



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

ROSE SUELLEN DE CASTRO LISBOA

**COMPARTILHAMENTO DE DADOS CIENTÍFICOS:** percepção e práticas dos docentes  
dos programas de pós-graduação do ITEC e ICSA

Belém

2021

ROSE SUELLEN DE CASTRO LISBOA

**COMPARTILHAMENTO DE DADOS DE PESQUISA: percepção e práticas dos docentes dos programas de pós-graduação do ITEC e ICSA**

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Informação, do Instituto de Ciências Sociais Aplicadas da Universidade Federal do Pará, como parte dos requisitos necessários para obtenção do título de Mestre em Ciência da Informação.

Linha de Pesquisa: Mediação e uso da Informação

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Danielly Oliveira Inomata

Coorientador: Prof<sup>º</sup>. Dr. Cristian Berrío-Zapata

Belém

2021

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD  
Sistema de Bibliotecas da Universidade Federal do Pará  
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

---

- L769c Lisboa, Rose Suellen de Castro.  
Compartilhamento de dados de pesquisa : percepção e práticas dos docentes dos programas de pós-graduação do ITEC e ICSA / Rose Suellen de Castro Lisboa. — 2021.  
111 f. : il. color.
- Orientador(a): Prof<sup>ª</sup>. Dra. Danielly Oliveira Inomata  
Coorientador(a): Prof. Dr. Cristian Berrio-zapata  
Dissertação (Mestrado) - Universidade Federal do Pará,  
Instituto de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Ciência da Informação, Belém, 2021.
1. Compartilhamento de dados. 2. Dados de pesquisa. 3. Ciência aberta. I. Título.

ROSE SUELLEN DE CASTRO LISBOA

**COMPARTILHAMENTO DE DADOS DE PESQUISA:** percepção e práticas dos  
docentes dos programas de pós-graduação do ITEC e ICSA

Dissertação apresentado ao Programa de Pós-graduação em Ciência da Informação, do Instituto de Ciências Sociais Aplicadas da Universidade Federal do Pará, como parte dos requisitos necessários para obtenção do título de Mestre em Ciência da Informação.

Data de aprovação:

Banca examinadora:

---

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Danielly Oliveira Inomata (Orientadora)  
Universidade Federal do Amazonas – UFAM

---

Prof<sup>º</sup>. Dr. Cristian Berrío-Zapata (Coorientador)  
Universidade Federal do Pará – UFPA

---

Prof. Dr. Ricardo César Gonçalves Sant’Ana  
Universidade Estadual Paulista - UNESP

---

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Luana Farias Sales Marques  
Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT

---

Dr<sup>ª</sup>. Paloma Marín Arraiza  
ORCID Inc.

Ao meu filho Gael, meu marido Nonato e meus  
pais Marinete e Raimundo.

## AGRADECIMENTOS

A Deus, porque até aqui tem me sustentado, não me deixando desistir dos meus sonhos.

Aos meus orientadores, Danielly Inomata e Cristian Berrío-Zapata, por toda atenção dada ao meu trabalho.

Aos professores da banca de qualificação, pela contribuição para o melhor direcionamento deste estudo.

À minha família, pelo apoio incondicional. Sem ela eu não teria conseguido. Meu marido Nonato e em especial ao meu filho Gael, obrigada pelas noites que passou ao meu lado enquanto eu estudava.

À minha mãe Marinete, que foi fundamental para eu ter tempo de estudar, meu muito obrigada.

Ao meu pai Raimundo e irmãos Ronne e Rosiane por todo pensamento positivo.

À minha grande parceira de mestrado, Celineide, que foi minha principal incentivadora.

Aos professores do programa que me apoiaram, em especial: Renata Lira, Fernando Assis e Franciele Redigolo.

Aos professores que se dispuseram a participar da minha pesquisa, pois sem eles, ela não seria possível.

À minha chefia, Célia Ribeiro, pela compreensão e apoio no período do mestrado.

Aos meus colegas de trabalho, que torceram por mim em todo o período do mestrado.

Agradeço a todos que de forma direta e indireta me ajudaram a concluir esta pesquisa.

## RESUMO

A Ciência Aberta pode ser considerada como uma mudança de paradigma dentro do movimento de Acesso Livre, contendo em um de seus pilares o acesso aberto aos dados de pesquisas científicas de forma a contribuir e colaborar com uso, reuso e reprodutibilidade desses dados. Mediante o compartilhamento, as pesquisas podem ser validadas e trazer mais transparência aos estudos desenvolvidos. Dentre as questões debatidas sobre o compartilhamento de dados de pesquisa está a percepção dos autores sobre essa prática, os quais podem tanto inibi-la como influenciá-la. Assim, o objetivo do presente estudo é investigar as percepções e práticas dos docentes permanentes de programas de pós-graduação do Instituto de Tecnologia (ITEC) e do Instituto de Ciências Sociais Aplicadas (ICSA) da UFPA quanto ao compartilhamento de dados de pesquisa. A metodologia, de cunho exploratório, por meio do estudo de caso, compõe a investigação quanto ao compartilhamento de dados de pesquisa, tendo a UFPA como *locus* da pesquisa, mediante questionário e entrevista. A pesquisa foi realizada por meio de 51 respondentes ao questionário e 10 participantes na entrevista. Os resultados apontaram que os docentes possuem conhecimento sobre compartilhamento de dados de pesquisa, mas não em relação aos fundamentos da Ciência Aberta. Em relação as práticas de compartilhamento de dados, a pesquisa aponta a ocorrência de maneira informal, sem a utilização de ferramentas de gerenciamento desses dados. Denota-se a importância deste estudo na UFPA, visto a incipiência sobre a temática e a necessidade de diretrizes e suporte em âmbito institucional, diante da relevância do tema em identificar e pesquisar fatores e características que podem aumentar o equilíbrio entre os participantes no processo e uso do compartilhamento de dados de pesquisa.

**Palavras-chave:** compartilhamento de dados; dados de pesquisa; ciência aberta.

## ABSTRACT

Open Science can be considered as a paradigm shift within the Open Access movement, containing in one of its pillars the open access to scientific research data in order to contribute and collaborate with the use, reuse and reproducibility of these data. Through sharing, research can be validated and bring more transparency to the studies developed. Among the issues debated about the sharing of research data is the authors' perception of this practice, which can both inhibit and influence it. Thus, the aim of this study is to investigate the perceptions and practices of permanent professors of graduate programs at the Institute of Technology (ITEC) and the Institute of Applied Social Sciences (ICSA) at UFPA regarding the sharing of research data. The methodology, of an exploratory nature, through the case study, composes the investigation regarding the sharing of research data, with the UFPA as the locus of the research, through a questionnaire and interview. The survey was conducted through 51 questionnaire respondents and 10 interview participants. The results showed that the professors have knowledge about research data sharing, but not in relation to the fundamentals of Open Science. Regarding data sharing practices, the research points to the occurrence informally, without the use of data management tools. The importance of this study at UFPA is highlighted, given the incipience on the theme and the need for guidelines and support at the institutional level, given the relevance of the theme in identifying and researching factors and characteristics that can increase the balance between the participants in the process and use of research data sharing.

**Keywords:** sharing research data; research data; open science.

## LISTA DE FIGURAS

Figura 1 - Paradigma da ciência .....	23
Figura 2 - Taxonomia da Ciência Aberta .....	25
Figura 3 - Tipologia de dados de pesquisa. ....	27
Figura 4 - Ciclo de vida de dados <i>DataONE</i> .....	34
Figura 5 - Ciclo de Vida dos Dados para Ciência da Informação – (CVD–CI) .....	35
Figura 6 - Modelo de Ciclo de Vida da curadoria dos dados da DCC .....	36
Figura 7 - População e amostra da pesquisa.....	48

## LISTA DE GRÁFICOS

Gráfico 1 - Grau de concordância com cada uma das afirmativas sobre compartilhamento de dados.....	70
Gráfico 2 - Retorno ao usarem os dados de pesquisa.....	72
Gráfico 3 - Onde você considera que seus dados de pesquisa devem ser preservados? .....	73

## LISTA DE QUADROS

Quadro 1 - Origem, formato e exemplos de dados de pesquisa. ....	28
Quadro 2 - Benefícios do gerenciamento de dados. ....	31
Quadro 3 - Principais etapas do ciclo de vida dos dados.....	33
Quadro 4 - Tópicos e questões norteadoras para elaboração de PDG.....	38
Quadro 5 - Princípios orientadores do FAIR.....	40
Quadro 6 - Relação entre os objetivos específicos e o instrumento metodológico .....	48
Quadro 7 - Panorama das percepções e práticas de compartilhamento de dados de pesquisa (ICSA) .....	80
Quadro 8 - Panorama das percepções e práticas de compartilhamento de dados de pesquisa (ITEC).....	81

## LISTA DE TABELAS

Tabela 1 - Respondentes por instituto .....	51
Tabela 2 - Faixa etária dos participantes da pesquisa.....	52
Tabela 3 - Relação de gênero entre os participantes .....	52
Tabela 4 - Tempo de envolvimento em pesquisa .....	53
Tabela 5 - Qual das alternativas melhor descreve o principal financiador da sua pesquisa.....	54
Tabela 6 - Nível de conhecimento sobre ciência aberta, dados de pesquisa e compartilhamento de dados de pesquisa .....	55
Tabela 7 - Termo que melhor descreve os tipos de dados de pesquisa por instituto.....	57
Tabela 8 - Tipos de dados de pesquisa por instituto.....	58
Tabela 9 - Volume estimado de dados produzidos em um ano .....	58
Tabela 10 - Quantidade de envolvidos na produção de dados de pesquisa.....	59
Tabela 11 - Locais de armazenamento .....	60
Tabela 12 - Você já compartilhou dados de pesquisa em acesso aberto? .....	61
Tabela 13 - Quanto dos seus dados de pesquisa você atualmente compartilha em acesso aberto? .....	62
Tabela 14 - Se você respondeu que compartilha seus dados de pesquisa em acesso aberto, onde você.....	63
Tabela 15 - Quais suas razões para não compartilhar ou compartilhar parcialmente os seus dados de pesquisa? .....	63
Tabela 16 - As seguintes afirmativas se referem ao compartilhamento e ao acesso a dados de pesquisa .....	65
Tabela 17 - Quão satisfeito você está com a organização dos seus dados de pesquisa nas seguintes situações.....	66
Tabela 18 - Qual sua familiaridade com a gestão de dados de pesquisa? .....	66
Tabela 19 - Você já elaborou um plano de gestão de dados.....	67
Tabela 20 - Você utiliza algum repositório para acessar dados de pesquisa de outros pesquisadores?.....	68
Tabela 21 - Você já utilizou em suas pesquisas dados abertos compartilhados por outros grupos de pesquisa?.....	68
Tabela 22 - Quem deve oferecer apoio a gestão de dados de pesquisa.....	74

## LISTAS SIGLAS

ITEC	Instituto de Tecnologia
BOAI	<i>Budapest Open Access Initiative</i>
CIF	Arquivo de Informações Cristalográficas
CVD	Ciclo de Vida dos Dados
CVD–CI	Ciclo de Vida dos Dados para Ciência da Informação
DCC	<i>Digital Curation Center</i>
DOI	<i>Digital Object Identifier</i>
<i>FAIR</i>	Localizável, Acessível, Interoperável e Reutilizável
FITS	Sistema de Transporte de Imagem Flexível
<i>Foster</i>	<i>Facilitate Open Science Training for European Research</i>
ICSA	Instituto de Ciências Sociais Aplicadas
PDF	<i>Portable Document Format</i>
PGD	Plano de Gestão de Dados
PPGAD	Programa de Pós-Graduação em Administração
PPGAU	Programa de Pós-Graduação em Arquitetura e Urbanismo
PPGCI	Pós-Graduação em Ciência da Informação
PPGCTA	Programa de Pós-Graduação em Ciência e Tecnologia de Alimentos
PPGE	Programa de Pós-Graduação em Economia
PPGEA	Programa de Pós-Graduação em Economia Aplicada
PPGEE	Programa de Pós-Graduação em Engenharia Elétrica
PPGEI	Programa de Pós-Graduação em Engenharia Industrial
PPGEM	Programa de Pós-Graduação em Engenharia Mecânica
PPGENAV	Programa de Pós-Graduação em Engenharia Naval
PPGEP	Programa de Pós-Graduação em Engenharia de Processos
PPGESA	Programa de Pós-Graduação em Engenharia Sanitária e Ambiental
PPGPatri	Programa de Pós-Graduação em Ciências do Patrimônio Cultural
PPGSS	Programa de Pós-Graduação em Serviço Social
PRODERNA	Programa de Pós-Graduação em Engenharia de Recursos Naturais da Amazônia
RTF	<i>Rich Text Format</i>
XML	<i>Extensible Mark-up Language</i>
ZVI	Imagem Microscópica Digital

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>15</b>
1.1 JUSTIFICATIVA.....	16
1.2 OBJETIVO .....	18
1.3 ESTRUTURA DOS CAPÍTULOS .....	19
<b>2 O MOVIMENTO DE ACESSO ABERTO E A CIÊNCIA ABERTA.....</b>	<b>20</b>
2.1 DADOS ABERTOS DE PESQUISA .....	26
2.2 GESTÃO DE DADOS DE PESQUISA.....	30
2.3 CICLO DE VIDA DOS DADOS DE PESQUISA .....	32
2.4 PLANO DE GESTÃO DE DADOS .....	37
2.5 COMPARTILHAMENTO DE DADOS DE PESQUISA .....	41
2.6 PERCEPÇÕES SOBRE COMPARTILHAMENTO DE DADOS DE PESQUISA.....	43
<b>3 METODOLOGIA.....</b>	<b>47</b>
3.1 INSTRUMENTOS DE COLETA DE DADOS.....	48
<b>4 RESULTADO E ANÁLISE DOS DADOS.....</b>	<b>51</b>
4.1 PRÁTICAS SOBRE COMPARTILHAMENTO DE DADOS DE PESQUISA .....	56
4.2 PERCEPÇÕES EM RELAÇÃO AO COMPARTILHAMENTO DE DADOS DE PESQUISA .....	69
4.3 APROFUNDAMENTO DA INVESTIGAÇÃO POR MEIO DA ENTREVISTA .....	74
<b>5 CONCLUSÕES.....</b>	<b>83</b>
<b>REFERÊNCIAS .....</b>	<b>85</b>
<b>APÊNDICE A – INTERPRETAÇÃO DO QUESTIONÁRIO.....</b>	<b>93</b>
<b>APÊNDICE B – ROTEIRO DE ENTREVISTA .....</b>	<b>110</b>

## 1 INTRODUÇÃO

A comunicação eficiente e eficaz é essencial para o desenvolvimento científico, visto que é mediante o processo da comunicação científica e por meio de publicações de resultados de pesquisas que se gera a produção do conhecimento científico. Esse processo gira em torno da acumulação do conhecimento que vem de novas observações que podem ser acrescentadas ao que havia sido produzido em um nível mais elevado, envolvendo trocas de informações que fomentam novo conhecimento. Assim, é essencial o compartilhamento de forma durável e acessível (MEADOWS, 1999; VIEIRA, 2013). Desse modo, para Mueller (2006, p. 27):

O movimento para acesso livre ao conhecimento científico pode ser considerado como o fato mais interessante e talvez importante de nossa época no que se refere à comunicação científica. Ao mesmo tempo, este movimento representa enorme desafio para a comunidade científica, à medida que, quanto mais amplo o seu sucesso, mais radical será a mudança provocada no sistema tradicional e profundamente arraigado de comunicação do conhecimento científico.

