos dados na coleção;
Estrutura Formativa e de Divulgação	<ol style="list-style-type: none"> 1. Formação da equipe editorial responsável pela coleção. 	<ol style="list-style-type: none"> 1. Elaboração de um plano de divulgação da coleção de dados do periódico. 2. Documento orientando autores sobre descrição do contexto do projeto de pesquisa que originou o artigo submetido no periódico, anonimização dos dados, dentre outros fatores a serem considerados antes de divulgar os dados. 	<ol style="list-style-type: none"> 1. Formação continuada da equipe editorial responsável pela coleção. 	<ol style="list-style-type: none"> 1. Orientações aos autores quanto ao depósito dos dados; 2. Elaboração de ações de divulgação da coleção e da política de fortalecimento da Ciência Aberta (no site do periódico, página da coleção de dados e redes sociais).

Fonte: elaborado pela autora.

A partir das ações descritas no Quadro 11 foram realizadas reuniões com a equipe do BGG para elaboração de modelos referentes às definições preliminares de criação da coleção. As ações de funcionamento, embora já previstas no referido Quadro 11, só serão efetivamente possíveis após a aprovação das instruções normativas pelo Conselho Universitário da UFG, o CONSUNI.

Portanto, as etapas da pesquisa-ação desenvolvida estiveram centradas nas ações e definições preliminares, que todos os periódicos que desejam endossar a abertura de dados científicos, por meio do repositório de dados da UFG, devem refletir. Nesse tocante, Pavão, Silva e Silveira (2020) apontam que para efetivação da abertura de dados referentes à artigos publicados em periódicos científicos, “[...] a política editorial deve contemplar orientações a respeito de onde o autor irá depositar seus dados, qual o conjunto mínimo de metadados para descrevê-los, assim como qual a indicação em relação à licença de uso e distribuição” (Pavão, Silva, Silveira, 2020, p. 169).

Nesse sentido, consideraram-se, também, as diretrizes *Transparency and Openness Promotion* (diretrizes TOP) que têm como intuito promover a transparência e abertura nas políticas e práticas editoriais de periódicos científicos e que orientam como padrões: 1 - citação dos dados, 2 - transparência dos dados, 3, e 5 - transparência dos métodos de coleta e análise, bem como dos códigos, 4 - transparência dos materiais de pesquisa, 6 - aceite ou encorajamento de pré registro dos estudos (como em repositórios *preprints*) ou 7 - registro dos planos de análise e 8 - políticas relativas à replicação dos dados (*Center for Open Science*, 2018, tradução SciELO).

No que concerne à abertura, especificamente, dos dados científicos, as normativas elaboradas neste estudo foram embasadas pelas recomendações 1 a 5. Cabe ressaltar que sobre a de número 1 – citação dos dados, esse é um recurso da própria ferramenta de *software* do repositório de dados da UFG, já implementada pelo Projeto *Dataverse*. Ao realizar o depósito e publicação do conjunto de dados, o mesmo apresenta uma citação no padrão *American Psychological Association* (APA), gerada automaticamente conforme metadados preenchidos e, caso o repositório o disponibilize, um DOI é atribuído ao conjunto de dados.

Para as definições preliminares da coleção, considera-se que o ponto de partida, assim como prevê a Recomendação da UNESCO, é “promover um entendimento comum sobre a Ciência Aberta, sobre os benefícios e desafios associados, assim como sobre os diversos caminhos para a Ciência Aberta” (UNESCO, 2022, p. 6). Desse modo, a primeira ação foi elaborar um texto que posicione o BGG diante desse objetivo da UNESCO e que esteja visível a todos os leitores (Apêndice A).

Assim, como a página *web* do periódico traz um *link* para sua coleção de dados no repositório, esta deve ser informativa sobre seu foco e apresentar um caminho de acesso para a página *web* do periódico (bidirecionalidade). No Apêndice B há a apresentação da finalidade da coleção na página do repositório de dados da UFG.

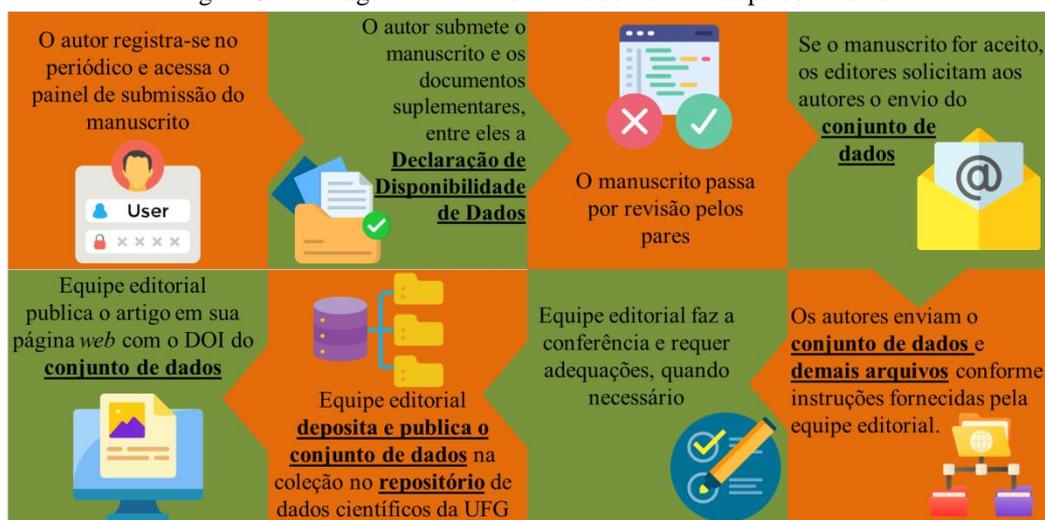
Nota-se no Apêndice B que na pesquisa-ação foi definido junto à equipe editorial que todos os autores devem enviar, no momento da submissão do artigo, o documento intitulado “Declaração de Disponibilidade de Dados”. Embora, inicialmente, o GT de Ciência Aberta da UFG, tenha sugerido o nome “Termo de Depósito de Dados”, para um documento com essa mesma finalidade, o levantamento documental apresentado na seção 4.1 evidenciou que, nas normativas de abertura de dados dos periódicos nacionais e internacionais que compõem a amostra do estudo, não aparece essa expressão.

Assim, na “Declaração de Disponibilidade de Dados” os autores já informarão se pretendem, ou não, compartilhar os dados científicos associados ao artigo que estão submetendo no BGG. O Apêndice C apresenta o documento elaborado.

O GT de Ciência Aberta da UFG já determinou que o funcionamento do repositório de dados, pelo menos nesse primeiro momento da implantação, não será auto depósito, ou seja, as coleções não serão alimentadas pelos próprios autores. Estes, por sua vez, após aprovação do manuscrito serão incentivados pela equipe editorial do periódico a enviarem o conjunto de dados. As orientações para preparação e envio destes dados estarão disponíveis na guia de “Instruções para os autores” na página *web* do periódico, conforme Apêndice D.

Em suma, a Figura 8 apresenta um resumo do fluxo de trabalho sobre o momento em que os dados são requeridos e depositados na coleção do BGG no repositório de dados da UFG.

Figura 8 – Fluxograma de envio de dados científicos pelos autores



Fonte: elaborado pela autora.

O Apêndice D faz menção a um formulário descritivo de metadados que será disponibilizado, também na página *web* do periódico, para *download* e preenchimento pelos autores. Os campos deste formulário foram definidos na pesquisa-ação, considerando as especificidades de um periódico científico (Apêndice E).

No formulário há um metadado intitulado como “Descrição dos dados”, anteriormente a autora havia utilizado outro título para esse campo, “Contexto”, sendo solicitada pela equipe editorial sua alteração. Claramente outras alterações foram sugeridas no decurso da pesquisa-ação, porém essa é uma discussão recente entre os pesquisadores que relacionam os metadados mais apropriados para identificar os conjuntos de dados e possibilitar sua reutilização.

Isto Huvila (2022) reflete sobre o alto número de conjuntos de dados mal documentados e de difícil acesso, apresentando, então, o conceito de parados, ou seja, dados que descrevem processos referentes aos dados. Por esse motivo, além de ajustar o título, editou-se sua descrição de forma que os autores forneçam informações suficientes sobre o contexto dos dados que possam ajudar a elucidar processos passados, em andamento e potenciais relacionados ao conjunto de dados.

Após os autores enviarem o conjunto de dados, conforme as instruções do Apêndice D e o formulário de metadados (Apêndice E), a equipe editorial procederá com o depósito na coleção do BGG no repositório de dados da UFG. Sabe-se que as equipes editoriais dos periódicos geridos por IES, geralmente, são formadas por uma quantidade limitada de membros, envolvidos com uma infinidade de outras atividades acadêmicas. A opção por aderir à política de compartilhamento de dados, embora uma tendência irreversível no universo científico, não pode desconsiderar a responsabilidade em promover acessibilidade de longo prazo e a compreensibilidade dos dados conforme as necessidades da comunidade a que se designa (*Core Trust Seal, 2022b*). Isso quer dizer que a equipe editorial deve eleger pessoal para realizar o depósito e curadoria dos dados. As propostas elaboradas até aqui têm o intento de que os autores forneçam a maioria das informações e arquivos que facilitem tais tarefas, por parte da equipe editorial. Afinal, os autores são os detentores do conhecimento acerca dos dados produzidos ou obtidos durante suas pesquisas.