Nesse sentido, visando disponibilizar publicações científicas de forma on-line e sem restrição, surge a Ciência Aberta, que possui como um dos seus pilares os dados abertos. A pesquisa científica do século XXI produz uma quantidade exorbitante de dados, usando novos métodos e instrumentação. No entanto, os dados de pesquisa não estão disponíveis e as práticas de compartilhamento de dados podem variar entre as áreas do conhecimento. Além disso, os dados da pesquisa possuem muitos formatos, são tratados de maneiras diferentes em alguns casos usam diversas abordagens e, muitas vezes, são difíceis de interpretar quando retirados de seu contexto (BORGMAN, 2012).

O compartilhamento de dados de pesquisas com a comunidade científica, e até mesmo com o público, permite que essas informações sejam examinadas, inspecionadas e reutilizadas em diferentes conjunturas. Nesse sentido, se faz necessário um suporte tecnológico e gerencial para apoiar a execução do processo de gestão desses dados, envolvendo todo o seu ciclo de vida, como é o caso do repositório digital de dados de pesquisa (SAYÃO; SALES, 2016).

Os dados de pesquisa possuem um tempo de vida que variam conforme a finalidade de uso e a política adotada. Alguns desses modelos possuem os seus próprios objetivos e particularidades, direcionados para áreas específicas do conhecimento e são importantes, pois

oferecem estrutura que representa as operações que precisarão ser realizadas mediante o registro e o tempo de vida dos dados, garantindo que eles tenham usabilidade otimizada e estendida (SAYÃO; SALES, 2015).

Para isso, é essencial que haja um plano de gerenciamento de dados (PGD), uma ferramenta de boas práticas de gerenciamento, que permite os dados serem acessados, salvos e reutilizados em pesquisas futuras. Outro aspecto importante a se considerar é que o dado pode ser lido e operado por máquinas, permitindo integração automática com outros sistemas de informação, facilitando o preenchimento e otimizando o tempo do pesquisador (VEIGA *et al.*, 2019).

No cenário atual, a UFPA está inserida no contexto do Movimento de Acesso Aberto, por meio do seu Repositório Institucional, da Biblioteca de Teses e Dissertações e do Portal do Livro Aberto que coletam, armazenam e disseminam a produção científica produzida pela instituição.

Mediante o exposto, os dados de pesquisa agregaram novo desdobramento na cadeia de informação construída ao longo do tempo pelas universidades, ligadas ao Movimento da Ciência Aberta (GALVINO; ROSA; OLIVEIRA, 2020). No entanto, questões a serem debatidas, como infraestrutura tecnológicas, barreira organizacionais e a própria percepção dos autores são fatores que impedem o compartilhamento de dados (KIM, 2013).

Tendo em vista que essa temática ainda é pouco discutida no contexto da UFPA e frente à grande produção de dados com novos métodos e instrumentos, questiona-se: **qual a percepção e as práticas dos docentes permanentes do ITEC e ICSA sobre o compartilhamento de dados de pesquisa?**

## **1.1 Justificativa**

A pesquisa científica atual é mais intensiva em dados e colaborativa do que no passado, conforme assinala Tenopir *et al.* (2011, p. 1, tradução nossa):

[...] é importante estudar as práticas de dados dos pesquisadores - acessibilidade de dados, descoberta, reutilização, preservação e, particularmente, compartilhamento de dados. O compartilhamento de dados é uma parte valiosa do método científico, permitindo a verificação dos resultados e ampliando os resultados de pesquisa anterior.

Com o exponencial aumento na produção de dados de pesquisa e, conseqüentemente, a necessidade de um gerenciamento, armazenamento e compartilhamento adequado, percebe-se o interesse crescente das instituições de pesquisa em trabalharem na construção e aperfeiçoamento de infraestrutura tecnológica, gerencial e confiável para o armazenamento e o compartilhamento dos dados de pesquisa. Nesse sentido, Hanson, Sugden e Alberts (2011, p. 649, tradução nossa) reforçam que:

A ciência é movida por dados. As novas tecnologias aumentaram significativamente a facilidade de coleta de dados e, conseqüentemente, a quantidade de dados coletados, ao mesmo tempo que possibilitaram que os dados fossem minerados e reanalisados de forma independente por terceiros. E a sociedade agora depende de dados científicos de diversos tipos; por exemplo, na resposta a surtos de doenças, gestão de recursos, resposta à mudança climática e melhoria do transporte. É óbvio que tornar os dados amplamente disponíveis é um elemento essencial da pesquisa científica. A comunidade científica se esforça para cumprir suas responsabilidades básicas em relação à transparência, padronização e arquivamento de dados.

Córdula e Araújo (2019) dizem que é necessário reconhecer a importância dos dados para o modelo atual da ciência para aperfeiçoar a visão que caracterizava dados de pesquisa, geralmente registrados em mídia impressa ou em formatos digitais, porém esquecidos em armazenamentos locais, como meros subprodutos dos processos de pesquisa. Observam também, que na era do *e-Science*, o compartilhamento de dados científicos é ainda uma temática pouco explorada, carente de pesquisas que se proponham a: estudar, discutir e compreender as questões que a envolvem.

Adicionalmente, pontua-se que esta pesquisa se mostra ainda mais atual e relevante para o momento em que o mundo vive, a pandemia do novo coronavírus (Sars-cov-2) causador da COVID-19, em que o ambiente digital ganhou ainda mais espaço para o desenvolvimento e comunicação científica.

Assim, a Ciência da Informação pode desempenhar um papel importante dentro de uma instituição, contribuindo para o cenário de acesso e uso de dados da melhor maneira possível, buscando identificar e pesquisar fatores e características que propiciem ampliação do equilíbrio entre os participantes envolvidos no processo e a máxima otimização do seu uso (SANT'ANA, 2016).

Como profissional da informação, na qualidade de bibliotecária em exercício e participante da comunidade universitária como discente do programa de Pós-Graduação em Ciência da Informação, percebendo amplitude dessa discussão, envolvendo a Ciência Aberta

em torno de compartilhamento de dados e a importância no desenvolvimento de estudos, técnicas e aplicabilidade no âmbito institucional, foi que a autora se motivou para o desenvolvimento dessa temática.

Nas atividades de pesquisa modernas, o compartilhamento de dados científicos está cada vez mais crescente. Desse modo, se faz necessário examinar o impacto em âmbito institucional e as motivações pessoais dos cientistas para o desenvolvimento dessa prática (KIM, 2013). Compreender o compartilhamento de dados é importante também para os profissionais da informação, que podem fornecer seus conhecimentos e habilidades para curadoria, facilitando o compartilhamento e a reutilização desses dados (BORGMAN, 2010).

Tendo em vista a relevância deste estudo para o aprofundamento de discussões sobre os dados de pesquisa no contexto da UFPA, considerando ínfima essa temática e notando que a instituição ainda não possui estratégia ou plano para o armazenamento, gerenciamento e compartilhamento de dados dos seus pesquisadores, entende-se que tais estudos podem fornecer subsídio para tomada de decisão e desenvolvimento de política de gerenciamento de dados no âmbito da UFPA.

## **1.2 Objetivo**

### **1.2.1 Objetivo Geral**

Investigar as percepções e práticas dos docentes permanentes de programas de pós-graduação do Instituto de Tecnologia (ITEC) e do Instituto de Ciências Sociais Aplicadas (ICSA) da UFPA quanto ao compartilhamento de dados de pesquisa.

### **1.2.2 Objetivos Específicos**

Para construção da investigação sobre as percepções e práticas desses docentes, instituiu-se os seguintes objetivos específicos:

- a) Identificar as percepções dos docentes permanentes dos programas de pós-graduação ITEC e do ICSA sobre compartilhamento de dados de pesquisa;
- b) Mapear as práticas de compartilhamento de dados de pesquisa;
- c) Apresentar o panorama das percepções e práticas por área do conhecimento/instituto.

### **1.3 Estrutura dos capítulos**

O presente trabalho está dividido em 4 seções com suas subseções, a saber: a seção 1, denominada Introdução, trata dos aspectos gerais sobre as temáticas abordadas, bem como apresenta o Problema de pesquisa, a Justificativa, o Objetivo geral, Objetivos específicos e esta estruturação.

A seção 2, Referencial Teórico, apresenta os principais assuntos sobre a temática, como: o movimento de Acesso Aberto e a Ciência Aberta, que se divide em: Dados Abertos de Pesquisa, Gestão de Dados de pesquisa, ciclo de vida dos dados de pesquisa, plano de gestão de dados, compartilhamento de dados de pesquisa e percepções sobre compartilhamento de dados de pesquisa.

A seção 3, Metodologia, descreve os procedimentos metodológicos delineados para alcançar os objetivos desta pesquisa e é seguida pela seção 4, Resultados, a qual apresenta os resultados de pesquisa alcançados através do questionário e entrevista. Por fim, a seção 5, Considerações finais, aborda as conclusões da referida pesquisa.

## 2 O MOVIMENTO DE ACESSO ABERTO E A CIÊNCIA ABERTA

A comunicação é essencial para o desenvolvimento científico, sendo tão importante quanto a própria pesquisa, pois está ligada ao intercâmbio de informação e conhecimento, envolvendo a produção, disseminação e uso do conhecimento (LEITE, 2006; MEADOWS, 1999).

O conhecimento científico é produto das atividades relacionadas com a investigação científica, e esta última tem no processo de comunicação um elemento inerente à sua própria natureza. Comunicar o conhecimento científico significa compartilhar o conhecimento produzido por meio da investigação científica. Pressupõe um fluxo de conhecimento (tácito ou explícito) entre uma fonte e um receptor por meio de um determinado canal. Nesse sentido, os canais de comunicação são os mecanismos pelo qual o conhecimento produzido é veiculado, formal ou informalmente (LEITE, 2006, p. 67).

No processo da comunicação científica vem ocorrendo mudanças relacionadas ao contexto social e tecnológico, como as que aconteceram em função dos periódicos em meados do século XVII, das sociedades científicas no século XVIII, da profissionalização da pesquisa e da especialização em áreas do conhecimento no século XIX e do mundo eletrônico no século XX (COSTA; LEITE, 2016; MEADOWS, 1999).

A partir do século XXI dois fatores propiciaram o surgimento das iniciativas de Acesso Aberto: o primeiro foi a insatisfação dos pesquisadores com o modelo tradicional de publicação científica, com imposição de barreiras para o acesso; e o segundo, a oportunidade proporcionada pela tecnologia em gerar agilidade e dinamização no processo da comunicação científica (COSTA; LEITE, 2016).

O Movimento de Acesso Aberto à produção científica manifestou-se como uma reação ao domínio das editoras no processo editorial de periódicos científicos, em decorrência do alto custo das assinaturas, que atingiu bibliotecas e usuários, limitando-os na possibilidade de serem informados sobre o conhecimento gerado e publicado em revistas científicas, o que ocasionou a crise dos periódicos (PINHEIRO, 2014).

O Movimento surge com a articulação de pesquisadores e instituições majoritariamente de países do norte do mundo. No entanto, sua perspectiva é oportuna para os países social e historicamente excluídos do centro da ciência. Entre os benefícios do acesso aberto destaca-se a promoção do acesso à literatura científica aos pesquisadores de instituições que não teriam como pagar pelo valor cobrado pelas editoras comerciais para a assinatura de periódicos e bases de dados. Além disso, espera-se que as iniciativas ampliem

a visibilidade da produção científica disponível em acesso aberto. No que diz respeito a produção, disseminação e acesso ao conhecimento científico, se comparadas à outras regiões, a América Latina possui características singulares que favorecem a sua adesão às iniciativas de acesso aberto (COSTA; LEITE, 2016, p. 2).

As primeiras iniciativas sobre o Movimento de Acesso Aberto foram: o ArXiv, de Los Alamos, nos EUA, e o *EPrints*, de Southampton, no Reino Unido, seguidos da *Budapest Open Access Initiative* (BOAI) em 2002. O movimento foi em direção ao uso de ferramentas, estratégias e metodologias que criam um modelo para representar o novo processo de comunicação científica, ao mesmo tempo, em que servem de base para interpretá-lo nas seguintes questões (COSTA, 2006):

- **software livre:** para o desenvolvimento de aplicações em computador;
- **arquivos abertos:** para interoperabilidade em nível global;
- **acesso aberto:** para a disseminação ampla e irrestrita de resultados da pesquisa científica.

O Acesso Aberto ou *Open Access* é a disponibilização da literatura acadêmica com o acesso totalmente irrestrito e gratuito, que envolve o cientista, acadêmico, professor, estudante entre outros interessados, visando desfazer barreiras que impedem o acesso à literatura, acelerando a pesquisa, fortalecendo a educação e difundindo o conhecimento de maneira geral, extraindo dela seu máximo proveito para a construção de um novo conhecimento (BUDAPEST..., 2002; MINISTÉRIO..., c2016).

O Ministério da Ciência, Tecnologia e Ensino Superior (c2016, p. [1]) de Portugal apresenta algumas vantagens do Acesso Aberto, tais como:

- Promove e acelera o progresso da investigação e da ciência;
- Aumenta a visibilidade, o acesso, a utilização e o impacto dos resultados de investigação;
- Melhora a monitorização, avaliação e gestão da atividade científica;
- Facilita a inovação e maximiza o impacto e o retorno social e económico;
- Torna os resultados da investigação acessíveis a cidadãos e organizações.

A ciência precisa de Acesso Aberto para facilitar seu progresso e acelerar seu desenvolvimento. Nesse contexto, a métrica padrão para medir o valor de uma publicação é a citação, os artigos que são altamente citados possuem valor mensurável e maior visibilidade (SWAN, 2007).

Esses apontamentos são importantes para reforçar o alcance de um *Open* completo, com acesso ao *corpus* científico de maneira mais rápida e em âmbito global. Com isso, Costa (2006, p. 41) aponta que:

[...] para um processo cíclico em que maior acessibilidade é igual a maior uso, que significa maior fator de impacto, que, por sua vez, significa maiores recompensas, que provocam mais produção, a qual, estando mais acessível, provoca maior uso e assim sucessivamente.

Um marco para esse movimento de Acesso Aberto é a criação de cinco vias para impulsionar o alcance do *Open Access* (FAUSTO, 2013; UNIVERSIDADE DE SALAMANCA, 2019):

- **via verde:** referente aos documentos arquivados em repositórios institucionais de livre acesso;
- **via dourada:** referente à produção e ampla disseminação de periódicos eletrônicos de Acesso Aberto;
- **via de diamante:** tenta cobrir a lacuna entre os modelos verde e dourado, já que este modelo leva em consideração os dois aspectos mais críticos de ambos os modelos: custos (Ouro) e qualidade (Verde). Sendo o único modelo que garante a sustentabilidade da publicação em acesso aberto;
- **via bronze:** referente a artigos que podem ser lidos livremente nas páginas dos editores, mas sem uma licença aberta explícita que permita sua distribuição e reutilização;
- **via híbrida:** referente a artigos baseados no modelo comercial denominado “O autor paga”, onde paga para publicar, e não para ler.

A Declaração de Berlim sobre Acesso Livre ao conhecimento produzido nas ciências e humanidades ([2003], p. [1]), ampliou as fronteiras do movimento do Acesso Aberto, ao relatar que as “contribuições de acesso livre incluem resultados de investigações científicas originais, dados não processados e metadados, fontes originais, representações digitais de materiais pictóricos e gráficos e material acadêmico multimídia”.

Galvino, Rosa e Oliveira (2020) relatam que dentro dessa mudança junto ao Movimento de Acesso Aberto, temos o movimento Ciência Aberta, que visa, entre outras coisas, o acesso e a disponibilização dos dados de pesquisas científicas.

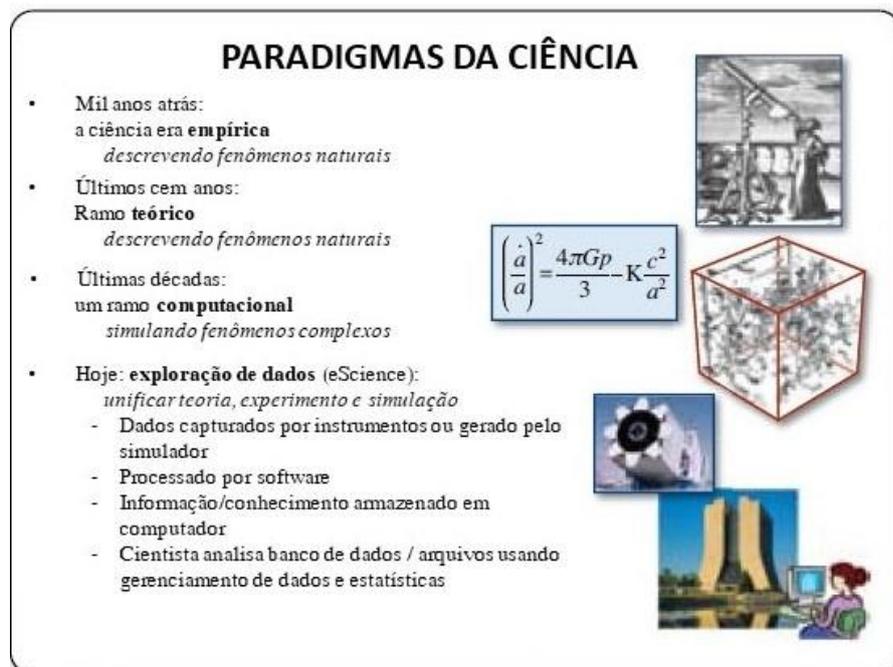
A ciência aberta é a prática da ciência que de tal forma outros podem colaborar e contribuir, onde os dados de pesquisa, notas de laboratório e outros

processos de pesquisa estão disponíveis gratuitamente, em termos que permitem a reutilização, redistribuição e reprodução da pesquisa e seus dados e métodos subjacentes (FOSTER, [20-?], p. [1], tradução nossa).

A Ciência Aberta relaciona-se com as novas formas de produção colaborativa, interativa e compartilhada da informação, do conhecimento e da cultura, possibilitando maior abertura e participação da sociedade (ALBAGLI, 2015).

Nessa perspectiva, a ciência contemporânea está pautada no quarto paradigma científico (*e-Science*): a chamada ciência dos dados. Que visa o estímulo na construção e uso de plataformas e tecnologias, envolvendo uma pluralidade de questões, como: as de ordem técnica e tecnológica, regulatórias, como as questões de proteção de direitos de propriedade intelectual, e os arranjos e mecanismos de governança que reconheçam e contribuam para lidar com as diferentes visões, culturas e interesses dos participantes, assim como a Ciência Aberta, nesse sentido, o compartilhamento aberto seria essencial para o desenvolvimento da *e-Science* (ALBAGLI; APPEL; MACIEL, 2014), conforme apresentado na Figura 1.

Figura 1 - Paradigma da ciência



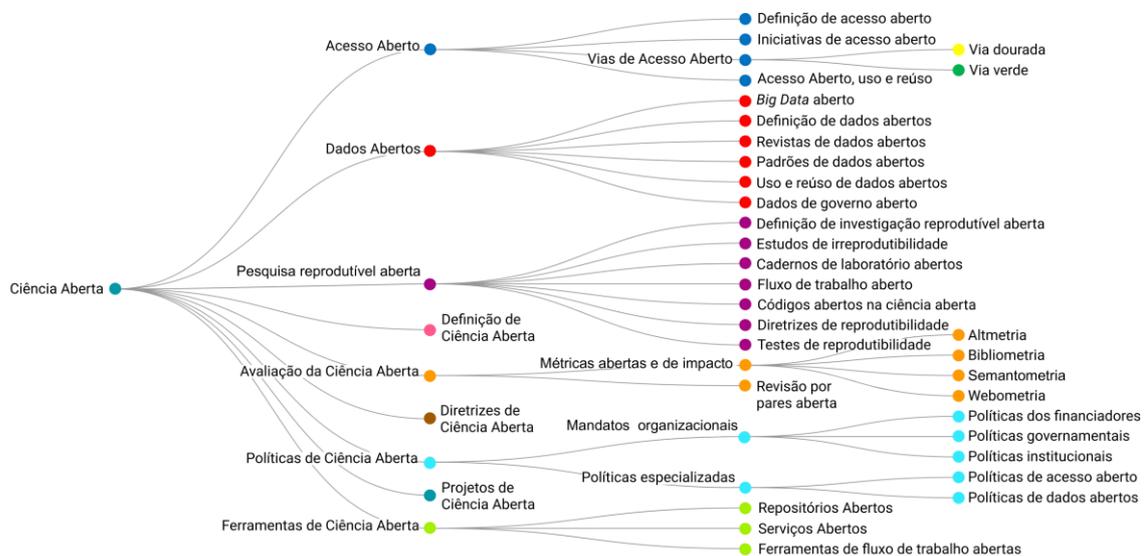
Fonte: Hey, Tansley e Tolle (2006, p. 19, tradução nossa)

Os paradigmas da ciência representam a evolução histórica e científica, conforme descreve Oliveira (2019):

- **1º paradigma:** baseia-se na experimentação, ocorrido há mais de 1000 (mil) anos, com a ciência experimental, que se originou na Grécia antiga e China, onde os filósofos e pesquisadores buscavam explicar as suas observações através de leis naturais, ao invés de causas sobrenaturais;
- **2º paradigma:** baseia-se na ciência teórica, ocorrido por volta do século XVII, cientistas como Newton tentaram fazer previsões para os novos fenômenos e verificação de novas hipóteses por meio da realização de experimentos, originando novas teorias.
- **3º paradigma:** baseia-se na ciência computacional, na simulação de fenômenos complexos. Com o advento dos computadores de alto desempenho na segunda metade do século XX.
- **4º paradigma:** baseia-se na quantidade de dados gerados no processo de pesquisa científica. A mudança é que dados brutos adjacentes de diferentes disciplinas são usados e usados como base para futuras pesquisas científicas.

No que tange à Ciência Aberta, Oliveira e Silva (2016, p. 6) dizem que a mesma “é o fio condutor de investigações científicas apoiadas por uma ciberinfraestrutura tecnológica e metodológica que permite o uso, reuso e reprodutibilidade de dados de pesquisa”. Uma maneira interessante de visualizar a Ciência Aberta é por meio das publicações do grupo *Facilitate Open Science Training for European Research (Foster)*, presente no projeto *Open Science* do Programa Comunitário de Investigação e Inovação, intitulado *Horizon 2020*. A Taxonomia criada por *Foster* reuniu um grupo de elementos que compõe as novas práticas da Ciência Aberta (ANGLADA; ABADAL, 2018; SILVA; SILVEIRA, 2019), conforme observado na Figura 2.

Figura 2 - Taxonomia da Ciência Aberta



Fonte: Ribeiro, Silveira e Santos (2020).

Diante do exposto na Figura 2, destaca-se no presente estudo, alguns elementos da taxonomia da Ciência Aberta que estão relacionados a abordagem principal desta pesquisa, a saber: os dados abertos, que envolve a questão do uso e reuso dos dados; pesquisa reprodutiva aberta, abordando diretrizes de reprodutibilidades, definição de investigação reprodutiva aberta, entre outros; e as ferramentas de ciência aberta, meios pelos quais ocorrem a difusão desse conhecimento.

Nesse cenário, Uhlir e Schröder (2007) analisaram os benefícios científicos e socioeconômicos de uma ciência mais aberta, tendo como foco o papel dos dados de pesquisas financiadas com recursos públicos. Os autores alinham algumas razões para o desenvolvimento de regimes de acesso mais abrangentes nas esferas institucionais, nacionais e internacionais, tendo o acesso livre como uma regra predominante para:

- reforçar a pesquisa científica aberta;
- incentivar a diversidade de análise e de opiniões;
- promover novas categorias de pesquisa;
- possibilitar a aplicação de ferramentas automatizadas online de descoberta de conhecimento;

- permitir a verificação de resultados prévios;
- tornar possível o teste de hipóteses e de métodos novos ou alternativos de análise;
- dar suporte a estudos sobre métodos de coleta de dados e de mensuração;
- facilitar a formação de novos pesquisadores;
- possibilitar a exploração, por outros pesquisadores, de tópicos não previstos pelos pesquisadores iniciais;
- permitir a criação de novos conjuntos de dados, de informações e de conhecimentos quando os dados de múltiplas fontes são combinados;
- ajudar a transferir informação factual para países em desenvolvimento, promovendo a capacitação de pesquisadores nesses países;
- promover a pesquisa interdisciplinar, intersetorial, interinstitucional e internacional.

Desse modo, a Ciência Aberta propõe ser aberta, gratuita, e envolve o volume intensivo de dados de pesquisa produzida. Assim, se faz necessário explanar sobre a temática de Dados Científicos Abertos, apresentada na seção a seguir.

## **2.1 Dados abertos de pesquisa**

A expansão do conceito de Acesso Aberto não está associada apenas às publicações acadêmicas tradicionais, como os artigos de periódicos. Seu progresso também inclui outros conteúdos como os dados de pesquisa. O compartilhamento com a comunidade científica e com a sociedade em geral, permite que essas informações sejam analisadas, conferidas e reutilizadas em diferentes contextos. Esse movimento em direção a uma Ciência Aberta, para além das publicações em formato livre, já se encontra evoluído no cenário internacional (SAYÃO; SALES, 2014; CAREGNATO *et al.*, 2019).

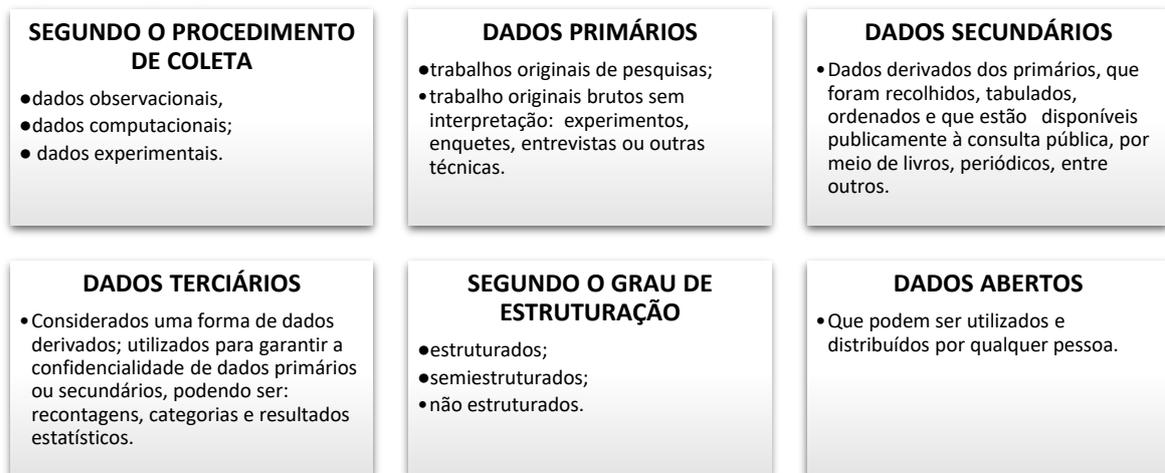
A capacidade crescente do computador, associada às possibilidades trazidas pela internet, deram lugar para novas aplicações de recursos básicos de pesquisa (os chamados dados de pesquisa), o que dá um novo impulso ao trabalho científico (SAYÃO; SALES, 2014). Além disso:

O reconhecimento do potencial informacional dos dados digitais, distribuídos em rede de computadores, para a ciência contemporânea transforma a visão

que caracterizava dados de pesquisa, registrados em mídia impressa ou mesmo em formatos digitais, como simples subprodutos dos processos de pesquisa. Nesse contexto, os dados eram considerados somente na sua configuração final, sem considerar os seus ciclos de vida, versões e linhagens e, via de regra, eram descartados ou armazenados em mídias ou servidores sem a devida gestão quando os projetos eram concluídos. Quase sempre eram tragados silenciosamente pelo tempo: pela obsolescência tecnológica, pela efemeridade dos formatos e pela fragilidade das mídias digitais (SAYÃO; SALES, 2014, p.78).