Ainda assim, a equipe editorial realizará a curadoria em algum nível (A, B, C e D), conforme a certificação *Core Trust Seal* (2022b):

A. Conteúdo depositado na coleção como enviado pelos autores.

B. Curadoria básica - verificação breve, adição de metadados básicos ou documentação.

C. Curadoria aprimorada - conversão para novos formatos durante o depósito, aprimoramento da documentação e metadados.

D. Curadoria em nível de dados - como em C acima, mas com edição adicional de dados a serem depositados.

Nesse primeiro momento de funcionamento, inclusive pelas limitações de pessoal já mencionadas, a curadoria da coleção do periódico em questão será realizada no nível C, porém as conversões de arquivos serão básicas, por exemplo, conversão de formatos de documentos abertos em *softwares* pagos (como .xlsx), para formatos de *softwares* livres (como .ods). Está previsto pelo GT de implantação da Ciência Aberta na UFG que sejam requeridos à administração superior da instituição, a disponibilização de profissionais com formação específica na gestão de dados para apoiar as equipes editoriais e demais pesquisadores da UFG.

Quanto ao aprimoramento da documentação, ao considerar o depósito após o aceite dos manuscritos, o conjunto de dados será publicado e receberá um DOI. Este, por sua vez, será incluído na versão final do artigo a ser publicado na página *web* do periódico. Desse modo, os metadados do conjunto de dados necessitará ser editado, de forma a serem incluídas informações do artigo associado, também chamada de publicação relacionada. Os metadados adicionais são os apresentados no Quadro 12.

Quadro 12 – Metadados a serem preenchidos pela equipe editorial do BGG

Metadado	Descrição	Forma de Preenchimento
Publicação Relacionada	Citação completa do artigo a que se relaciona o conjunto de dados	Formato APA ou ABNT do artigo associado
Data de publicação do artigo	Data em que o artigo foi publicado no BGG	DD/MM/AAAA
Informações do periódico	Informações sobre edição do BGG na qual o artigo foi publicado	Volume, número, ano
DOI da publicação relacionada	Identificador único do artigo a que se relaciona o conjunto de dados	Link completo

Fonte: elaborado na pesquisa-ação.

Antes, porém, que seja feita a edição dos metadados para a inclusão dos itens descritos no Quadro 12 é necessário elucidar o passo-a-passo para depósito na coleção pela equipe editorial. Nesse ínterim, a autora elaborou um guia de depósito de dados a ser utilizado pelos depositantes e que será detalhado a seguir.

Primeiramente a equipe editorial deverá realizar um cadastro no repositório de dados da UFG. Ao acessar (dadosdepesquisa.ufg.br), basta clicar no botão “Inscreva-se, no canto

superior da página, e preencher o formulário apresentado. Um e-mail de confirmação será enviado para validar o cadastro e quando essa ação for realizada é necessário que os administradores do repositório forneçam permissões para manuseio da coleção específica do BGG. Cabe ressaltar que a coleção já foi criada pela autora com as especificações definidas na pesquisa-ação, conforme a Figura 9.

Figura 9 – Coleção de dados do BGG

Fonte: repositório de dados da UFG.

Observa-se na Figura 9 que a sub coleção do BGG foi criada dentro da coleção “Periódicos Científicos UFG” (grifo A) e já possui uma descrição de sua finalidade. Além disso, há um *link* para a página *web* do periódico (grifo B), assim como, também, outro *link* para a coleção na página do periódico, conforme visto no Apêndice A, e que garante a bidirecionalidade entre as páginas.

No grifo C, a equipe editorial poderá adicionar os conjuntos de dados previamente enviados pelos autores do BGG, cujos artigos tenham sido aceitos pelos pareceristas. Ao clicar no botão “Adicionar dados” é preciso escolher a opção “Novo conjunto de dados”, a tela a ser exibida será semelhante à apresentada na Figura 8.

Figura 10 – Adicionando conjuntos de dados na coleção do BGG

Metadados de Citações **A**

Título *
 B

Autor *

Nome * **Afiliação *** **C**

Esquema de Identificador **Identificador**

Entre em Contato *

Nome **Afiliação**

E-mail

Descrição *

Este campo é compatível apenas com determinadas [tags HTML](#).

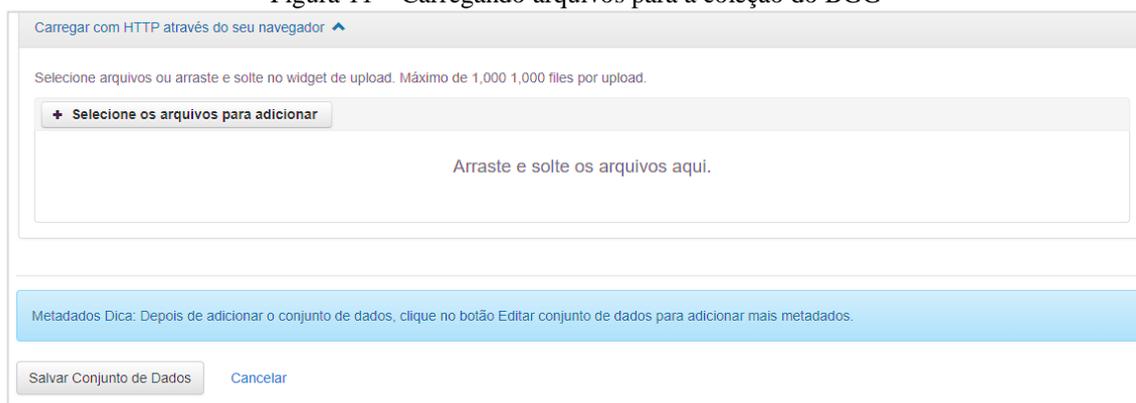
Texto *

[Continua...](#)

Fonte: repositório de dados da UFG.

Antes de carregar (fazer *upload*) os arquivos enviados pelos autores na coleção, a equipe editorial preencherá os metadados de citação (grifo A), também, fornecidos pelos autores ao enviarem o formulário (Apêndice E) preenchido. Após preencher o título do artigo relacionado ao conjunto de dados, deve-se clicar no botão marcado no grifo B, então, o título será alterado para “Dados de replicação para: ‘título do artigo’”. Nota-se no grifo C que é possível acrescentar campos para autores, informações de contato e outras, no restante do formulário, como palavras-chave que deverão ser adicionadas individualmente, além de permitir a associação de mais de uma publicação relacionada. Isso é muito útil nos casos em que um conjunto de dados gera vários manuscritos. Após o preenchimento de todos os campos marcados como obrigatórios (a partir da definição conjunta com a equipe editorial), é o momento de carregar todos os arquivos para a coleção (Figura 11).

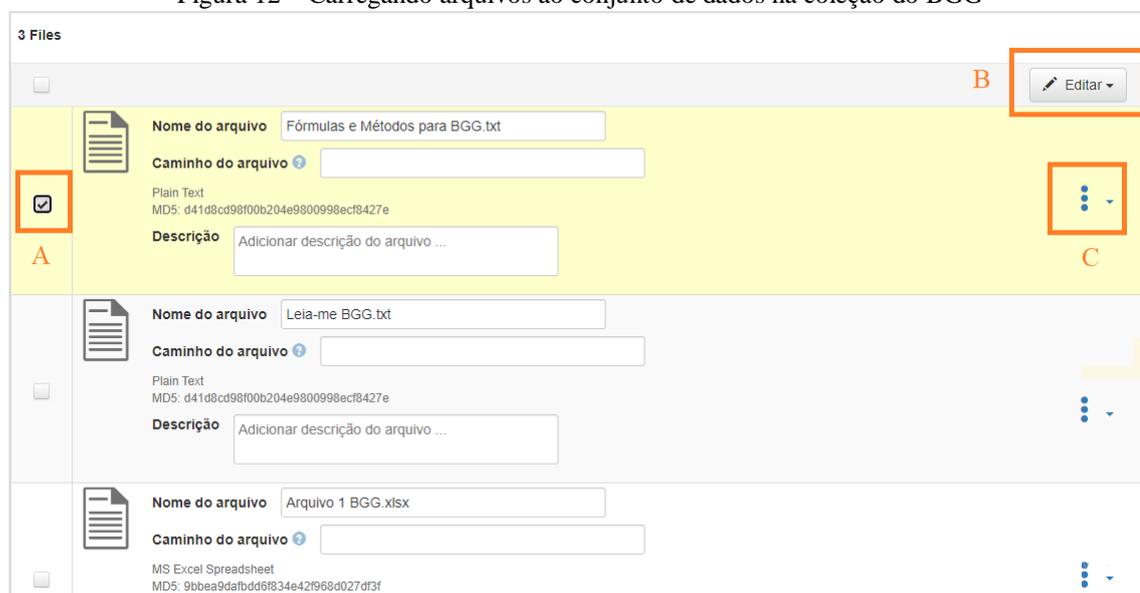
Figura 11 – Carregando arquivos para a coleção do BGG



Fonte: repositório de dados da UFG.

No botão “Selecione os arquivos para adicionar” é possível fazer o carregamento de todos os arquivos que compõem o conjunto de dados e, posteriormente, acrescentar uma descrição a eles (Figura 12).

Figura 12 – Carregando arquivos ao conjunto de dados na coleção do BGG



Fonte: repositório de dados da UFG.

Após confirmar as informações, conforme Figura 12, pode-se selecionar um ou mais arquivos (grifo A), sendo possível editá-lo(s) (grifo B): excluindo-o(s) ou restringindo acesso. Nesse caso é necessário fornecer informações sobre como e se os usuários podem obter acesso aos arquivos restritos neste conjunto de dados (Figura 13).

Figura 13 – Aplicando restrições a arquivo(s) específico(s) dentro do conjunto de dados

Restrições de arquivos ✕

Restringindo o acesso aos arquivos publicados. Você pode adicionar ou editar os Termos de Acesso para o conjunto de dados e permitir que as pessoas solicitem acesso a arquivos restritos. **Se você desativar a solicitação de acesso, deverá adicionar informações sobre o acesso ao campo Termos de acesso.**

Saiba como restringir o acesso a arquivos e conjuntos de dados no [Guia do usuário](#).

Solicitar acesso ⓘ Ativar solicitação de acesso

Termos de Acesso ⓘ

Salvar alterações Cancelar

Fonte: repositório de dados da UFG.