Os dados de pesquisa possuem informações cruciais para a composição dos resultados científicos. Nesse sentido, o Relatório da *Organisation for Economic Co-operation and Development* (2007, p. 13, tradução nossa) os descreve como: “registros factuais usados como fonte primária para a pesquisa científica e que são comumente aceitos pelos pesquisadores como necessários para validar os resultados do trabalho científico”. A seguir, a Figura 3, apresenta como Silva (2019) caracteriza a tipologia dos dados de pesquisa.

Figura 3 - Tipologia de dados de pesquisa.



Fonte: Adaptado de Silva (2019).

Mediante a Figura 3, percebe-se que a amplitude sobre o entendimento de dados de pesquisa sugere um conceito complexo que pode se manifestar numa multiplicidade de formas. A noção de dados pode variar entre pesquisadores e principalmente entre áreas do conhecimento. A constatação de que os dados são gerados para diferentes propósitos, por comunidades acadêmicas e científicas distintas e por processos díspares, intensifica ainda mais essa percepção de diversidade (SAYÃO; SALES, 2015).

Para Sayão e Sales (2015), os tipos de dados podem ser: números, imagens, textos, vídeos, áudio, software, algoritmos, equações, animações, modelos e simulações. Certos tipos

de dados têm valor imediato e duradouro, enquanto outros adquirem valor ao longo do tempo. Alguns dados são capturados em momento específico e irrecuperável, enquanto outros são passíveis de se reproduzir. De acordo com *National Science Board* (2005) os dados de pesquisa surgem de várias formas diferentes que incluem sua natureza, reprodutibilidade e o nível de processamento em que foi submetido. Eles também podem ser distinguidos por suas origens, que podem ser: observacionais, computacionais ou experimentais. Essas diferenciações são importantes para o arquivamento e preservação do conhecimento gerado.

Nessa direção, a *Imperial College London* (c2020) caracteriza os dados segundo sua origem, formato e exemplos. Conforme apresentado no Quadro 1:

Quadro 1 - Origem, formato e exemplos de dados de pesquisa.

<b>ORIGEM</b> Os dados de pesquisa podem ser gerados ou coletados para diferentes fins e por meio de diferentes processos	<b>FORMATO</b> Os dados de pesquisa vêm em muitos formatos variados	<b>EXEMPLOS</b> Os dados de pesquisa (pesquisa tradicional e eletrônica) podem incluir todos os seguintes:
<p><b>OBSERVACIONAL:</b> dados capturados em tempo real, geralmente insubstituíveis, por exemplo, dados do sensor, dados de pesquisa, dados de amostra, neuroimagens;</p> <p><b>EXPERIMENTAL:</b> dados de equipamentos de laboratório, muitas vezes reproduzíveis, mas podem ser caros para reproduzir, por exemplo, sequências de genes, cromatogramas, dados de campo magnético toróide;</p> <p><b>SIMULAÇÃO:</b> dados gerados a partir de modelos de teste onde o modelo e os metadados são mais importantes do que os dados de saída, por exemplo, modelos climáticos, modelos econômicos;</p> <p><b>DERIVADO OU COMPILADO:</b> os dados são reproduzíveis, mas caros, por exemplo, mineração de texto e dados, banco de dados compilado, modelos 3D;</p> <p><b>REFERÊNCIA OU CANÔNICO:</b> um conglomerado (estático ou orgânico) ou coleção de conjuntos de dados menores (revisados por pares) provavelmente publicados e com curadoria, por exemplo, bancos de dados de sequência de genes, estruturas químicas, portais de dados espaciais.</p>	<p><b>DISCIPLINA ESPECÍFICA:</b> Sistema de Transporte de Imagem Flexível (FITS) em astronomia, Arquivo de Informações Cristalográficas (CIF) em química;</p> <p><b>INSTRUMENTAL ESPECÍFICO:</b> Formato de dados do microscópio confocal, formato de imagem microscópica digital (ZVI);</p> <p><b>MODELOS:</b> 3D, estatísticos;</p> <p><b>MULTIMÍDIA:</b> jpeg, tiff, dicom, mpeg, quicktime;</p> <p><b>NUMÉRICO:</b> Pacote Estatístico para Ciências Sociais (SPSS), Stata, Excel;</p> <p><b>SOFTWARE:</b> Java e C;</p> <p><b>TEXTO:</b> Arquivos de texto simples, Word, Portable Document Format (PDF), Rich Text Format (RTF), Extensible Mark-up Language (XML)</p>	<p>Fitas de áudio, fitas de vídeo; Coleção de objetos digitais adquiridos e gerados durante o processo de pesquisa; Conteúdo de um aplicativo de software (entrada, saída, arquivos de log para software de análise, código-fonte, esquemas); Arquivos de dados; Conteúdo do banco de dados (vídeo, áudio, texto, imagens); Documentos (texto, Word), planilhas; Cadernos de laboratório, cadernos de campo, diários; Metodologias e fluxos de trabalho; Modelos, algoritmos, scripts; Fotos, filmes; Questionários, transcrições, livros de código; Slides, artefatos, espécimes, amostras; Procedimentos e protocolos operacionais padrão; Respostas de teste.</p>

Fonte: Adaptado de *Imperial College London* (c2020).

O Quadro 1 mostra os tipos de origem, dados e exemplos de dados da pesquisa que são as evidências e sustentam os resultados de pesquisa. Os dados e conjuntos de dados de pesquisas geram informações necessárias para conferir a veracidade, autenticidade e capacidade de reprodutibilidade do conhecimento publicado (SAYÃO; SALES, 2014).

O *Imperial College London* (c2020) reforça essa ideia ao relatar que os dados de pesquisa fundamentam as conclusões de pesquisa (exceto aquelas que são puramente teóricas) e incluem dados que foram coletados, observados, gerados, criados ou obtidos de fontes comerciais, governamentais ou outras, para posterior análise e síntese para produzir resultado de pesquisa original. Assim, esses resultados podem ser usados para gerar outras pesquisas científicas.

Quanto maior for a capacidade dos sistemas de informação de oferecer dados de pesquisas de acesso livre e tratados por metadados, de forma que possam ser interpretados e reutilizados pelo maior número possível de pesquisadores de diferentes áreas, maior será o grau de: transparência, reprodutibilidade e eficiência do processo de geração de conhecimento científico, assim como, a amplitude de aplicação dos projetos de pesquisa para a sociedade também será mais extensa (SAYÃO; SALES, 2014).

Portanto, o potencial relativo aos dados redesenha, por meio do reuso, os fluxos tradicionais de comunicação científica, estabelecendo novos padrões de socialização e de trabalho cooperativo independente de barreiras geográficas e disciplinares. O valor do dado de pesquisa está diretamente relacionado à possibilidade de uso e ao seu potencial de ser reinterpretado em outras áreas e contextos diferentes da que originalmente o gerou (SAYÃO; SALES, 2014).

Diante o exposto, percebe-se a importância dos dados de pesquisa e seu armazenamento. No entanto, para que ocorra de maneira eficiente e eficaz é necessário um bom gerenciamento de dados. Cumpre ressaltar que em algumas discussões os seguintes termos são utilizados: dados científicos, dados de pesquisa e dados de pesquisa científica. Neste estudo, adotou-se o termo dados de pesquisa.

## 2.2 Gestão de dados de pesquisa

Na ciência contemporânea, Sales e Sayão (2018) salientam que os dados deixaram de ser meros subprodutos de pesquisa e passaram a ser protagonistas das novas metodologias de pesquisa científica. As tecnologias digitais e as redes de computadores aceleraram e transformaram os processos das pesquisas, tornando-as mais transparentes e abertas.

Portanto, é essencial o gerenciamento dos dados de pesquisa, não apenas com o propósito de organizá-los, acessá-los e usá-los, mas como uma atitude sustentável e responsável de se conduzir a pesquisa científica. Ressalta-se, que o pesquisador é quem decidirá o que fazer com os seus dados de pesquisa, isto é, se os manterá fechados e em acesso restrito, ou aberto e livre para todos (LIMA, 2020).

Para ilustrar esse gerenciamento, apresenta-se o caso da Universidade de São Paulo ([2019]) que define a gestão de dados científicos como um conjunto de atividades que visa coletar, armazenar, gerenciar e compartilhar dados oriundos de pesquisa científica. A gestão eficaz viabiliza a racionalização de recursos, mediante o reuso e compartilhamento de dados com a finalidade de auxiliar pesquisadores em relação ao planejamento, organização e segurança; documentação e compartilhamento; preparação dos conjuntos de dados para depósito; preservação dos dados e questões relacionadas a direitos autorais, licenciamento, propriedade intelectual interesses comerciais legítimos, privacidade e questões éticas.

As diferentes áreas científicas possuem padrões, práticas e políticas distintas em relação aos seus dados de pesquisa. O estabelecimento de infraestrutura técnica e gerencial que facilitem a integração dos conjuntos de dados de diferentes domínios e a criação de meios de colaboração entre as diferentes comunidades, são importantes para o efetivo uso e reuso desses recursos (SAYÃO; SALES, 2014)

O *Imperial College London* (c2020) relata que boas práticas de gerenciamento de dados de pesquisa trazem muitos benefícios para o pesquisador, sua disciplina e sociedade. No entanto, manter o controle dos mesmos torna-se cada vez mais difícil à medida que novos dados são coletados e os já existentes são processados. O bom gerenciamento de dados de pesquisa pode reduzir o fardo de corrigir ou desvendar problemas no futuro e leva a cinco benefícios, conforme apresenta o Quadro 2:

Quadro 2 - Benefícios do gerenciamento de dados.

<b>BENEFÍCIOS</b>	<b>DESCRIÇÃO</b>
<b>FRUSTRAÇÃO REDUZIDA</b>	<p>O gerenciamento de dados torna mais fácil:</p> <ul style="list-style-type: none"> <li>• escrever artigos e teses usando dados registrados e documentados de forma consistente ao longo de um projeto;</li> <li>• continuar usando os dados após a saída do pesquisador responsável do colégio;</li> <li>• escolha dados para arquivamento de longo prazo e descarte para economizar espaço.</li> </ul>
<b>RISCO REDUZIDO</b>	<p>Gerenciar dados ajuda a prevenir:</p> <ul style="list-style-type: none"> <li>• perda de dados;</li> <li>• incapacidade de provar os resultados da pesquisa;</li> <li>• repetição cara da coleta de dados;</li> <li>• violações acidentais de privacidade e legislação ética;</li> <li>• incapacidade de apoiar a comercialização de resultados de pesquisa.</li> </ul>
<b>MELHOR QUALIDADE DE PESQUISA</b>	<p>A robustez e integridade de seus dados permitirá que pesquisas futuras sejam construídas em bases sólidas.</p>
<b>REPUTAÇÃO MELHORADA</b>	<p>O Acesso Aberto a dados e informações é um tema em expansão em todas as áreas da sociedade. A transparência da pesquisa com financiamento público é importante tanto no nível individual quanto institucional. Para o acadêmico individual, dados bem gerenciados podem ser compartilhados com confiança, levando a:</p> <ul style="list-style-type: none"> <li>• citações adicionais de artigos e conjuntos de dados;</li> <li>• impacto demonstrável por meio do uso comercial de dados;</li> <li>• percepções inesperadas de novas técnicas e combinação de conjuntos de dados.</li> </ul>
<b>FINANCIADOR, EDITOR E CONFORMIDADE COM A POLÍTICA INSTITUCIONAL</b>	<p>Pelos benefícios citados anteriormente, os financiadores e editores de pesquisas, possuem políticas que exigem que os dados sejam gerenciados adequadamente. Alguns financiadores estão começando a considerar o gerenciamento de dados anterior ao conceder novos financiamentos, enquanto outros estão retendo os pagamentos finais do subsídio se os dados não forem gerenciados adequadamente.</p>

Fonte: Adaptado de *Imperial College London* (c2020).

Dado o exposto no Quadro 2, percebe-se a importância de boas práticas de gerenciamento dos dados de pesquisa desde seu início, visando minimizar frustração, reduzir riscos, melhorar a qualidade e reputação da pesquisa, assim como trazer benefícios para financiadores, editores e para o próprio autor.

Nesse aspecto, *Foster* (2018) apresenta uma taxonomia sobre a gestão de dados de pesquisa, com cinco elementos que contribuem para essa gestão: Gestão de Dados de Pesquisa, Políticas de Gestão de Dados de Pesquisa, Serviços de Gestão de Dados de Pesquisa, Padrões em Gestão de Dados de Pesquisa e Ferramentas em Gestão de Dados de Pesquisa, recursos valiosos para governança, segurança e qualidade dos dados.

*Foster* (2018) apresenta ainda a taxonomia das questões legais ligadas ao direito de propriedade intelectual, leis e licenças, que lidam com os problemas de saber se determinado comportamento está ou não em conformidade com a lei. Tendo em vista as questões éticas e legais, nem todo tipo de dado deve ser disponibilizado, conforme reforçado no documento *The state of open data* (2017, p. 16, tradução nossa), ao afirmar que no desenvolvimento das políticas de compartilhamento de dados de pesquisa “os resultados da pesquisa acadêmica devem ser abertos quando possível e fechados quando necessário”.

Dessa forma, desenvolver diretrizes e boas práticas de gerenciamento de dados possibilita o seu compartilhamento, portanto, podem ser usados para novas pesquisas ou para duplicar e validar pesquisas existentes ao longo de sua vida. O gerenciamento de dados precisa ser planejado no início da pesquisa para que as práticas possam ser implementadas ao longo do ciclo de pesquisa (*UK DATA SERVICE*, c2020).

### **2.3 Ciclo de vida dos dados de pesquisa**

Segundo *Medeiros* (2018) o Ciclo de Vida dos Dados (CVD) cobre a gestão de dados, que vai desde sua coleta até o armazenamento de longo prazo, perpassando por uma série de processamento de limpeza, curadoria, anotação, indexação e transformação, enquanto *UK Data Service* (c2020) afirma que alguns dados possuem vida útil mais longa do que a pesquisa que os cria, visto que os pesquisadores podem continuar a trabalhar nos dados após o término do financiamento e os projetos de acompanhamento podem analisar ou adicionar dados aos dados e os dados podem ser reutilizados por outros pesquisadores, reiniciando, dessa forma, um novo ciclo.

Assim, *Aventurier* (2016) relata que o ciclo de vida dos dados de pesquisa possui várias representações, envolvendo as principais etapas, destacando 5 elementos, conforme apresentado no Quadro 3.

Quadro 3 - Principais etapas do ciclo de vida dos dados

ETAPAS	DESCRIÇÃO
<b>Criação dos dados</b>	<ul style="list-style-type: none"> <li>• Concepção da pesquisa</li> <li>• Plano de gestão de dados (formato e estocamento)</li> <li>• Plano para compartilhar os dados</li> <li>• Localização dos dados existentes</li> <li>• Coletar dados (experimentações, observações, medir, simulação é modelização)</li> <li>• Capturamento e criação de metadados seguindo os padrões do domínio</li> </ul>
<b>Tratamento dos dados</b>	<ul style="list-style-type: none"> <li>• Transcrição, tradução, entrada de dados, digitalização,</li> <li>• Verificação, validação e limpeza de dados,</li> <li>• Anonimização dos dados quando é necessário</li> <li>• Descrição dos dados</li> <li>• Gestão e estocagem dos dados</li> </ul>
<b>Análise dos dados</b>	<ul style="list-style-type: none"> <li>• Interpretação dos dados</li> <li>• Derivação dos dados</li> <li>• Produzir resultados</li> <li>• Preparar os dados para acompanhar a publicação científica</li> <li>• Preparar os dados para preservação</li> </ul>
<b>Preservação dos dados</b>	<ul style="list-style-type: none"> <li>• Migração dos dados pelo formato mais adequado (em geral um formato aberto ou não proprietário)</li> <li>• Migração dos dados no supor mais adequado</li> <li>• Backup e armazenamento de dados</li> <li>• Criação de metadados e documentação</li> <li>• Arquivar os dados</li> </ul>
<b>Acesso aos dados</b>	<ul style="list-style-type: none"> <li>• Distribuir os dados</li> <li>• Compartilhar dados</li> <li>• Controle de acesso</li> <li>• Aspectos jurídicos de reutilização</li> <li>• Promoção dos dados</li> </ul>
<b>Reutilização dos dados</b>	<ul style="list-style-type: none"> <li>• Acompanhamento</li> <li>• Nova pesquisa</li> <li>• Fazer revisão ou síntese sobre pesquisa</li> <li>• Examinar os resultados</li> <li>• Ensinar e aprender</li> </ul>

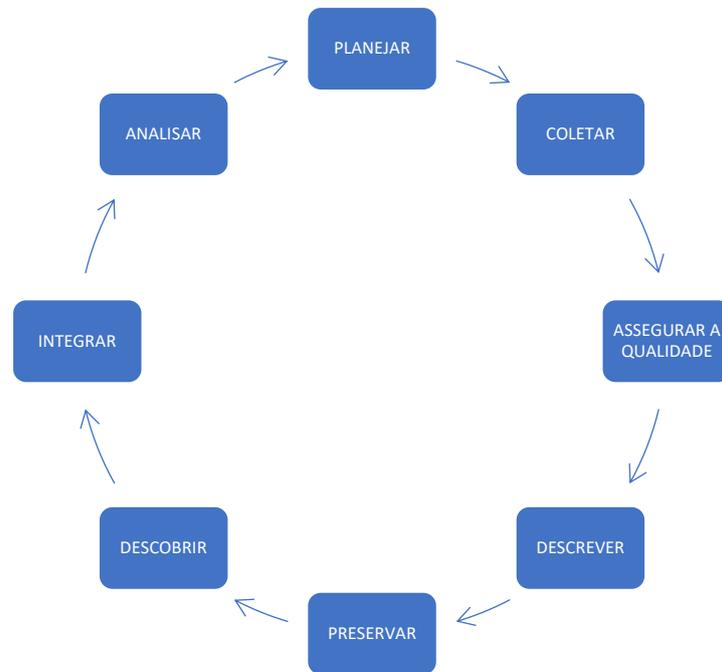
Fonte: Adaptado de Aventurier (2016)

Desse modo, cada etapa do ciclo de vida é fundamental para que os dados sejam organizados e documentados, possibilitando assim a preservação e compartilhamento, que são cruciais para a investigação científica, reforçando o aprendizado e inovação.

Assim sendo, destaca-se no presente estudo, que existem modelos de ciclo de vida dos dados com a finalidade de sistematizar as etapas, recomendar boas práticas e oferecer uma visão dos processos que devem ser realizados, cada um com atribuições e variações distintas, como mostra os modelos apresentados nas Figuras 4, 5 e 6:

O modelo apresentado pela *DataONE* aborda os aspectos mais abrangentes ligados a gestão dos dados para uma preservação, utilização e reutilização bem sucedida, conforme apresenta a Figura 4.

Figura 4 - Ciclo de vida de dados *DataONE*



Fonte: Adaptado de *DataONE* ([2019]).

Como observado na Figura 4 o modelo da *DataONE*, possui oito componentes do ciclo de vida dos dados de pesquisa do *DataONE* ([2019])

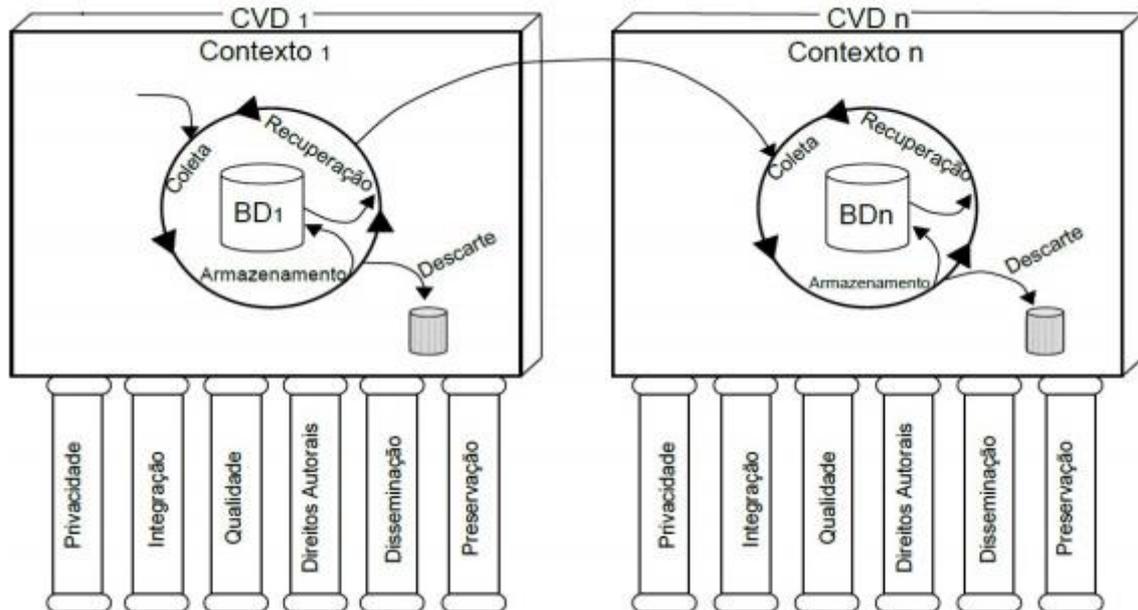
- **Planejar (*Plan*):** é a descrição dos dados que serão compilados e como os dados serão gerenciados e acessíveis ao longo de sua vida;
- **Coletar (*Collect*):** são as observações feitas manualmente, com sensores ou outros instrumentos em que os dados são colocados em formato digital;
- **Assegurar (*Assure*):** garante a qualidade dos dados por meio da verificação e inspeção;
- **Descrever (*Describe*):** é o componente em que os dados são descritos de forma precisa e completa usando padrões de metadados apropriado;
- **Preservar (*Preserve*):** é a etapa em que os dados são submetidos a um arquivo de longo prazo apropriado (ou seja, Centro de dados);

- **Descobrir (*Discover*):** é o processo em que os dados potencialmente úteis são localizados e obtidos, com as informações relevantes sobre os dados (metadados);
- **Integrar (*Integrate*)** é a fase em que os dados de fontes distintas são combinados para formar um conjunto homogêneo de dados que pode ser facilmente analisado;
- **Analisar (*Analyze*)** é a análise dos dados.

As etapas do ciclo de vida dos dados de pesquisa podem ser usadas de forma não linear ou apenas em parte, além disso, cientistas ou equipes podem criar novos dados no processo de descoberta, integração, análise e síntese de dados existentes.

O modelo apresentado por Sant’Ana (2016) considera as características da Ciência da Informação, por possuir um papel importante nesta temática e oferece um arcabouço teórico capaz de contribuir na construção de novos conhecimentos e técnicas, como é possível observar na Figura 5.

Figura 5 - Ciclo de Vida dos Dados para Ciência da Informação – (CVD-CI)

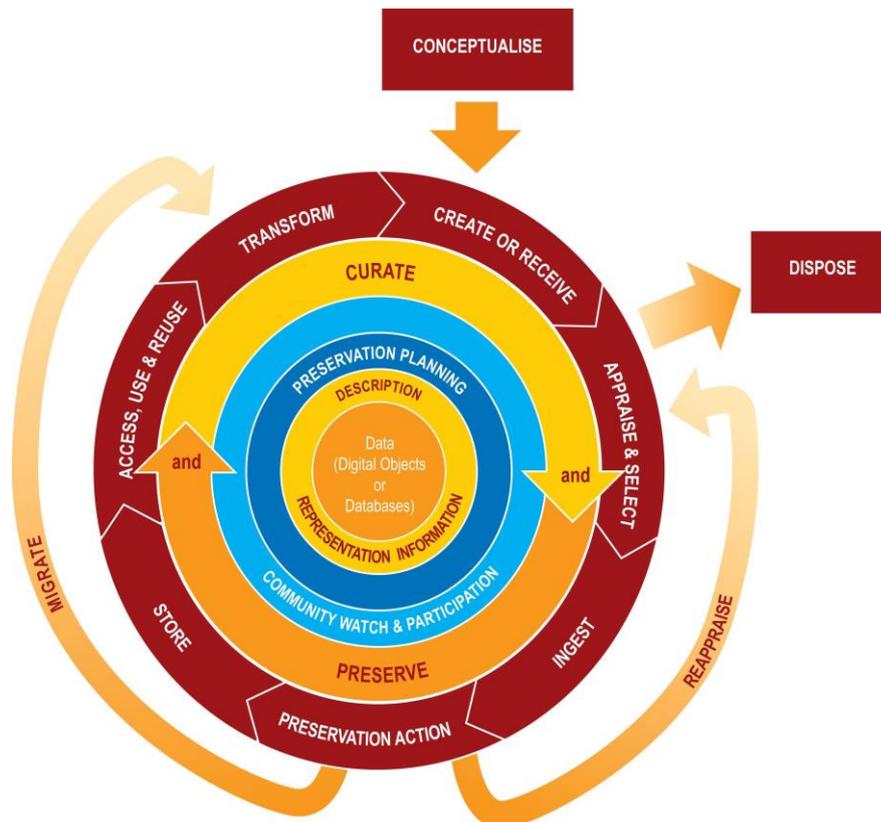


Fonte: Sant’Ana (2016)

O modelo exposto na Figura 5, contém quatro fases (coleta, armazenamento, recuperação e descarte) e seis fatores (privacidade, integração, qualidade, direitos autorais, disseminação e preservação) que sintetizam as atividades que envolve o acesso, uso e manutenção dos dados.

Já o modelo apresentado pela *Digital Curation Center* (DCC) aborda o ciclo de vida de curadoria, conforme apresenta a Figura 6:

Figura 6 - Modelo de Ciclo de Vida da curadoria dos dados da DCC



Fonte: Higgins (2008)

Esse modelo fornece uma visão geral gráfica de alto nível dos estágios necessários para curadoria bem sucedida e preservação de dados desde o início e pode ser usado para planejar atividades em uma organização, garantindo que todas as etapas necessárias sejam realizadas em sequência correta. Esse modelo permite que a funcionalidade seja mapeada e possibilita a definição de funções e responsabilidades, visando uma estrutura de padrões e tecnologias a serem implementadas, com potencial para ajudar no processo de identificação de etapas adicionais que podem ser necessárias, ou ações que não são exigidas por certas situações ou disciplinas e garantindo que os processos e políticas sejam devidamente documentados (HIGGINS, 2008).

Os modelos de Ciclo de Vida dos Dados e da Curadoria auxiliam e contribuem para assegurar a preservação, acesso, uso e reuso dos dados de pesquisa, tornando a gestão dos dados mais eficiente.

## 2.4 Plano de Gestão de Dados

O Plano de Gestão de Dados (PGD) “é um documento formal, dinâmico e atualizável com perguntas orientadoras que estimulam o pesquisador a planejar, de maneira intencional, como se dará a gestão de dados ao longo de todo o seu ciclo de vida - período que é maior que a execução e finalização da pesquisa” (FIOCRUZ, [20--], p. [1]).