Aos arquivos adicionados ao conjunto de dados é possível, também, aplicar *tags* que facilitam sua identificação, sendo necessário clicar no símbolo marcado no grifo C da Figura 12 e, posteriormente, selecionar o identificador desejado (Figura 14) e salvar. Após essas configurações, basta clicar no botão “Salvar conjunto de dados” no final da página.

Figura 14 – Selecionando uma *tag* de identificação ao(s) arquivo(s)

Editar tags ✕

Selecione as tags de arquivo existentes ou crie novas tags para descrever seus arquivos. Cada arquivo pode ter mais de uma tag.

Tags selecionadas Nenhuma tag selecionada

Tags de arquivo

Tag de arquivo personalizado

- Documentação
- Data
- Código

Selecione esta tag para adicioná-la a esta tag. Adicionará essa tag como uma opção de arquivos neste conjunto de dados.

Salvar alterações Cancelar Aplicar

Fonte: repositório de dados da UFG.

Após o salvamento do conjunto de dados, antes da publicação, são passíveis algumas alterações, conforme Figura 15.

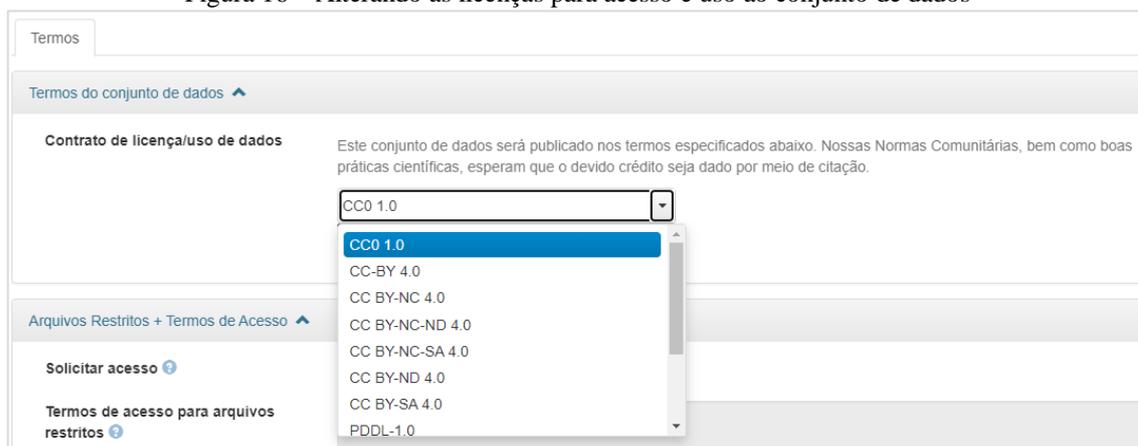
Figura 15 – Vista do conjunto de dados antes da publicação



Fonte: repositório de dados da UFG.

Na Figura 15, grifo A, nota-se que o conjunto de dados recebe um DOI, mesmo estando ainda na fase de rascunho, e será ativado quando for publicado (grifo B). No grifo C da mesma figura podem ser acrescentados novos arquivos ou feitas alterações naqueles já carregados. Ainda no grifo C, podem ser acrescentados ou editados os metadados e alterados os termos, ou licenças. Ao clicar em Termos (Figura 15) e no botão “Editar requisitos de termos”, será visualizada a tela de edição desse requisito (Figura 16).

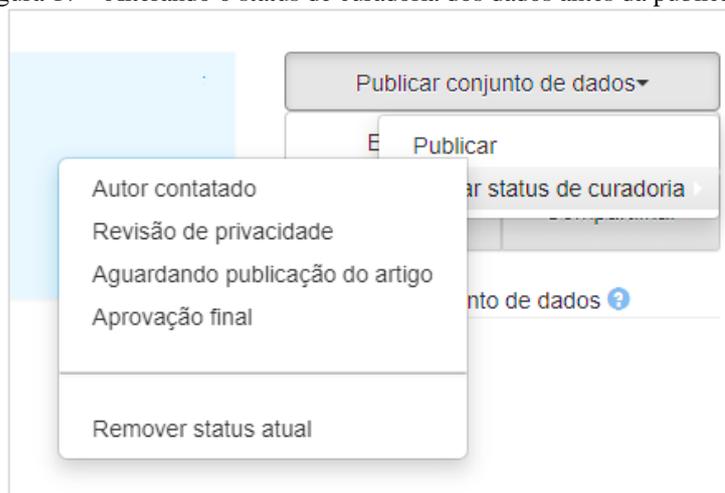
Figura 16 – Alterando as licenças para acesso e uso ao conjunto de dados



Fonte: repositório de dados da UFG.

As licenças são informadas pelos autores no formulário de metadados (Apêndice E). Após essas edições e configurações, deve-se salvar e publicar o conjunto de dados, sendo que assim, o DOI será validado, ou ainda no botão “Publicar” alterar o status da curadoria (Figura 17). Esta última ação pode ser necessária caso a opção seja publicar o conjunto de dados somente após a publicação do artigo ou se tenha que gerenciar a publicação de vários conjuntos de dados. O status de curadoria ajuda na organização.

Figura 17 – Alterando o status de curadoria dos dados antes da publicação



Fonte: repositório de dados da UFG.

Após a publicação do conjunto de dados, caso sejam feitas edições em quaisquer dos campos do grifo C da Figura 15, deve-se clicar novamente em “Publicar” e a ação realizada ficará registrada em “Versões”, com descritivo detalhado do que foi alterado.

Todo o constructo da pesquisa-ação descrito até aqui integram as ações e definições preliminares para que a coleção do BGG comece a ser utilizada, ou seja, que a equipe editorial passe a requerer o envio dos dados adjacentes aos artigos aprovados para que, então, possam ser depositados no repositório da UFG. Para que isso seja possível, conjuntamente à equipe editorial, concluiu-se que, primeiramente, se necessitará de uma ação de conscientização sobre os temas da Ciência Aberta e abertura de dados científicos, já que, na opinião da equipe, são temas novos em sua área do conhecimento. Inclusive, na primeira reunião da pesquisa-ação, na qual foi apresentado o estudo em desenvolvimento e realizada uma contextualização teórica, a equipe editorial afirmou não estar a par desse movimento amplo da Ciência Aberta, em especial, quando se mencionou sobre *preprints* e abertura de dados.

Sendo assim, elaboraram-se peças de divulgação sobre o tema para sensibilizar os leitores e autores do BGG. A proposta é que tais peças possam ser compartilhadas tanto na

página *web* do BGG, quanto em seus perfis nas redes sociais uma vez que explicam o que é Ciência Aberta, sob o viés da recomendação da UNESCO, trazem algumas vantagens sobre compartilhamento de dados, esclarecem sobre a atribuição de DOI aos conjuntos compartilhados, permitindo citação, entre outros aspectos (Apêndice F).

Nesta seção apresentaram-se as normativas para compartilhamento de dados associados aos artigos aceitos para publicação no BGG, bem como, a documentação mínima necessária para compreensibilidade dos conjuntos de dados e a definição, por parte dos autores, das licenças para acesso e uso dos dados. A coleção foi efetivamente criada dentro do repositório de dados da UFG e elaborou-se um guia para depósito de dados na referida coleção, além de esclarecer sobre o nível de curadoria que será realizado. Por fim, foram desenvolvidas imagens de divulgação científica sobre os temas tratados no estudo, com foco nos autores e leitores do BGG. Chega-se ao final do trabalho com a certeza que os objetivos foram atendidos, mas que há um longo caminho para a consolidação da Ciência Aberta na UFG, que necessitará do comprometimento e responsabilidade de toda a comunidade acadêmica rumo a essa mudança.

5 CONSIDERAÇÕES FINAIS

Começar a redigir esta seção foi um desafio, sobretudo, porque inicialmente ela estava intitulada como “Conclusões”. Alterar esse nome, por certo, tornou a missão menos árdua, afinal este é um trabalho não concluído. Aliás, existe algum estudo que chegue ao seu final, completamente pronto e acabado? Grandes teóricos, como Piaget²¹, concordariam que não! Claramente este autor, em específico, tratava da construção do conhecimento humano, porém a analogia se encaixa bem à construção do conhecimento científico. Faz-se pesquisador, encontram-se problemáticas para estudos a partir da influência dos campos científicos em que se está inserido e este ciclo se repete inúmeras vezes. Portanto, o estudo que aqui se apresentou, apoiado nas inquietações de filósofos da ciência e de pesquisadores contemporâneos da Comunicação e Ciência da Informação não está finalizado. Ele se apresenta como uma reflexão, com encaminhamentos práticos, aos pesquisadores da Universidade Federal de Goiás, em especial, editores de periódicos científicos, com o viés da Ciência Aberta.

A Ciência Aberta é um termo amplo que se associa tanto às discussões profundas sobre a socialização do conhecimento e sua privatização, quanto às práticas científicas de produção, comunicação e popularização da ciência. Neste estudo, a Ciência Aberta é tida como um dos movimentos mais expressivos da revolução científica atual: a *eScience*, a era da explosão dos dados, das formas de potencializar seu compartilhamento, maximizar seu uso e, conforme pontua Albagli (2015) possibilitar a colaboração entre pesquisadores e a participação da sociedade na Ciência.

Com esse entendimento e segundo Kuhn (1998) as revoluções científicas representaram, durante toda a história das descobertas científicas, alterações nas práticas metodológicas, ontológicas, semânticas etc. de se praticar o que Freire-Maia (2007) chamou de ciência-processo. Nesse sentido, embora em outros países do mundo esse processo de abertura do fazer científico esteja mais encaminhado, no Brasil representa uma mudança significativa que requer, antes de qualquer outra ação, a conscientização dos pesquisadores, porque práticas de Ciência Aberta representarão uma nova concepção paradigmática.