Para a efetiva gestão de dados de pesquisa, o planejamento é uma fase essencial. Ele se inicia quando a pesquisa ainda está sendo delineada e deve considerar como os dados serão gerenciados durante o desenvolvimento do projeto e como eles serão compartilhados depois. Dessa forma é necessário formalizar as ações e compromissos que serão estabelecidos em relação aos dados desde os seus primeiros estágios. (SAYÃO; SALES, 2015, p. 15).

O PGD consiste no instrumento de formalização do gerenciamento de dados de pesquisa. Nele é descrito o ciclo de vida dos dados para que eles sejam coletados, processados e/ou gerados. Com base na *European Commission* (2016, p. 4), um PGD deve incluir informação sobre:

- o tratamento de dados de pesquisa durante e após o final do projeto;
- indicação de quais dados serão coletados, processados e / ou gerados;
- indicação de qual metodologia e padrões serão aplicados;
- indicação se os dados serão compartilhados / tornados em Acesso Aberto;
- como os dados serão curados e preservados (inclusive após o final do projeto).

Para Sayão e Sales (2015), o PDG deve ser tratado como uma carta de intenção, em que se considera o que realmente é necessário para a preservação, compartilhamento e reuso dos dados, visto que ele não é um documento burocrático e não deve ser pensado apenas como mais uma tarefa administrativa com um texto padrão para todos os projetos.

Nessa direção, a *Science Europe* (2018) apresenta seis requisitos essenciais para o desenvolvimento de um PDG sólido, expostos no Quadro 4.

Quadro 4 - Tópicos e questões norteadoras para elaboração de PDG.

TÓPICOS	QUESTÕES DE ORIENTAÇÃO
1. Descrição de dados e coleta ou reutilização de dados existentes	<ul style="list-style-type: none"> <li>• como os novos dados serão coletados ou produzidos e / ou como os dados existentes serão reutilizados?</li> <li>• quais dados (por exemplo, os tipos, formatos e volumes) será coletado ou produzido?</li> </ul>
2. Documentação e qualidade de dados	<ul style="list-style-type: none"> <li>• quais metadados e documentação (por exemplo, a metodologia de coleta de dados e forma de organização dados) acompanharão os dados?</li> <li>• quais medidas de controle de qualidade de dados serão utilizadas?</li> </ul>
3. Armazenamento e backup durante o processo de pesquisa	<ul style="list-style-type: none"> <li>• como os dados e metadados serão armazenados e copiados durante o processo de pesquisa?</li> <li>• como será a segurança de dados e proteção de dados confidenciais atendidos durante a pesquisa?</li> </ul>
4. Requisitos legais e éticos, códigos de conduta	<ul style="list-style-type: none"> <li>• se os dados pessoais forem processados, como será o cumprimento da legislação sobre dados pessoais para que a segurança de dados seja garantida?</li> <li>• como outras questões legais, como propriedade intelectual direitos e propriedade, será gerenciado? Qual legislação é aplicável?</li> <li>• como possíveis questões éticas, códigos de conduta serão levadas em consideração?</li> </ul>
5. Compartilhamento de dados e preservação de longo prazo	<ul style="list-style-type: none"> <li>• como e quando os dados serão compartilhados? São possíveis restrições ao compartilhamento de dados ou motivos de embargo?</li> <li>• como os dados para preservação serão selecionados e onde os dados serão preservados a longo prazo (por exemplo, um repositório de dados ou arquivo)?</li> <li>• quais métodos ou ferramentas de software serão necessários para acessar e usar os dados?</li> <li>• como será a aplicação de um identificador único e persistente (como um <i>Digital Object Identifier</i> (DOI)) para cada conjunto de dados ser assegurado?</li> </ul>
6. Responsabilidades e recursos de gerenciamento de dados	<ul style="list-style-type: none"> <li>• quem (por exemplo, função, posição e instituição) será responsável pelo gerenciamento de dados (ou seja, o administrador de dados)?</li> <li>• quais recursos (por exemplo, financeiros e de tempo) serão dedicados ao gerenciamento de dados garantindo que os dados serão <i>FAIR</i> (localizável, acessível, interoperável e reutilizável)?</li> </ul>

Fonte: Adaptada de *Science Europe* (2018).

Diante o exposto no Quadro 4, nota-se que esses elementos e suas descrições são requisitos essenciais para que o pesquisador desenvolva um PDG eficiente. Esse documento não deve ser burocrático, mas sim um meio adequado para planejar e conduzir a produção de dados de pesquisa.

Destaca-se, também, as principais razões para criação de um PDG, com base em Sayão e Sales (2015, p. 16):

- Ajustar o seu projeto de pesquisa às políticas mandatórias da sua instituição e/ou dos órgãos de fomento à pesquisa;
- Assegurar a integridade da pesquisa e o seu potencial de replicação;
- Assegurar que os dados e demais registros de pesquisa sejam acurados, completos, autênticos e confiáveis;
- Aumentar a sua eficiência como pesquisador – um plano que organize os dados e seu armazenamento permite que você foque na sua pesquisa. Você

estará mais capacitado a localizar e usar os seus dados e compartilhá-los com os seus colaboradores;

- Permitir que os seus dados sejam compreensíveis agora e no futuro – se os dados são bem documentados antes e durante a formação da coleção de dados, eles serão mais facilmente entendidos e reutilizados;
- Economizar tempo e recursos a longo prazo;
- Aumentar a segurança dos dados e minimizar os riscos de perda;
- Evitar a duplicação de esforços na coleta ou regeneração dos dados, possibilitando que outros pesquisadores se beneficiem dos seus dados e os interprete em outros contextos e com novas visões;
- Aumentar a visibilidade da pesquisa – se os seus dados foram planejados para estarem organizados e corretamente arquivados, eles poderão ser identificados, recuperados e citados, aumentando a visibilidade da sua pesquisa e o seu prestígio como pesquisador;
- Tornar mais fácil a preservação e o arquivamento – ajustando antecipadamente a geração de dados tomando como referência as práticas, formatos e demais padrões mais adequados ao arquivamento e à preservação de longo prazo, torna a gestão de dados mais fácil e menos custosa; além do mais, tornam os dados mais aderentes aos requisitos dos repositórios e centros de dados.

Percebendo a importância desse planejamento, a comunidade científica busca tornar os PDGs acionáveis por máquinas, agregando e ampliando o reconhecimento das partes interessadas: pesquisadores, financiadores, gerentes de repositório, administradores de pesquisa e outros. Permitindo que parte do PDG seja gerado e compartilhado automaticamente, dessa forma reduzindo as tarefas administrativas e melhorando a qualidade da informação dentro de um PDG (MIKSA *et al.* 2019). Nesse sentido, surgiram ferramentas online para a geração desse instrumento, como: *DMPonline* da *Digital Curation Center* (DCC) ou *Datastewarship Wizard* que visa PGDs legíveis por máquina (MARÍN-ARRAIZA; PUERTA-DÍAZ; GREGORIO-VIDOTTI, 2019).

Ademais, a gestão de dados de pesquisa precisa estar articulada para atender aos princípios **FAIR** (*Findable, Accessible, Interoperabl e Reusable*) que enfatiza a capacidade de ação da máquina, isto é, a capacidade do sistema computacional encontrar, acessar, interoperar e reutilizar os dados com nenhuma ou mínima intervenção humana, pois, depende-se cada vez mais do suporte tecnológico para lidar com os dados como resultado do aumento no volume, complexidade e velocidade de criação dos mesmos (*GO FAIR*, [20--]). A seguir, o Quadro 5 apresenta os quatro princípios *FAIR*.

Quadro 5 - Princípios orientadores do *FAIR*.

PRINCÍPIOS	INDICADORES
<p style="text-align: center;"><b><i>FINDABLE / LOCALIZÁVEL</i></b></p> <p>O primeiro passo para (re) usar dados é encontrá-los. Metadados e dados devem ser fáceis de encontrar, tanto para humanos quanto para computadores. Metadados legíveis por máquina são essenciais para a descoberta automática de conjuntos de dados e serviços</p>	F1. (meta) dados são atribuídos a um identificador globalmente único e persistente;
	F2. os dados são descritos com metadados ricos (definidos por R1 abaixo);
	F3. metadados incluem clara e explicitamente o identificador dos dados que descreve;
	F4. (meta) dados são registrados ou indexados em um recurso pesquisável.
<p style="text-align: center;"><b><i>ACCESSIBLE / ACESSÍVEL</i></b></p> <p>Uma vez que o usuário encontra os dados necessários, ela / ele precisa saber como eles podem ser acessados, possivelmente incluindo autenticação e autorização.</p>	A1. (meta) dados são recuperáveis por seu identificador usando um protocolo de comunicação padronizado;
	A1.1 o protocolo é aberto, gratuito e universalmente implementável;
	A1.2 o protocolo permite um procedimento de autenticação e autorização, quando necessário;
	A2. metadados são acessíveis, mesmo quando os dados não estão mais disponíveis.
<p style="text-align: center;"><b><i>INTEROPERABLE / INTEROPERÁVEL</i></b></p> <p>Os dados geralmente precisam ser integrados com outros dados, além disso, os dados precisam interoperar com aplicativos ou fluxos de trabalho para análise, armazenamento e processamento.</p>	I1. (meta) dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento;
	I2. (meta) dados usam vocabulários que seguem os princípios FAIR;
	I3. (meta) dados incluem referências qualificadas a outros (meta) dados.
<p style="text-align: center;"><b><i>REUSABLE / REUSÁVEL</i></b></p> <p>O objetivo final do FAIR é otimizar a reutilização de dados. Para conseguir isso, metadados e dados devem ser bem descritos para que possam ser replicados e / ou combinados em diferentes configurações.</p>	R1. meta (dados) são ricamente descritos com uma pluralidade de atributos precisos e relevantes;
	R1.1. (meta) dados são liberados com uma licença de uso de dados clara e acessível;
	R1.2. (meta) dados estão associados à proveniência detalhada;
	R1.3. (meta) dados atendem aos padrões da comunidade relevantes para o domínio.

Fonte: Adaptado de *Go FAIR* ([20--]) e Wilkinson *et al.* (2016).

Diante do exposto no Quadro 5, percebe-se que os princípios *FAIR* exigem mudanças em cultura e prática de pesquisa e a implementação de um ecossistema com serviços e componentes de dados.

Portanto, compreende-se que o desenvolvimento e implementação de um PDG é o meio pelo qual os projetos de investigação conseguem garantir a produção de dados *FAIR*, visto que é através dele que são definidas todas as questões relacionadas com a descoberta, acesso, interoperabilidade e reutilização dos dados produzidos no projeto. Para isso, o PDG deve estar em conformidade com os requisitos que compõem os princípios *FAIR* (MACIEL, 2020).

## **2.5 Compartilhamento de dados de pesquisa**

A partir da discussão acerca dos dados de pesquisa, Borgman (2012) argumenta que os pesquisadores desenvolvem muito trabalho físico e intelectual para coletar, gerenciar e analisar seus dados e publicar os resultados de pesquisa. Tal argumento também foi reforçado por Van Den Eynden *et al.* (2011, p. 3, tradução nossa). Para os autores, “os dados de pesquisa são um recurso valioso, geralmente exigindo muito tempo e dinheiro para serem produzidos. Muitos dados têm um valor significativo além do uso para a pesquisa original”.

Os dados são a força vital da pesquisa em qualquer campo, mas o significado dos dados varia conforme o propósito, método, instrumento, comunidade e muitas outras considerações locais e globais. Alguns desses dados podem estar em um formato compartilhável e outros não; assim como, alguns são de valor reconhecido para a comunidade e outros não; alguns pesquisadores desejam compartilhar todos os seus dados o tempo todo, ao mesmo tempo que outros não desejam compartilhar nenhum de seus dados. Mesmo assim, a maioria das pessoas, às vezes, está disposta a compartilhar alguns de seus dados. Essas visões concorrentes, variedade de tipos, origens de dados e de circunstâncias locais contribuem para a complexidade e dificuldade em compartilhá-los (BORGMAN, 2012).

Ainda para o autor, o compartilhamento de dados é uma tarefa difícil, porém, não significa que se deve abandonar a esperança de que os dados sejam compartilhados com algumas pessoas por algum tempo. Os desafios estão pautados em entender quais dados podem ser compartilhados, por quem, com quem, em que condições, o porquê e com quais efeitos. Que constituem a política e prática do compartilhamento de dados, que visa gerar benefícios como o de reproduzir ou verificar pesquisa; disponibilizar os resultados da pesquisa com financiamento público para a sociedade; permitir que outros pesquisadores façam novas descobertas sobre os dados existentes e avançar o estado da pesquisa e inovação (BORGMAN, 2012).

Nesse contexto, Roa-Martínez, Vidotti e Santana (2017) dizem que o compartilhamento de dados pode ser descrito como o ato de disponibilizar dados de pesquisa de várias maneiras, de troca privada ou direta, a pedido da parte interessada o criador dos dados, ou por meio do local de armazenamento (por exemplo, um repositório de dados público) para uso dos pesquisadores, com diferentes finalidades.

Para Borgman (2012), os dados de pesquisa não estão prontamente disponíveis, visto que o compartilhamento é comum em algumas áreas, como astronomia e genômica. Assim, as práticas de compartilhamento de dados em outros campos variam amplamente. Desse modo, os dados da pesquisa assumem diversas formas, são tratados de maneiras diferentes, usando muitas abordagens e, às vezes, são difíceis de interpretar quando removidos de seu contexto inicial.

Porém, existem quatro razões para compartilhar dados, extraído exemplos das ciências sociais e humanas (BORGMAN, 2012, p. 1059, tradução nossa):

- (1) para reproduzir ou verificar a pesquisa;
- (2) para tornar os resultados da pesquisa financiada publicamente disponíveis ao público;
- (3) para permitir outros fazerem novas perguntas sobre os dados existentes, e
- (4) para fazer avançar o estado da pesquisa e inovação.