Sabe-se que o movimento pela Ciência Aberta é precedido por outros que tinham como finalidade a socialização dos conhecimentos científicos, sem custos ou outras barreiras, como o Projeto Gutenberg e o movimento de acesso aberto. Este último passa a integrar as

²¹ O autor aborda essa noção em diversas obras, referindo-se à teoria construtivista.

agendas da Ciência Aberta, principalmente, por existirem, ainda, barreiras que contrariam o objetivo inicial de democratização do conhecimento científico.

Exemplo que corrobora a afirmação anterior é a existência de periódicos científicos que oferecem acesso somente aos artigos cujos autores pagaram as taxas de publicação (via híbrida do acesso aberto), ou ainda, aqueles que mesmo oferecendo acesso ilimitado ao seu conteúdo para leitores, cobram taxas de processamento de artigo (APC) dos autores (via dourada do acesso aberto).

No contexto dos periódicos científicos geridos pelas Instituições de Ensino Superior (IES) brasileiras, ainda é comum que publiquem em acesso aberto, via diamante, sem cobrar taxas nem de autores e nem leitores. Embora comece a se discutir a cobrança de APC dos autores, haja vista a limitação de pessoal dedicado ao trabalho editorial e que comumente é voluntário, este estudo centrou-se em outra discussão, a abertura dos dados científicos associados aos artigos publicados nos periódicos.

Esta abertura possibilita a transparência da ciência, a reprodutibilidade ou replicabilidade de experimentos anteriores, atribuição de validade aos resultados descritos nos artigos e a utilização dos dados para outras análises, o que reduz novos custos com coletas que já tenham sido realizadas. Fato é que os pesquisadores da Ciência Aberta, como Borgman (2015) e Córdula e Araújo (2019) podem citar inúmeras outras vantagens, entre elas o aumento do impacto e visibilidade das pesquisas, entre outras que já foram elencadas no decorrer do estudo, nesta etapa compete concluir que uma abertura efetiva de dados científicos requer que tais dados sejam localizáveis, acessíveis, interoperáveis com outros sistemas e reutilizáveis, atendendo, portanto, aos princípios FAIR (*findable, accessible, interoperable, reusable*).

Logo, quando se debruçou sobre o objetivo de traçar orientações para que periódicos científicos criem coleções de dados abertos, diversas vezes retornou-se aos princípios FAIR, seja para definir as normativas editoriais da abertura de dados, seja para refletir sobre o povoamento da coleção no repositório da UFG e, também, para possibilitar sua efetiva usabilidade.

A partir dessa premissa, com o subsídio dos requisitos de repositórios de dados confiáveis, *Core Trust Seal* (2022), o levantamento documental realizado no repositório *Dataverse* da Universidade de *Harvard* e nas páginas *web* dos periódicos que possuem coleções no referido repositório permitiu identificar as lacunas na definição de políticas de abertura de dados e que se mostraram muito relevantes na elaboração das normativas do periódico piloto da UFG, o Boletim Goiano de Geografia (BGG).

Nesse sentido, realizaram-se algumas etapas de pesquisa-ação que permitiram, a partir dos conceitos apresentados por Thiollent (1986), pensar nas inferências “micro”, ou seja, num grupo menor, com vistas à generalização. Por conseguinte, espera-se que o caminho traçado junto ao BGG possibilite a criação de coleções de dados abertos de outros periódicos científicos da UFG e de outras instituições.

Entretanto, algumas dificuldades apresentaram-se nesse caminho. A maior delas, de ordem burocrática, uma vez que, institucionalmente, ainda não foram aprovadas as resoluções que respaldam o uso do repositório de dados da UFG. Os resultados apresentados configuram-se, então, como protótipo e foram elaborados com cuidado e detalhamento das etapas, possibilitando seu funcionamento.

Apesar da referida resolução ainda não ter sido aprovada no CONSUNI, compreende-se que isso ocorrerá em breve, uma vez que há o compromisso formal e ético da UFG perante a chamada aberta de incubação de repositório de dados científicos, a qual foi contemplada. Além disso, publicamente²², a instituição já se manifestou sobre a criação do Centro de Ciência Aberta que incorporará além da questão da abertura dos dados produzidos ao longo dos estudos, agendas voltadas para promoção de um entendimento sobre a Ciência Aberta a toda comunidade acadêmica e como efetivar práticas científicas que estejam em consonância com esse movimento.

Desse modo, no que concerne à chamada de incubação de repositórios, a UFG já realizou a personalização, configuração e instalação do seu repositório de dados científicos, se responsabilizando pela infraestrutura técnica, tecnológica e tomando ciência das necessidades de seu uso em consonância com os princípios FAIR.

Logo, mesmo com as limitações, os objetivos estabelecidos no início do estudo foram atendidos. O caminho construído contempla orientações para criação das coleções de dados abertos, normativas para povoamento e uso e tutoriais para depósito e curadoria dos dados. Qualquer periódico científico da UFG, ou externo, poderá utilizar essas orientações para refletir a criação de suas próprias coleções de dados abertos. Em relação aos periódicos externos, caso suas instituições de origem não dispuserem de repositório próprio, podem utilizar opções abertas e gratuitas, como o *Dataverse* da Universidade de *Harvard*, o *Figshare*, o *Zenodo*, o *Dspace*, entre outros. Independente da escolha, importante é a

²² <https://ufg.br/n/157513-criacao-de-centro-de-ciencia-aberta-e-discutida-na-ufg#:~:text=Ao%20final%20da%20reuni%C3%A3o%2C%20ficou,e%20%C3%B3rg%C3%A3os%20administrativos%20da%20Universidade.>

compreensão de que um repositório, bem como uma coleção de dados, deve garantir acesso em longo prazo e a compreensibilidade dos dados depositados.

Com a efetiva criação do Centro de Ciência Aberta da UFG espera-se que o funcionamento da coleção de dados proposta neste estudo seja consolidado e que, a posteriori, outras investigações venham avaliar e validar as ações propostas. Ademais, retornar-se a um ponto mencionado no início deste trabalho, o caráter de mudança que essa realidade de abertura de dados impõe. Acredita-se que a conscientização e formação da comunidade acadêmica sejam as ações mais emergentes.

Durante o estudo foram sugeridas peças de divulgação científica sobre as temáticas do estudo e sobre o posicionamento do BGG acerca desses temas. Essas peças foram elaboradas para serem publicizadas na página *web* do periódico e nas redes sociais. Entretanto, sabe-se que somente isso não é suficiente para gerar um entendimento comum sobre tais temas e para ganhar adeptos às práticas científicas mais abertas. Os pesquisadores precisarão enxergar vantagens em compartilhar seus dados e as agências de fomento representam, a exemplo da FAPESP, e representarão ainda mais, um incentivo de encorajamento.

Além disso, acredita-se que com o passar do tempo e com a popularização do uso (e reuso) dos dados científicos disponíveis em coleções de dados de periódicos científicos, mais pessoas vislumbrarão os benefícios do compartilhamento e apoiarão a causa. Novos estudos sobre o uso de repositórios de dados pelos pesquisadores nos dirão!

6 REFERÊNCIAS

ALBAGLI, Sarita. Ciência Aberta em questão. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (Org.). **Ciência Aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015.

ANDRADE, Viviane Toraci Alonso de. **Comunicação científica na sociedade em rede: uma plataforma de Ciência Aberta para o Brasil**. 2014. 227 f. Tese (Doutorado) – Departamento de Comunicação Social, Universidade Federal de Pernambuco, Recife, 2014.

APPEL, Andre Luiz. **Dimensões tecnopolíticas e econômicas da comunicação científica em transformação**. Tese (Doutorado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio De Janeiro, Rio de Janeiro, 2019.

BAPTISTA, Ana Alice; COSTA, Sely Maria de Souza; KURAMOTO, Hélio; RODRIGUES, Eloy. Comunicação científica: o papel da *open archives initiative* no contexto do acesso livre. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, [S. l.], v. 12, n. 1, p. 1–17, 2007. DOI: 10.5007/1518-2924.2007v12nesp1p1. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12nesp1p1>. Acesso em: 20 mai. 2023.

BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2011.

BAZI, Rogério. Produção da informação nos campos da Ciência da Informação e Comunicação jornalística: possíveis interfaces. **Intexto**, Porto Alegre: UFRGS, v. 1, n. 18, p. 1-14, janeiro/maio, 2007.

BOAI. Budapest Open Access Initiative. **BOAI10 Portuguese (Brazilian) Translation**. Disponível em: <https://www.budapestopenaccessinitiative.org/boai10/portuguese-brazilian-translation/>. Acesso em 02 set. 2022.

BOMFÁ, Cláudia Regina Ziliotto; Castro, João Ernesto E. Desenvolvimento de revistas científicas em mídia digital – o caso da Revista Produção Online. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 39-48, maio/ago. 2004. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652004000200004 Acesso em 04 de maio de 2021.

BOURDIEU, Pierre. **Os usos sociais da ciência**: Por uma sociologia clínica do campo científico. Tradução de Denice Barbara Catani. São Paulo: Fundação Editora da UNESP, 2003.

BOURDIEU, Pierre. **Para uma sociologia da ciência**. Tradução de Pedro Elói Duarte. Lisboa: Edições 70, 2004.

BORGMAN, Christine L. **Big data, little data, no data**: scholarship in the networked world. Cambridge, Massachusetts: MIT Press. 2015.

BRANDAO, Carlos Rodrigues. A pesquisa participante e a participação da pesquisa: um olhar entre tempos e espaços a partir da América Latina. In: BRANDÃO, Carlos Rodrigues; STRECK, Danilo (orgs.). (Org.). **Pesquisa participante: a partilha do saber**. Aparecida: Idéias & Letras, 2006. Disponível em: <https://apartilhadavida.com.br/wp-content/uploads/escritos/PESQUISA/PESQUISA%20PARTICIPANTE/A%20PARTICIPA%C3%87%C3%83O%20DA%20PESQUISA%20E%20A%20PESQUISA%20PARTICIPANTE%20-%20rosa%20dos%20ventos.pdf>. Acesso em 16 nov. 2021.

BRASIL. **Lei nº 9.610, de 19 de fevereiro de 1998**. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Brasília, DF: Presidência da República, 1998. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9610.htm. Acesso em: 05 set. 2022.

BRASIL. **Decreto de 15 de setembro de 2011**. Institui o Plano de Ação Nacional sobre Governo Aberto e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/dsn/dsn13117.htm. Acesso em: 05 set. 2022.

BRASIL. **Decreto nº 7.724, de 16 de maio de 2012**. Regulamenta a Lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição. Brasília, DF: Presidência da República, 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7724.htm#:~:text=DECRETO%20N%C2%BA%207.724%2C%20DE%2016%20DE%20MAIO%20DE%202012&text=Regulamenta%20a%20Lei%20n%C2%BA%2012.527,216%20da%20Constitui%C3%A7%C3%A3o. Acesso em: 05 set. 2022.

BRASIL. **Decreto de 12 de março de 2013**. Altera o Decreto de 15 de setembro de 2011, que institui o Plano de Ação Nacional sobre Governo Aberto. Brasília, DF: Presidência da República, 2013. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/Dsn/Dsn13594.htm. Acesso em: 05 set. 2022.

BRASIL. **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, 2016a. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 06 set. 2022.

BRASIL. **Decreto nº 8.789 de 29 de junho de 2016**. Dispõe sobre o compartilhamento de bases de dados na administração pública federal. Brasília, DF: Presidência da República, 2016b. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8789.htm. Acesso em: 06 set. 2022.

BRASIL. **Lei nº 13.709 de 14 de agosto de 2018**. Institui a Política Nacional de Governo Aberto e o Comitê Interministerial de Governo Aberto. Brasília, DF: Presidência da República, 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D10160.htm. Acesso em: 06 set. 2022.

BRASIL. **Decreto n.º 10.160 de 9 de dezembro de 2019**. Dispõe sobre o compartilhamento de bases de dados na administração pública federal. Brasília, DF: Presidência da República, 2016b. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8789.htm. Acesso em: 06 set. 2022.

BUFREM, Leilah Santiago. **Comunicação do conhecimento científico**. Brasília: CAPES - UAB; Rio de Janeiro: Departamento de Biblioteconomia, FACC/UFRJ, 2019.

CARIBÉ, Rita de Cássio do Vale. **Comunicação Científica para o público leigo no Brasil**. 2011. 320f. Tese (Doutorado em Ciências da Informação). Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília.

CARIBÉ, Rita de Cássia do Vale. Comunicação científica: reflexões sobre o conceito. **Informação & Sociedade: Estudos**, João Pessoa, v.25, n.3, p. 89-104, setembro/dezembro, 2015.

CASTRO, Regina C. Figueiredo. Impacto da internet no fluxo da comunicação científica em saúde. **Revista de Saúde Pública: São Paulo**, v. 40 (n.º especial), p. 57-63, ago. 2006.

CENTER FOR OPEN SCIENCE. **Diretrizes para promoção da transparência e abertura nas políticas e práticas de periódicos “Diretrizes TOP”**. Disponível em: <https://osf.io/us5yg>. Acesso em 24 mai. 2023.

CIÊNCIA ABERTA. **Política Nacional de Ciência Aberta**. Portugal. Disponível em: <https://www.ciencia-aberta.pt/pnca>. Acesso em 04 set. 2022.

CHALMERS, Alan F. **O que é ciência afinal?** Tradução de Raul Filker. Editora Brasiliense, 1993.

CHIZZOTTI, A. **Pesquisa qualitativa em ciências humanas e sociais**. 3. ed. Petrópolis: Vozes, 2006.

CNPQ. Conselho Nacional de Desenvolvimento Científico e Tecnológico. **CNPq/MCTI-FNDCT N° 61/2022 - Ciência Cidadã e Divulgação Científica: Programa Ciência no Mar, Ciência Antártica e Ações Nacionais voltadas à Década das Nações Unidas da Ciência Oceânica para o Desenvolvimento Sustentável**. Disponível em: http://memoria2.cnpq.br/web/guest/chamadas-publicas?p_p_id=resultadosportlet_WAR_resultadoscnpqportlet_INSTANCE_0ZaM&filtro=abertas&detalha=chamadaDivulgada&idDivulgacao=11167. Acesso em 06 set. 2022.

CONSELHO UE. Conselho fornece orientação política sobre cooperação internacional, Ciência Aberta e missões europeias. **Conselho da União Europeia**. Disponível em: <https://www.consilium.europa.eu/pt/press/press-releases/2022/06/10/council-provides-political-orientations-on-international-cooperation-open-science-and-european-missions/>. Acesso 04 set. 2022.

CÓRDULA, Flavio Ribeiro; ARAÚJO, Wagner Junqueira de. O compartilhamento de dados científicos na era do e-Science. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire de (Org.). **Dados científicos: perspectivas e desafios**. João Pessoa, PB: Editora UFPB, 2019.

CORE TRUST SEAL. **About**. 2022a. Disponível em: <https://www.CoreTrustSeal.org/about/>. Acesso 01 set. 2022.

CORE TRUST SEAL. **Core Trust Seal Trustworthy Digital Repositories Requirements 2023-2025 Extended Guidance**. (V01.00). Zenodo. 2022b. Disponível em: <https://doi.org/10.5281/zenodo.7051096>. Acesso 03 abr. 2023.

DIAS, Guilherme Ataíde; ANJOS, Renata Lemos dos; RODRIGUES, Adriana Alves. Os princípios FAIR: viabilizando o reuso de dados científicos. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire de (Org.). **Dados científicos: perspectivas e desafios**. João Pessoa, PB: Editora UFPB, 2019.

FACHIN, Gleisy Regina Bories; HILLESHEIM, Araci Isaltina de Andrade. **Periódico científico – padronização e organização**. Florianópolis: Editora da UFSC, 2006.

FAPESP. Fundação de Amparo à Pesquisa do Estado de São Paulo. **Portaria CTA nº 01/2019**. Disponível em: <https://fapesp.br/12632/portaria-cta-no-012019>. Acesso em 20 fev. 2023.

FEDERER, Lisa M., *et al.* Data sharing in PLOS ONE: An analysis of Data Availability Statements. **PLoS ONE**, v. 13, n. 5. DOI: <https://doi.org/10.1371/journal.pone.0194768>. Acesso em: 12 set. 2022.

FELCHER, Carla Denize Ott; FERREIRA, André Luis Andrejew; FOLMER, Vanderlei. Da pesquisa-ação à pesquisa participante: discussões a partir de uma investigação desenvolvida no Facebook. **Experiências em Ensino de Ciências**, v.12, n. 7, 2017.

FERREIRA, Raquel Viana. **Compartilhamento de dados de pesquisa em repositórios digitais: o cenário latino americano**. 2019. Dissertação (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília, 2019.

FIOCRUZ. **Ciência Aberta na Fiocruz**. 2021. Disponível em: <https://portal.fiocruz.br/ciencia-aberta-na-fiocruz>. Acesso em 04 jan. 2021.

FREIRE-MAIA, Newton. **A Ciência por dentro**. Petrópolis, Editora Vozes, 2007.

FUNCATE. Edital de chamada FINEP/PCAL/XBDB n.º 002/2009. **Fundação de Ciência, Aplicações e Tecnologia Espaciais**. São José dos Campos: FUNCATE, 2009. Disponível em: <https://kuramoto.wordpress.com/2009/04/20/edital-fineppcalxbdb-n%C2%BA-0012009-foi-revogado/>. Acesso em: 06 set. 2022.

GARRIDO, Isadora Dos Santos; RODRIGUES, Rosangela Schwarz. Portais de periódicos científicos online: organização institucional das publicações. **Perspectivas em Ciência da Informação**, v. 15, n. 2, p. 56–72, ago. 2010. Acesso em: 1 set. 2022.

GATTEI, Stefano. **Thomas Kuhn's "linguistic turns" and the legacy of logical empiricism: incommensurability, rationality, and the search for truth** Hampshire: Ashgate Publishing Company, 2008.

GIBBS, Graham. **Análise de dados qualitativos**. Tradução de Roberto Cataldo Costa. Porto Alegre: Artmed, 2009.

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

GUANAES, Paulo Cezar Vieira. Quais direitos podem ser alegados sobre dados de pesquisa? In: SHINTAKU, Milton; SALES Luana Farias; COSTA, Michelli (Org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: ABEC, 2020.

GOMES, Suely Henrique de Aquino; SANTOS, Andréa Pereira dos; REIS, Felipe. Cielo da comunicação científica: estrutura e evolução. In: GOMES, Suely Henrique; et al (Org). **Letramento Informacional: entendendo a ciência e a comunicação científica**. Goiânia: Gráfica UFG, 2020.

GOULEMOT, Jean Marie. Da leitura como produção de sentidos. In: ROGER, Chartier (org.). **Práticas de Leitura**. 5. ed. São Paulo: Estação Liberdade, 2011.

GOVERNMENT OF CANADA. Canada's 2018-2020 National Action Plan on Open Government. 2020. Disponível em: <https://open.canada.ca/en/content/canadas-2018-2020-national-action-plan-open-government#toc8>. Acesso em 04 set. 2022.

HARVARD DATAVERSE. **Join a growing community of Harvard and worldwide researchers who share data in the Harvard Dataverse Repository**. 2022. Disponível em: <https://support.dataverse.harvard.edu/>. Acesso em 01 set. 2022.

HENNING, Patricia; MOREIRA, João. Ciência Aberta, dados abertos e princípios FAIR: uma contribuição dos Países Baixos. In: SHINTAKU, Milton; SALES Luana Farias; COSTA, Michelli (Org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: ABEC, 2020.

HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (Organizadores). **O quarto paradigma: descobertas científicas na era da eScience**. Tradução de Leda Maria Marques Dias Beck. São Paulo: Oficina de Textos, 2011.

HUVILA, Isto. Improving the usefulness of research data with better paradata. **Open Information Science**, 2022. Disponível em: <https://doi.org/10.1515/opis-2022-0129>. Acesso em 24 mai. 2023.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). Disponível em: www.cbissn.ibict.br/. Acesso em 06 de junho de 2021.

KUHN, Thomas S. **A Estrutura das Revoluções Científicas**. 5 ed. São Paulo: Editora Perspectiva, 1998.

KURAMOTO, Hélio. Acesso livre: um caso de soberania nacional? In: TOUTAIN, Lídia Maria Batista Brandão (Org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007.

LAENDER, Alberto Henrique Frade *et al.* Abertura e gestão de dados: desafios para a ciência brasileira. Rio de Janeiro: **Academia Brasileira de Ciências**, 2020. Disponível em: <https://www.abc.org.br/wp-content/uploads/2020/09/ABC-Abertura-e-Gest%C3%A3o-de-Dados-desafios-para-a-ci%C3%A2ncia-brasileira.pdf>. Acesso em 21 mai. 2023.

LE COADIC, Yves François. **A ciência da informação**. Tradução de Maria Yêda F. S. Filgueiras Gomes. Brasília: Briquet de Lemos, 1996.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto**. Brasília: IBICT, 2009.

MEADOWS, Arthur Jack. **A comunicação científica**. Tradução de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos, 1999.

MENDES, Suênia Oliveira. **Periódicos científicos em acesso aberto: uma análise do povoamento do Directory of Open Access Journals (DOAJ)**. 2019. 223f. Tese (Doutorado em Ciências da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2019.

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE. **Le Plan national pour la science ouverte: les résultats de la recherche scientifique ouverts à tous, sans entrave, sans délai, sans paiement**. France, 2019. Disponível em: <https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-tous-49241>. Acesso 04 set. 2022.

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO. **Instrução normativa nº 4, 12 de abril de 2012**. Institui a Infraestrutura Nacional de Dados Abertos (INDA). Diário Oficial da União, Edição 72, do dia 13/04/2012, Seção 1. Disponível em: <https://dados.gov.br/pagina/instrucao-normativa-da-inda>. Acesso em 06 set. 2022.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA INOVAÇÕES E COMUNICAÇÕES. **Portaria nº 3.679, de 07 de agosto de 2019**. Dispõe sobre a recriação do grupo de trabalho para elaboração de proposta de política nacional para Ciência Aberta. Diário Oficial da União, Edição 156, do dia 14/08/2019, Seção 1. Página 16. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-3.679-de-7-de-agosto-de-2019-210511903>. Acesso em 06 set. 2022.

MIRANDA, Angélica C. D.; DAMASIO, Edilson; FIRME, Simone Machado (Organizadores). **Ciência Aberta: visão e contribuição a partir dos Periódicos Científicos**. Rio Grande: Editora da Furg, 2020.

MIRANDA, Ana Cláudia Carvalho de; CARVALHO, Edirsana Maria Ribeiro de; COSTA Maria Ilza da. O impacto dos periódicos na comunicação científica. **Biblos: Revista do Instituto de Ciências Humanas e da Informação**, v. 32, n. 1, p. 01-22, jan./jun. 2018.

MONTEIRO, Elizabete Cristina de Souza de Aguiar. **Plano de Gestão de dados**. Seção Técnica de Referência e Atendimento ao Usuário e Documentação STRAUD. Marília, SP: Biblioteca da FFC/Unesp, 2020. Disponível em: <https://www.marilia.unesp.br/Home/Biblioteca/pgd-biblioteca.pdf>. Acesso em 11 set. 2022.

MOURA, Mariluce. Universidades públicas respondem por mais de 95% da produção científica do Brasil. **Blog Ciência na rua**. 11 abr. 2019. Disponível em: <https://ciencianarua.net/universidades-publicas-respodem-por-mais-de-95-da-producao-cientifica-do-brasil/>. Acesso em 06 set. 2022.

MUELLER, Suzana Pinheiro Machado. A ciência, o sistema de comunicação científica e a literatura científica. In: Campello, Bernardete Santos; Cendón, Beatriz Valadares. (Orgs.) **Fontes de Informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000a. p. 21-34.

MUELLER, Suzana Pinheiro Machado. O periódico científico. In: Campello, Bernardete Santos; Cendón, Beatriz Valadares. (Orgs.) **Fontes de Informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000b. p.73-96.

NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE. **Reproducibility and Replicability in Science**. Washington, DC: The National Academies Press, 2019. DOI: <https://doi.org/10.17226/25303>. Acesso em 09 set. 2022.

NATIONAL RESEARCH COUNCIL. **Bits of Power: Issues in Global Access to Scientific Data**. Washington, DC: National Academies Press, 1997.

NWO. Open Science. **Netherlands Organization for Scientific Research**. Disponível em: <https://www.nwo.nl/en/open-science>. Acesso em 04 set. 2022.

OGP. **Open Government Partnership**. Declaração de Governo Aberto. [S.l]: OGP, 2011. Disponível em: <https://www25.senado.leg.br/web/atividade/sinfleg/relatorio-de-gestao/saiba-mais-23#:~:text=Lan%C3%A7ada%20em%202011%2C%20a%20Open,em%20rela%C3%A7%C3%A3o%20aos%20dados%20p%C3%ABlicos..> Acesso em: 6 set. 2022.

ODI. **Open Data Institute Canadá**. Disponível em: <https://theodi.org/>. Acesso em 04 set. 2022.

OLIVEIRA, Adriana Carla Silva de; SILVA, Edilene Maria da. Ciência Aberta: dimensões para um novo fazer científico. **Informação & Informação**, Londrina, v. 21, n. 2, p. 5 – 39, mai./ago., 2016.

OLIVEIRA, Lais Pereira de. Movimento acesso livre e aberto: origens, desenvolvimento, prerrogativas e produtos resultantes. In: GOMES, Suely Henrique; et al (Org). **Letramento Informacional: entendendo a ciência e a comunicação científica**. Goiânia: Gráfica UFG, 2020.

OLIVEIRA, Maxwell Ferreira de. **Metodologia científica: um manual para a realização de pesquisas em Administração**. Catalão: UFG, 2011.

OPEN GOVERNMENT. U.S. Help create the 5th U.S. National Action Plan for Open Government. **Open Government Initiatives**. Disponível em: <https://open.usa.gov/national-action-plan/co-creation/>. Acesso em 04 set. 2022.

PACKER, Abel L. O Programa SciELO e o Acesso aberto via Dourada [The SciELO Program and Gold Road Open Access]. Translated from English. In: PERUZZO, C. M. K., MARTINS, M. L. and GABRIOTI, R., eds. **Revistas Científicas de Comunicação Ibero-Americanas na Política de Divulgação do Conhecimento: tendências, limitações e os desafios de novas estratégias**. Braga, Portugal: UMinho Editora, 2021. Disponível em: <https://doi.org/10.21814/uminho.ed.43.4>. Acesso em 19 mai. 2023.

PAVÃO, Caterina Groposo; SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia da. Gestão de dados em periódicos científicos. In: SHINTAKU, Milton; SALES Luana Farias; COSTA, Michelli (Org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: ABEC, 2020.

REZENDE, Laura Vilela Rodrigues; BARBOSA, Sonia. Using the Dataverse project to move towards fair principles. In: SALES, Luana Farias; OLIVEIRA, Viviane Santos de Veiga; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 31-46. DOI: 10.22477/9786589167242.cap3

REZENDE, Laura Vilela Rodrigues, FALGUERAS, Ernest ABADAL. Estado da arte dos marcos regulatórios brasileiros rumo à Ciência Aberta. **Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da informação**, Florianópolis, v. 25, p. 01-25. DOI: 10.5007/1518-2924.2020.e71370. Acesso em: 05 set. 2022.

RODRIGUES, Marcello Mundim; DIAS, Guilherme Ataíde; LOURENÇO, Cíntia de Azevedo. Repositórios de dados científicos na América do Sul: uma análise da conformidade com os Princípios FAIR. **Em Questão**, Porto Alegre, v. 28, n. 2, e-113057, abr./jun. 2022. Disponível em: <https://www.redalyc.org/journal/4656/465669993013/html/>. Acesso em: 18 jul. 2023.

ROMANOWSKI, Joana Paulin; ENS, Romilda Teodora. As pesquisas denominadas do tipo “estado da arte” em educação. **Diálogo Educacional**, Curitiba, v. 6, n. 19, p. 37-50, set./dez.

2006. Disponível em:

<https://periodicos.pucpr.br/index.php/dialogoeducacional/article/view/24176>. Acesso em: 19 nov. 2021.

SALES, Luana Farias; SAYÃO, Luís Fernando. A grande e a pequena ciência: análise das diferenças na gestão de dados de pesquisa. **Inf. & Soc.: Est.**: João Pessoa, v.29, n.3, p. 151-170, jul./set. 2019.

SANTAREM SEGUNDO, José Eduardo. Web semântica e linked data como aliados na consolidação da Ciência Aberta. In: SHINTAKU, Milton; SALES Luana Farias; COSTA, Michelli (Org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: ABEC, 2020.

SANTOS, Anne Danielle Soares Clinio dos; SANTOS Paula Xavier dos. Cenário Internacional. **Curso Panorama Histórico da Ciência Aberta**. Campus Virtual: FIOCRUZ, 2022.

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.