Essas justificativas diferem pelos argumentos para compartilhar, beneficiários, motivações e incentivos das partes interessadas e envolvidas. Nessa perspectiva, Van Den Eynden *et al.* (2011) relatam que o compartilhamento de dados de pesquisa, mediante a facilidade que os dados digitais possuem em serem armazenados, divulgados e acessíveis de forma online, possibilita as instituições aumentarem o impacto e a visibilidade de forma significativa, por meio do compartilhamento de dados de pesquisa. Além disso, por meio desse compartilhamento, o pesquisador:

- incentiva a investigação científica e o debate;
- promove a inovação e potenciais novos usos de dados;
- leva a novas colaborações entre usuários de dados e criadores de dados;
- maximiza a transparência e responsabilidade;
- permite o escrutínio dos resultados da pesquisa;
- incentiva a melhoria e validação da pesquisa e métodos;
- reduz o custo de duplicação da coleta de dados;
- aumenta o impacto e a visibilidade da pesquisa;
- promove a pesquisa que criou os dados e seus resultados;
- pode fornecer um crédito direto ao pesquisador como um resultado da pesquisa por direito próprio;
- fornece recursos importantes para educação e treinamento. (VAN DEN EYNDEN *et al.*, 2011, p. 3, tradução nossa)

Van Den Eynden *et al.* (2011, p. 3, tradução nossa) destacam, também, que existem várias maneiras de compartilhar dados de pesquisa, incluindo:

- depositando-os em um data center especializado, data arquivo ou banco de dados;
- submetê-los a um jornal para apoiar uma publicação;
- depositando-os em um repositório institucional;
- disponibilizá-los online por um projeto ou site institucional;
- disponibilizando-os informalmente entre pesquisadores em uma base ponto a ponto.

Cada uma dessas formas de compartilhar dados têm vantagens e desvantagens: os data centers podem não ser capazes de aceitar todos os dados apresentados a eles; os repositórios institucionais podem não ser capazes de manter os dados de longo prazo ou dar suporte para dados de pesquisa mais complexos; e os sites são frequentemente efêmeros com pouca sustentabilidade (VAN DEN EYNDEN *et al.*, 2011).

Desse modo, é de extrema importância compartilhar os dados de pesquisa em meios confiáveis. O *FAIRsharing*<sup>1</sup> e *re3data.org*<sup>2</sup> são alguns dos meios para localizar repositórios adequados. Outros repositórios generalistas são: *Figshare*<sup>3</sup>, *Mendeley Data*<sup>4</sup> e o *Zenodo*<sup>5</sup>. Nesse sentido, é relevante selecionar meios que gerem um identificador persistente, por exemplo o *Digital Object Identifier* (DOI), e que possuam um plano de preservação de dados permanentes e atendam aos Princípios *FAIR*.

## 2.6 Percepções sobre compartilhamento de dados de pesquisa

As percepções sobre compartilhamento de dados de pesquisa é o ponto central deste estudo. Por isso, buscou-se mapear, na literatura, pesquisas sobre a temática em questão. Nas próximas linhas apresenta-se os estudos de Veiga (2017), Carvalho e Leite (2019), Kurata, Matsubayashi, Mine (2017), Lima (2020), Lu e Ke (2020).

O estudo de Veiga (2017), menciona que o Acesso Aberto às publicações e aos dados de pesquisa apresentam várias questões a serem solucionadas, dentre elas, a adesão do

---

<sup>1</sup> <https://fairsharing.org/>

<sup>2</sup> <https://www.re3data.org/>

<sup>3</sup> <https://figshare.com/>

<sup>4</sup> <https://data.mendeley.com/>

<sup>5</sup> <https://zenodo.org/>

pesquisador ao compartilhamento de sua produção científica. A referida pesquisa objetivou a investigação da percepção dos neurocientistas e seu comportamento de compartilhamento, bem como os estímulos e as barreiras à sua adesão.

Os principais resultados identificados pela autora mostram que em relação aos **custos** para artigos científicos, a principal influência para não compartilhar foi: “Preocupações com *copyright*”; e para dados de pesquisa: “Preocupações com perda de oportunidade de publicação”. Já em fatores de **benefícios**, a principal motivação para o compartilhamento de artigos científicos foram: “Visibilidade”, “Acessibilidade” e “Altruísmo”. Para dados de pesquisa, foi o “Altruísmo”. O principal fator **contextual** alegado pelos pesquisadores para não compartilhar artigos científicos e dados de pesquisa foi o “Desconhecimento do repositório”. Concluiu-se que a principal barreira para o compartilhamento de artigos científicos e dados de pesquisa tanto no Brasil quanto em Portugal é a desinformação: da plataforma a ser utilizada para este fim, das possibilidades de versões e embargos que ela suporta e dos mecanismos existentes para que o pesquisador possa compartilhar seus artigos e dados sem perder oportunidades de publicação e sem infringir os contratos com as editoras.

Por sua vez, a pesquisa de Carvalho e Leite (2019), teve como objetivo identificar as práticas de produção/obtenção, compartilhamento e (re)uso de dados de pesquisa em Química, Antropologia e Educação. Um dos resultados obtidos mostrou que pesquisadores nas três disciplinas analisadas consideraram que os dados de pesquisa que produzem/obtem são úteis a outros pesquisadores e possuem disposição para o compartilhamento desses dados. Outrossim, a percepção dos pesquisadores foi que a prática do compartilhamento de dados de pesquisa ainda se restringe: aos próprios grupos de pesquisa, a docentes, discentes e aos pesquisadores que realizam pesquisas correlatas.

A pesquisa aponta ser perceptível nos resultados que plataformas digitais, como repositórios de dados de pesquisa, ocupam funções centrais para o compartilhamento e (re)uso de dados nas três disciplinas analisadas, na qual se evidenciou uma demanda potencial de pesquisadores na disponibilização de dados nesses repositórios. Entretanto, há muitos desafios para o estabelecimento de práticas de compartilhamento e (re)uso de dados de pesquisa por meio de repositórios, principalmente, no Brasil.

Nesse contexto, tem-se muito a discutir sobre o papel de pesquisadores, instituições de ensino e pesquisa, agências de fomento, editores de periódicos científicos e da sociedade em geral na promoção da Ciência Aberta. Porém, surge o questionamento sobre quais seriam as

diretrizes necessárias para a gestão de dados de pesquisa, em termos de planejamento de pesquisa, captura, descrição, avaliação de qualidade, compartilhamento e preservação de dados.

Kurata, Matsubayashi e Mine (2017) buscaram fornecer uma visão geral das práticas e percepções de pesquisadores japoneses em ciências naturais sobre o uso e compartilhamento de dados. O estudo chegou a resultados que forneceram uma visão geral da relação entre dados de pesquisa e práticas de pesquisa e que constituem três camadas centradas em “dados de pesquisa”.

A primeira camada é formada (emoldurada) por três posições sobre os dados: “(A) Acesso aberto para dados”, “(B) Acesso restrito para dados” e “(C) Interpretação de dados”. Esta camada indica uma percepção mais superficial dos pesquisadores em relação aos dados. A segunda camada é “(D) Processamento e preservação de dados”, considerada como a incorporação de práticas reais de pesquisa que apoiam a percepção do compartilhamento de dados. A terceira camada fornece a base para a conscientização sobre o compartilhamento de dados e práticas de pesquisa. Em outras palavras: essa camada funciona ou interfere no compartilhamento, ou no uso de dados abertamente

Nessa perspectiva, Lima (2020), desenvolveu uma pesquisa, a partir de indagações elementares, sobre a percepção dos pesquisadores com relação aos dados de pesquisa e acerca dos princípios da Ciência Aberta. O estudo buscou investigar quais os desafios e dificuldades encontradas pelos pesquisadores no gerenciamento dos dados, compartilhamento de dados e na adoção das práticas da Ciência Aberta. Por esse prisma, os *stakeholders* (pesquisadores, bibliotecários e instituição) precisam estar cientes e compreender esse processo, sendo as ações formais de educação e informação um meio para concretizar a gestão dos dados. A pesquisa alcançou os seguintes resultados: os motivos para não compartilhar os dados são variados e o principal motivo trata-se do desconhecimento e por não saber como proceder o compartilhamento, apesar dos pesquisadores concordarem com o compartilhamento e acesso aos dados proposto pela Ciência Aberta, não significa que eles compartilham os dados, e quando o fazem, trata-se de uma exceção e não uma regra, especialmente em algumas áreas.

A pesquisa de Lu e Ke (2020) explora o entendimento de estudiosos taiwaneses sobre as percepções e práticas de gerenciamento de dados de pesquisa, a qual investigou problemas, incluindo fontes e tipos de dados de pesquisa, armazenamento e pesquisa de dados, metadados de dados de pesquisa, plano de gerenciamento de dados, treinamento, suporte, compartilhamento e reutilização de dados. Como resultado do item “práticas de dados de

pesquisa armazenamento e gerenciamento” os acadêmicos evidenciaram o armazenamento de seus dados de pesquisa em repositórios de dados criados por suas instituições. De modo geral, a pesquisa chegou ao seguinte posicionamento:

- (1) institutos de pesquisa e bibliotecas acadêmicas devem desenvolver RDM Serviços;
- (2) agências de financiamento e pesquisa institutos devem formular políticas de RDM ou mandatos; e
- (3) institutos de pesquisa, acadêmicos bibliotecas e agências de financiamento devem estabelecer repositórios de dados de pesquisa equipados com a funcionalidade de controle de acesso. (LU; KE, 2020, p. 5, tradução nossa).

Diante dos estudos mencionados percebe-se a importância da presente pesquisa no âmbito da UFPA, pois, conforme relata Lima (2020), levantamentos como o do presente estudo possibilita obter uma noção sobre o entendimento e as práticas dos pesquisadores em relação à gestão de dados de pesquisa e dos temas correlatos. Nestes termos, é preciso sensibilizar, conscientizar e informar aos pesquisadores sobre o tema, ressaltando os benefícios da adesão a essa prática. Além disso, as informações coletadas aqui servem de insumos para tornar palpáveis a construção de diretrizes institucionais a respeito de dados de pesquisa na instituição.

### 3 METODOLOGIA

A classificação desta pesquisa é de cunho exploratório e relaciona-se ao objetivo de investigar as percepções e práticas dos docentes permanentes de programas de pós-graduação do ITEC e ICOSA sobre o compartilhamento de dados para entender, identificar e apresentar um panorama sobre essa temática na percepção dos pesquisados.

O estudo de caso, de acordo com Schramm (1971), é a essência de uma pesquisa que busca esclarecer uma decisão ou um conjunto de decisões mediante o motivo pelo qual foram tomadas, implementadas e quais resultados geraram. Boaventura (2004) alega que por meio desse tipo de pesquisa é possível a aplicação prática de conhecimentos para a solução de problemas sociais.

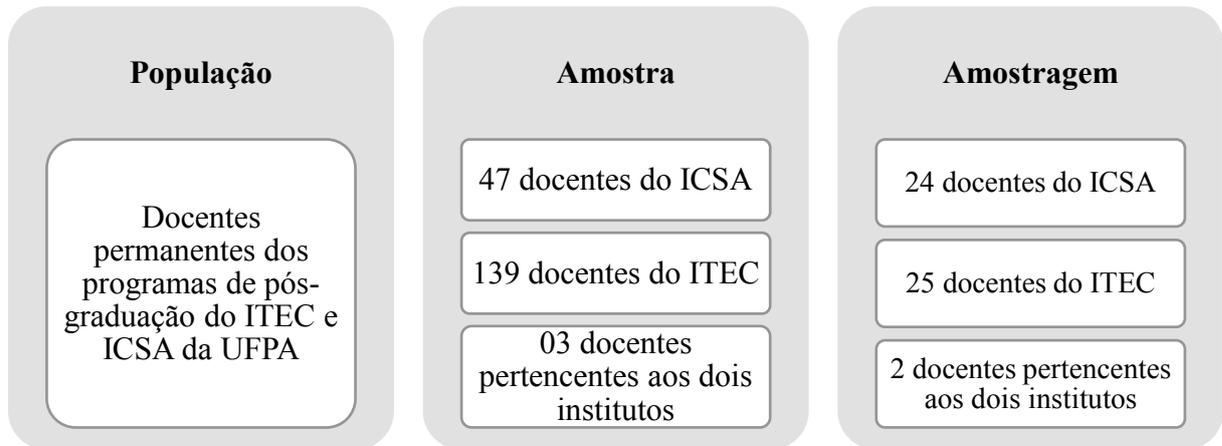
Para Yin (2015) o estudo de caso representa uma investigação empírica, por envolver coleta direta de dados de um grupo de informantes em determinada realidade. Ademais, compreende um método abrangente, que ocorre a partir de uma lógica de planejamento, coleta e análise de dados, podendo incluir estudos de caso único ou múltiplos.

Para o estudo de caso, a amostra estudada foi constituída por docentes permanentes dos programas de pós-graduação selecionados. A escolha desses institutos se deu pelo fato de o ITEC ser o instituto com maior número de pós-graduações e o ICOSA por ser um dos maiores institutos da universidade, além de serem institutos que englobam diferentes áreas do conhecimento, possibilitando assim, uma visão ampliada dessas áreas.

O ITEC compreende as áreas de engenharias, enquanto o ICOSA as áreas de ciências sociais aplicadas. Ambas possuem diferentes métodos para o desenvolvimento e construção de suas pesquisas que enriquece a presente pesquisa com as peculiaridades de cada área.

Para o desenvolvimento desta pesquisa utilizou-se a amostragem intencional não probabilística, a qual, segundo Deslauriers e Kèrìsit (2008) não é constituída ao acaso, mas em razão dos aspectos específicos que o pesquisador busca compreender, conforme sintetiza a Figura 7.

Figura 7 - População e amostra da pesquisa



Fonte: Elaborado pela autora (2021)

Assim, a abordagem desta pesquisa está relacionada a análises qualitativas para subsidiar a investigação sobre as percepções dos docentes por meio de perguntas realizadas a partir de questionário e entrevista.

### 3.1 Instrumentos de coleta de dados

Utilizou-se dois instrumentos de coleta de dados (questionário e roteiro de entrevista), construídos com base na literatura e em acordo com os objetivos específicos da pesquisa, conforme apresentados no Quadro 6.