SANTOS, Paula Xavier dos; GUANAES, Paulo. Ciência Aberta, dados abertos: desafio e oportunidades. **Trabalho, educação e saúde**: Rio de Janeiro. v.16, n.1, Editorial, jan./abr. 2018. Disponível em: <https://doi.org/10.1590/1981-7746-sol00120>. Acesso em 20 mai. 2023.

SCHIESSL, Ingrid Torres. **A relação entre os conjuntos de dados de pesquisa depositados em repositórios de dados e os periódicos científicos**. 2020. 132p. Dissertação (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília, 2020.

SCIELO. Scientific Electronic Library. **Linhas prioritárias de ação 2019-2023**. Atualiza as linhas de prioritárias ação para o Programa Scielo para o período 2019-2023, a fim de orientar o desenvolvimento de periódicos e coleções nacionais da Rede 2021a. Disponível em: <https://wp.scielo.org/wp-content/uploads/linhas-prioritarias-acao-2019-2023.pdf>. Acesso em 06 set. 2022.

SHINTAKU, Milton; SALES, Luana Farias (Orgs.) **Ciência Aberta para editores científicos**. Botucatu, SP: ABEC, 2019. 108 p.

SHINTAKU, Milton; APPEL, André Luiz; OLIVEIRA, Alexandre Faria de. Tecnologias para gestão de dados de pesquisa segundo preceitos FAIR. In: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 129 - 146. DOI: 10.22477/9786589167242.cap10.

SOUZA, Marcia Izabel Fugisawa; VISOLI, Marcos Cezar; TORRES, Tércia Zavaglia. **Catálogo de Dataset no Repositório de Dados da Embrapa: a Experiência do Projeto-piloto de Implantação de Gestão de Dados de Pesquisa no Laboratório Multiusuário de Bioinformática da Embrapa**. Campinas : Embrapa Informática Agropecuária, 2020.

SIBBR. **Sistema de Informação sobre a Biodiversidade Brasileira** (SiBBr). Disponível em: <https://sibbr.gov.br/>. Acesso em 18 set. 2022.

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

SILVEIRA, Lúcia da. *et al.* Ciência aberta na perspectiva de especialistas brasileiros: proposta de taxonomia. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, [S. l.], v. 26, p. 01-30, 2021. DOI: 10.5007/1518-2924.2021.e79646. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/79646>. Acesso em 23 mar. 2023.

SILVEIRA, Lúcia da. *et al.* Taxonomia da Ciência Aberta: revisada e ampliada. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, [S. l.], v. 28, p. 1–22, 2023. DOI: 10.5007/1518-2924.2023.e91712. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/91712>. Acesso em: 27 jul. 2023.

THIOLLENT, Michel. **Metodologia da pesquisa-ação**. 2 ed. São Paulo: Cortez, 1986.

TRIPP, David. Pesquisa-ação: uma introdução metodológica. **Educação e Pesquisa**, São Paulo, v. 31, n. 3, p. 443-466, set.- dez. 2005. DOI: <https://doi.org/10.1590/S1517-97022005000300009>. Acesso em 14 nov. 2021.

UKRI. Shaping our open access policy. **UK Research and Innovation**. United Kingdom, 2022. Disponível em: <https://www.ukri.org/what-we-offer/supporting-healthy-research-and-innovation-culture/open-research/open-access-policies-review/how-we-decided-on-our-open-access-policy/>. Acesso em 04 set. 2022.

UNESCO. Conferência Geral da Organização das Nações Unidas para a Educação, a Ciência e a Cultura. **Recomendação da UNESCO sobre Ciência Aberta**. Brasília, Escritório da UNESCO Brasília, 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em 29 abr. 2022.

UZWYSHYN, Ray. Research Data Repositories: The What, When, Why, and How. **Computers in Libraries**, Medford, v. 36, n. 3, 2016. Disponível em: <https://www.infotoday.com/cilmag/apr16/Uzwysyn--Research-Data-Repositories.shtml>. Acesso em: 18 jul. 2023.

VANZ, Samile Andréa de Souza; SILVA FILHO, Rubens da costa. O Protagonismo das Revistas na Comunicação Científica: Histórico e Evolução. In: Carneiro, Felipe Ferreira Barros; Ferreira Neto, Amarílio; Santos, Wagner dos. (Orgs.) **A Comunicação Científica em Periódicos**. 1. Ed. Curitiba: Appris, 2019. p. 20-57 (Versão Kindle).

VEIGA, Viviane Santos de Oliveira et al. (2019). Plano de gestão de dados fair: uma proposta para a Fiocruz. **Liinc em Revista**: Rio de Janeiro, v.15, n.2, p. 275-286, novembro 2019. DOI: <https://doi.org/10.18617/liinc.v15i2.5030>. Acesso em 11 set. 2022.

WIPO. World Intellectual Property Organization. **What is Intellectual Property?** 2022a. Disponível em: <https://www.wipo.int/about-ip/en/>. Acesso em 10 set. 2022.

WIPO. World Intellectual Property Organization. **What is copyright?** 2022b. Disponível em: <https://www.wipo.int/copyright/en/>. Acesso em 10 set. 2022.

WOELFLE, Michael; OLLIARO, Piero; TODD, Matthew H. Open science is a research accelerator. **Nature Chemistry**, v. 3, n. 10, p. 745–748, 23 set. 2011.

Doi:10.1038/nchem.1149. Disponível em:

<http://www.nature.com/nchem/journal/v3/n10/full/nchem.1149.html>. Acesso em: 12 nov. 2021.

YAMAMOTO, Oswaldo H. et al. Avaliação de periódicos científicos brasileiros da área da psicologia. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 163-177, maio/ago. 2002. Disponível em: <http://revista.ibict.br/ciinf/article/view/971/1008>. Acesso em 06 de junho de 2021.

APÊNDICE A - Texto para página *web* do BGG, menu “Sobre”

Conformidade com a Ciência Aberta

O Boletim Goiano de Geografia (BGG) está comprometido com práticas científicas que estejam em consonância com a Ciência Aberta, como a disponibilização de todos seus artigos em acesso aberto, o que favorece a democratização mundial do conhecimento.

Nesse sentido, encorajamos os autores a disponibilizarem publicamente todos os dados ou códigos usados para coletá-los ou analisá-los e que sejam relevantes para os artigos publicados no BGG. Defendemos que os dados que geraram os artigos que publicamos devem estar tão visíveis e acessíveis ao público quanto possível.

Como regra geral, todos os autores que submetam artigos para avaliação no BGG devem enviar a Declaração de Disponibilidade de Dados, assinada por todos os coautores, nela deverão explicitar seu posicionamento sobre o compartilhamento dos dados científicos, subjacentes ao manuscrito submetido ao BGG.

Tais dados, quando disponibilizados pelos autores, são acessíveis, identificáveis e visíveis dentro da coleção do BGG no repositório de dados da UFG (<https://dadosdepesquisa.ufg.br/dataverse/bgg/>).

Mais orientações sobre preparação e envio dos dados devem ser consultadas na guia “Diretrizes para autores”.

APÊNDICE B – Texto de apresentação da coleção de dados do BGG

O Boletim Goiano de Geografia (BGG) é um periódico científico vinculado ao Instituto de Estudos Socioambientais (IESA). O BGG foi criado em 1981 e desde então tem como objetivo difundir conhecimento sobre a geografia e áreas afins por meio de artigos originais e ensaios teóricos resultantes de pesquisas de reconhecido mérito técnico-acadêmico.

Esta coleção abriga dados gerados em pesquisas científicas cujos manuscritos estão publicados no BGG. Todo conteúdo do BGG está disponível gratuitamente no site (www.revistas.ufg.br/bgg).

APÊNDICE C - Declaração de Disponibilidade de Dados a ser preenchida pelos autores

Declaração de Disponibilidade de Dados

Autores são encorajados a disponibilizar os dados (arquivos de texto, áudio, imagens, planilhas, códigos etc.) subjacentes ao manuscrito científico aceito para publicação no Boletim Goiano de Geografia.

Em consonância com a Ciência Aberta, tal compartilhamento contribui com a transparência, preservação e reuso dos dados científicos e eventual reprodutibilidade das pesquisas.

Desse modo, assinale a sentença a seguir que melhor expressa a disponibilidade dos dados subjacentes ao seu manuscrito.

Obs.: No caso de manuscritos que utilizaram dados disponíveis em bases públicas (como dados do IBGE, dados governamentais etc.) E dados gerados durante a pesquisa, podem ser assinaladas mais de uma sentença.

() Este artigo não envolve análise de dados científicos.

() Os dados serão disponibilizados pela equipe editorial para depósito na coleção do BGG no repositório da UFG.

Nesse caso, é preciso ratificar o consentimento:

Certifico que o(s) autor(es) autoriza(m) a redistribuição/publicação dos conjuntos de dados associados ao manuscrito para fins de transparência, eventual replicação ou novos estudos.

() Foram utilizados dados em acesso aberto que já estão disponíveis em um repositório que emite um identificador único.

Nesse caso, informe o DOI: _____

() Foram utilizados dados que já estão disponíveis abertamente em uma base de dados que não emite um identificador único. (Ex.: Dados governamentais).

Nesse caso informe o link para acesso: _____

() Os dados serão fornecidos apenas mediante solicitação direta aos autores.

Nesse caso informar como contatar os autores: _____.

() Os dados não podem ser divulgados publicamente.

Nesse caso informe a justificativa: _____

Nesse caso os autores desejam compartilhar pelo menos os metadados: Sim Não

Essa declaração deve ser assinada por todos os autores.

APÊNDICE D – Conteúdo para guia de Instruções aos Autores – Preparação e envio dos dados

Compartilhamento dos Dados

Em consonância com a Ciência Aberta, incentivamos que os autores disponibilizem todos os arquivos de texto, áudio, imagens, planilhas, códigos etc., coletados ou obtidos nas análises e que sejam relevantes para os artigos publicados no BGG.

Assim que seu artigo for aceito para publicação, será sugerido o envio dos dados à equipe editorial. O conjunto de dados será depositado na coleção do BGG no repositório de dados da UFG (<https://dadosdepesquisa.ufg.br/dataverse/bgg/>), de acordo com os princípios FAIR (findable, accessible, interoperable and reusable), ou seja, de modo que sejam localizáveis, acessíveis, interoperáveis e passíveis de serem reutilizados, e com seu(s) nome(s), sendo que será atribuído um identificador unico de objeto digital (DOI), específico para seus conjuntos de dados ou códigos, o que permitirá que sejam citados por outros autores que venham a utilizá-los.

As exceções se aplicam à responsabilidade legal relacionada ao compartilhamento de dados, como dados sensíveis, direitos autorais, patentes ou outros. No entanto, espera-se que os metadados referentes aos conjuntos de dados restritos sejam compartilhados.

Como regra geral, todos os autores que tiverem seus artigos aceitos para publicação devem enviar a Declaração de Disponibilidade de Dados, assinada por todos os coautores, caso ainda não tenham enviado na submissão do artigo. O envio dos conjuntos de dados é desejável, mas, por enquanto, não obrigatório.

Aqueles que endossam práticas científicas abertas e aceitam enviar seus conjuntos de dados, observem as instruções listadas a seguir.

Instruções para preparação dos dados a serem enviados à equipe editorial, **caso o manuscrito tenha sido aprovado**.

O método de disponibilização do conjunto de dados será informado pela equipe editorial diretamente aos autores, mas é importante a preparação dos arquivos conforme as orientações a seguir, com vistas à sua reutilização.

1 – Conjunto de dados: os arquivos devem incluir todos os dados coletados ou obtidos por meio das análises, como imagens, planilhas, arquivos de texto ou outros, sendo que devem ser privilegiados formatos que podem ser acessados por softwares livres.

No caso de uso de variáveis, estas devem ser nomeadas de forma a facilitar sua identificação.

Se os autores usarem conjuntos de dados existentes (por exemplo, Survey), é preciso relatar no próprio artigo os comandos utilizados para obtê-los e sua localização.

2 – Formulário descritivo com metadados: realizar o download, preenchimento e envio do arquivo. Para os conjuntos de dados com restrições de acesso, é possível encaminhar apenas esse formulário.

3 – Arquivo descritivo de métodos e/ou fórmulas: um arquivo de texto que descreve exhaustivamente e detalhadamente os métodos de coleta, processamento dos dados, bem como de tratamento e as eventuais fórmulas matemáticas usadas nas análises. Podem ser usados comentários para descrever os comandos ou variáveis. Deve-se fornecer informações necessárias sobre todas as etapas realizadas (por exemplo, codificar variáveis, remover determinados valores etc.). O objetivo é oferecer um caminho detalhado para replicar os resultados do trabalho de pesquisa. Além disso, deve incluir informações sobre os recursos computacionais necessários (hardwares, softwares e suas versões).

4 – Códigos-fonte: caso tenham sido utilizados códigos de programação para obtenção dos dados. Informar no arquivo “leia-me”, descrito a seguir, o software necessário para executar o código e sua versão.

5 – Arquivo leia-me: um arquivo de texto que contenha informações de contato dos autores (nome, instituição e e-mail) e uma breve descrição sobre todos os outros arquivos.

APÊNDICE E - Conjunto de metadados para coleção do BGG

Metadado	Descrição	Forma de Preenchimento	Obrigatoriedade	Sua Resposta
Título	Título do artigo associado	Texto de até 30 palavras	Sim	
Título alternativo	Título do artigo traduzido para outra língua	Texto de até 30 palavras	Não	
Autor	Identificação do autor do artigo ou da(s) pessoa(s) responsável pela coleta e organização dos dados.	Nome completo, ultimo grau acadêmico, instituição de vínculo, e-mail, ORCID*	Sim	
Coautor 1	Identificação do coautor do artigo ou colaborador(es) na coleta e organização dos dados.	Nome completo, ultimo grau acadêmico, instituição de vínculo, e-mail, ORCID*	Não	
Coautor 2	Identificação do coautor do artigo ou colaborador(es) na coleta e organização dos dados.	Nome completo, ultimo grau acadêmico, instituição de vínculo, e-mail, ORCID*	Não	
Descrição dos dados	Um resumo descrevendo a finalidade, natureza e o escopo do conjunto de dados. Além disso, informações sobre como os dados foram coletados, tratados e usados no seu artigo.	Texto de até 200 palavras	Sim	
Área do Conhecimento	Escolha uma opção: Ciências Exatas e da Terra; Ciências Biológicas; Engenharias; Ciências da Saúde; Ciências Agrárias; Ciências Sociais; Ciências Humanas; Linguística, Letras e Artes	De acordo com o CNPq	Sim	
Período de tempo coberto	Recorte temporal do conjunto de dados	Intervalo temporal	Sim	
Cobertura Geográfica	Cobertura Geográfica dos dados	Texto de até 30 palavras		
Palavras-chave	Referentes ao conjunto de dados	3 a 4 palavras separadas por vírgula	Sim	
Licença**	Informe a licença aplicada ao uso dos dados, ver opções no final do formulário	Creative Commons Ex.: CC BY	Sim	
Idioma	Referente ao conjunto de dados	Texto	Sim	
Informação de Financiamento	Caso a pesquisa que originou o artigo científico e os dados tenha sido financiada por algum órgão	Agência/N.º Processo	Não	
Tipos de dados	Tipos de dados constantes no conjunto de dados	Questionários, planilhas, dados legíveis por softwares específicos etc.	Sim	
Software	No caso de dados legíveis por máquinas, indicar o software	Nome do software/versão	Não	

Outros conjuntos de dados relacionados	Material externo relacionado ao conjunto de dados, como apêndices ou informações de amostragem	DOI ou link	Não	
Notas	Observações importantes sobre o conjunto de dados que não tenham sido contempladas por outro metadado	Texto de até 200 palavras		

* O *Open Researcher and Contributor ID* (ORCID) é um identificador único dos pesquisadores e utilizado internacionalmente para vincular as produções aos seus autores, sem ambiguidades.

** Licenças aplicáveis ao conjunto de dados (GUANAES , 2020) adaptado pela autora.

CC0 1.0 - Permite cópia, modificações, distribuição e execução dos dados, mesmo para fins comerciais, sem necessidade de solicitar autorização ou citar a autoria;

CC BY (Recomendável pelo BGG) - Permite cópia, distribuição, remixagem, adaptação e criação a partir dos dados originais, mesmo para fins comerciais. Exige atribuição de crédito à obra original.

CC BY SA – Permite cópia, distribuição, remixagem, adaptação e criação a partir dos dados originais, mesmo para fins comerciais. Exige atribuição de crédito à obra original. Novos conjuntos de dados obtidos a partir do original devem ser compartilhados sob essa mesma licença.

CC BY-NC-SA – Permite cópia , remixagem, adaptação e criação a partir do conjunto de dados original para fins não comerciais. Exige atribuição de crédito à obra original. Novos conjuntos de dados obtidos a partir do original devem ser compartilhados sob essa mesma licença.

CC BY-ND – Permite cópia e reuso para qualquer finalidade, inclusive comercial. Exige atribuição de crédito à obra original. Não permite compartilhamento de novos conjuntos de dados obtidos a partir do original.

CC BY-NC – Permite cópia, remixagem, adaptação e criação a partir do conjunto de dados original para fins não comerciais. Exige atribuição de crédito à obra original. Novas criações não precisam vir sob a mesma licença.

CC BY-NC-ND - É a mais restritiva das licenças. Só permite cópia do conjunto de dados original, o qual não pode ser alterado ou usado para fins comerciais. Deve-se atribuir o crédito pela criação original.

APÊNDICE F – Carrossel²³ de divulgação sobre Ciência Aberta e compartilhamento de dados científicos

VOCÊ SABE O QUE É CIÊNCIA ABERTA?

CIÊNCIA ABERTA

Combina vários movimentos e práticas que têm o objetivo de disponibilizar abertamente conhecimento científico, torná-lo acessível e reutilizável para todos, aumentar as colaborações científicas e o compartilhamento de informações para o benefício da ciência e da sociedade.

UNESCO (2022)

O BOLETIM GOIANO DE GEOGRAFIA

JÁ ADERE A UMA DESSAS PRÁTICAS: O ACESSO ABERTO! TODO NOSSO CONTEÚDO É DE LIVRE ACESSO. NÃO HÁ TAXA PARA AUTORES E NEM PARA LEITORES

CIÊNCIA ABERTA

A PARTIR DAS RECOMENDAÇÕES DA UNESCO...
PODEMOS MAIS!!!

O BGG começa a incentivar o compartilhamento dos dados associados aos artigos aceitos para publicação!

dadosdepesquisa.ufg.br/dataverse/bgg

Compartilhar dados científicos:

- aumenta a visibilidade das pesquisas;
- promove novas possibilidades de usos dos dados;
- reduz custos de novas coletas;
- amplia a transparência...

E MUITO MAIS!!!

²³ Na linguagem das redes sociais, carrossel refere-se a uma postagem com até dez mídias.

E COMO FICA A QUESTÃO DA AUTORIA?

O conjunto de dados compartilhado recebe um **DOI** que viabiliza a citação correta e identificação em caso de reuso.

→

Além disso, você escolhe a licença **Creative Commons** aplicável ao seu conjunto de dados.

E não precisa abrir dados com restrições éticas, legais ou que envolvam patentes!

“É PRECISO ABRIR QUANDO POSSÍVEL, MAS FECHAR QUANDO NECESSÁRIO.”

Comissão Europeia sobre gestão de dados em pesquisa

Quer saber mais?

Acesse nossa página e conheça nossa política de compartilhamento de dados

revistas.ufg.br/bgg

PPGeo UFG

BOLETIM GOIANO DE GEOGRAFIA