Quadro 6 - Relação entre os objetivos específicos e o instrumento metodológico

Objetivos específicos	Instrumento metodológico
Identificar as percepções dos docentes sobre compartilhamento de dados de pesquisa;	Questionário (questões 3.1 a 3.4) Roteiro de entrevista (questões 2 a 8)
Mapear as práticas de compartilhamento de dados de pesquisa;	Questionário (questões 2.1 a 2.13) Roteiro de entrevista (questões 9 a 13)
Apresentar o panorama das percepções e práticas por área do conhecimento/por programa de pós-graduação.	Questionário e roteiro de entrevista

Fonte: Elaborado pela autora (2021)

A construção das questões foram baseadas no trabalhos de Vanz *et al.* (2018) com o questionário “Práticas e percepções sobre acesso aberto a dados de pesquisa” e Tenopir *et al.* (2011) a partir do questionário intitulado “Avaliação de dados de pesquisa: construindo uma compreensão de suas necessidades de dados”.

Ainda que essa metodologia tenha sido baseada em pesquisas já validadas em âmbito nacional e internacional, foi submetido aos orientadores e à banca de qualificação e conforme instruções foram realizados alguns ajustes de layout, reformulação e reescrita de questões consideradas redundantes. Por recomendação da banca de qualificação, foi retirado o glossário de conceitos em relação ao tema abordado, pois poderia influenciar as respostas. Dessa forma, o questionário constitui-se de 34 questões divididas em quatro seções:

- Seção 1 – Apresentação geral deste estudo: contém breve descrição deste estudo e a opção de aceite em participar da pesquisa;
- Seção 2 – Questões referentes ao perfil do pesquisador: constituiu-se de nove questões para conhecimento do perfil dos respondentes;
- Seção 3 – Questões referentes as suas práticas sobre dados de pesquisa: composta por dezesseis questões relacionadas às práticas de compartilhamento de dados de pesquisa docentes participantes desta pesquisa;
- Seção 4 – Questões referentes as suas percepções em relação ao compartilhamento e uso de dados: possui quatro questões. Destaca-se que as questões 3.5 e 3.6 foram destinadas para o aceite em participar da entrevista.

Adicionalmente, aplicou-se a ferramenta 5W2H para detalhar quais questões do questionário se relacionam com os objetivos desta pesquisa. Esta ação ajudou na compreensão dos tipos de dados a serem coletados para o alcance dos objetivos (Apêndice A). A ferramenta representa as seguintes perguntas: O que? Por que? Onde? Quando? Quem? Como? e Quanto? ou seja, questionamentos que devem ser feitos e respondidos ao investigar ou relatar um fato ou situação.

Para complementar a coleta de dados foi elaborado um roteiro de entrevista semiestruturado, adaptado das pesquisas realizadas por Carvalho (2018) e Lima (2020), contendo, em sua versão final, 13 perguntas, distribuídas da seguinte forma:

- 1 questão para conhecimento geral sobre acesso aberto a dados de pesquisa;

- 7 questões destinadas ao conhecimento sobre percepção de compartilhamento de dados de pesquisa;
- 4 questões a respeito das práticas de compartilhamento de dados de pesquisa;
- 1 questão para o entrevistado fazer algum comentário que julgue necessário.

A coleta de dados iniciou em 01 de junho de 2021 e finalizou em 05 de julho de 2021. As entrevistas ocorreram entre 10 de junho de 2021 a 05 de julho de 2021. Devido à pandemia da Covid-19 as entrevistas sucederam de forma síncrona, por meio da plataforma do Google Meet. Destaca-se que todas as entrevistas foram gravadas com anuência dos participantes.

#### 4 RESULTADO E ANÁLISE DOS DADOS

Nesta seção será apresentada a análise dos resultados obtidos por meio da aplicação de questionário e do roteiro de entrevistas realizadas com docentes permanentes dos programas de pós-graduação do ITEC e ICSA sobre percepção e práticas de compartilhamento de dados de pesquisa, com a participação de 51 professores pesquisadores, conforme detalhado na Tabela 1.

Tabela 1 - Respondentes por instituto

Institutos	Total de docentes permanentes	Nº de participantes	
		Questionário	Entrevista
ICSA	47	24 (51%)	7 (29%)
ITEC	139	25 (18%)	2 (8%)
ITEC E ICSA	03	2 (67%)	1 (50%)
<b>Total</b>	189	51 (27%)	10 (20%)

Fontes: Dados da pesquisa (2021).

Dentre os 189 docentes permanentes, somente 51 responderam ao questionário e desses 10 aceitaram participar da entrevista. Como apresentado na Tabela 1, o maior número de participantes foi do ICSA com 51% (n=24) respondentes no questionário e 29% (n=7) na entrevista, enquanto que o ITEC teve 18% (n=25) de participação no questionário e 8% (n=2) na entrevista.

No que tange ao perfil desses docentes, são apresentadas, informações quanto à faixa etária, gênero e tempo de envolvimento em pesquisa, todos os participantes da pesquisa são doutores, dentre os 51 pesquisadores, 15 deles possuem pós-doutorado.

Em relação à idade, conforme a Tabela 2, a faixa entre 20-25 não obteve nenhuma indicação. 4% (n=2) dos respondentes assinalaram possuir entre 26-35 anos, 31% (n=16) entre 36-45 anos, 33% (n=17) indicaram ter idade entre 46-55 anos, 28% (n=15) entre 56-65 anos e 2% (n=1) afirmaram ter mais de 65 anos. Assim, entre os 30 respondentes que indicaram já ter compartilhado dados de pesquisa (Tabela 12), percebeu-se maior incidência nas faixas etárias entre 36-45 e 46-55 anos somando 20 indicações.

Tabela 2 - Faixa etária dos participantes da pesquisa

<b>Faixa etária</b>	<b>Total de respostas</b>
20-25	0 (0%)
26-35	2 (4%)
36-45	16 (31%)
46-55	17 (33%)
56-65	15 (29%)
Mais de 65	1 (2%)
<b>Total</b>	<b>51 (100%)</b>

Fonte: Dados da pesquisa (2021).

Já em relação ao gênero, a partir dos dados da Tabela 3, observa-se que 41% (n=21) são do sexo feminino e 59% (n=30) do masculino, demonstrando uma pequena predominância masculina em relação ao gênero feminino. Esse resultado pode ser justificado pelo fato de nos dois institutos a predominância ser no sexo masculino, onde no ITEC de 142 docentes, 74% (n=105) são do gênero masculino e 26% (n=37) do feminino. No ICESA, 60% (n=30) são do sexo masculino e 40% (n=20) do feminino.

Tabela 3 - Relação de gênero entre os participantes

<b>Gênero</b>	<b>Total de respostas</b>
Feminino	21 (41%)
Masculino	30 (59%)
Não informar	0 (0%)
<b>Total</b>	<b>51 (100%)</b>

Fonte: Dados da pesquisa (2021).

Quanto ao tempo envolvido em pesquisa, percebeu-se que os docentes atuam em média a mais de 10 anos na condução de pesquisas: 6% (n=3) assinalaram < 5 anos envolvidos em pesquisa; 18% (n=9) entre 5-10 anos; 22% (n=11) entre 11-15 anos; 16% entre 16-20 anos e a maior representatividade, com 39% (n=20) escolheu a opção: tempo > 20 anos (Tabela 4).

Tabela 4 - Tempo de envolvimento em pesquisa

<b>Tempo de envolvido em pesquisa</b>	<b>Total de respostas</b>
< 5 anos	3 (6%)
5-10 anos	9 (18%)
11 – 15 anos	11 (22%)
16-20 anos	8 (16%)
> 20 anos	20 (39%)
Nunca estive envolvido em pesquisa	0 (0%)
<b>Total</b>	<b>51 (100%)</b>

Fonte: Dados da pesquisa (2021).

O perfil dos docentes participantes desta pesquisa demonstra uma maturidade e experiência acadêmica, visto que a maioria afirmou ter mais de 20 anos envolvidos em pesquisa, o que explica porque a faixa etária é acima dos 30 anos de idade.

Em breve pesquisa no Currículo *Lattes* dos 20 docentes que possuem mais de vinte anos no envolvimento em pesquisa, pode-se identificar que em média eles possuem 26 anos na carreira de docente e 17 anos com o título de doutor. Vale ressaltar, também, que dentre eles, 70% (n=14) pertencem ao ITEC e 30% (n=6) ao ICOSA. Esses resultados podem ser justificados pelo fato de o ICOSA possuir cinco programas e três deles terem menos de 6 anos de criação, como: Programa de Pós-Graduação em Administração (PPGAD) criado em 2020; Programa de Pós-Graduação Economia Aplicada (PPGEA) criado em 2019 e o Pós-Graduação em Ciência da Informação (PPGCI) criado em 2016.

Com relação ao vínculo dos docentes, identificou-se a ocorrência de vínculo em mais de um programa. Por esse motivo não foram quantificados os números de docentes por programa. Os resultados mostraram que houve representatividade docente dos cinco programas de pós-graduação do ICOSA: Programa de Pós-Graduação em Serviço Social (PPGSS); Programa de Pós-Graduação em Economia (PPGE); Pós-Graduação em Ciência da Informação (PPGCI); Programa de Pós-Graduação em Economia Aplicada (PPGEA) e Programa de Pós-Graduação em Administração (PPGAD).

No caso do ITEC, verificou-se a participação de professores de 10 programas: Programa de Pós-Graduação em Ciência e Tecnologia de Alimentos (PPGCTA); Programa de Pós-Graduação em Arquitetura e Urbanismo (PPGAU); Programa de Pós-Graduação em

Engenharia Elétrica (PPGEE); Programa de Pós-Graduação em Ciências do Patrimônio Cultural (PPGPatri); Programa de Pós-Graduação em Engenharia Sanitária e Ambiental (PPGESA); Programa de Pós-Graduação em Engenharia Mecânica (PPGEM); Programa de Pós-Graduação em Engenharia Naval (PPGENAV); Programa de Pós-Graduação em Engenharia de Recursos Naturais da Amazônia (PRODERNA); Programa de Pós-Graduação em Engenharia de Processos (PPGEP) e Programa de Pós-Graduação em Engenharia Industrial (PPGEI).

No que se refere ao financiador de pesquisa, tomando como base o estudo realizado por Monteiro, Lucas e Lecardelli (2019), compreendeu-se que os financiadores de pesquisas apresentam recomendações ou exigências para que os projetos financiados por eles incluam plano de gestão de dados científicos, visando o compartilhamento em acesso aberto. A iniciativa envolve a cooperação entre instituições de pesquisa, instituições de desenvolvimento, instituições de ensino e a própria sociedade, que reflete uma nova forma de pesquisa científica baseada em trabalho colaborativo e interativo no compartilhamento e reuso de dados, propício ao movimento de acesso aberto.

Desse modo, mediante a importância dos financiadores para o desenvolvimento do compartilhamento de dados de pesquisa, buscou-se identificar qual o principal financiador de pesquisa dos docentes permanentes dos programas de pós-graduação do ITEC e ICOSA.

No questionário, as alternativas eram passíveis de serem marcadas conforme julgassem necessário. Os resultados da pesquisa são apresentados na Tabela 5.

Tabela 5 - Qual das alternativas melhor descreve o principal financiador da sua pesquisa

	<b>ITEC</b>	<b>ICOSA</b>	<b>Total</b>
A própria instituição	12 (44%)	15 (56%)	27 (39%)
Agência nacional	16 (62%)	10 (38%)	26 (38%)
Agência internacional	1 (25%)	3 (75%)	4 (6%)
Empresa privada	5 (100%)	0 (0%)	5 (7%)
Agência regional	1 (33%)	2 (67%)	3 (4%)
Outros	2 (50%)	2 (50%)	4 (6%)

Fonte: Dados da pesquisa (2021).

A partir dos dados da Tabela 6, observa-se que: 39% (n=27) dos respondentes apontam como principal financiador a própria instituição, seguindo das agências nacionais, apontada por 38% (n=26) dos docentes. Nesse sentido, nota-se a importância da própria instituição no desenvolvimento de uma política norteadora para a gestão de dados de pesquisa, ponto de partida para o avanço do compartilhamento de dados de pesquisa de forma eficaz na instituição.

Percebe-se que nas áreas que envolvem os programas do ITEC os pesquisadores possuem financiamento de empresas privadas (n=5), enquanto no ICISA não se teve nenhuma sinalização. Dentre a opção “outros”, houve a indicação de financiamento com recursos próprios.

Adicionalmente, buscou-se saber qual o nível de conhecimento que o respondente tinha sobre as temáticas: **ciência aberta**, **dados de pesquisa** e **compartilhamento de dados de pesquisa** (Tabela 6). Para avaliar estes conhecimentos estabeleceu-se uma escala de 1 (nível mais baixo de conhecimento) a 5 (nível mais alto de conhecimento) para cada temática.

Tabela 6 - Nível de conhecimento sobre ciência aberta, dados de pesquisa e compartilhamento de dados de pesquisa

	Nível 1	Nível 2	Nível 3	Nível 4	Nível 5	Total de respostas
Ciência aberta	3 (6%)	11 (22%)	11 (22%)	14 (27%)	12 (24%)	51 (100%)
Dados de pesquisa	0 (0%)	3 (6%)	12 (24%)	15 (29%)	21 (41%)	51 (100%)
Compartilhamento de dados de pesquisa	0 (0%)	6 (12%)	16 (31%)	15 (29%)	14 (27%)	51 (100%)

Fonte: Dados da pesquisa (2021).

A ciência aberta foi a única temática que recebeu 6% (n=3) indicações de nível 1, o mais baixo de conhecimento. Foi também a temática com maior número de indicações no nível 2, em relação as outras, com 22% (n=11) marcações. O nível 3 obteve, também, 22% (n=11) das indicações. Já o nível 4 e 5 tiveram 27% (n=14) e 24% (n=12) indicações, respectivamente. Conforme demonstra estudo de Silva e Silveira (2019b) a Ciência Aberta é importante, visto a forma de rediscutir o desenvolvimento científico, produzindo nova maneira de gerar conhecimento, no qual as mudanças institucionais podem levar certo tempo nessa adaptação.

Em relação aos dados de pesquisas, obteve-se os seguintes resultados: 6% (n=3) docentes indicaram possuir nível 2 de conhecimento, enquanto 24% (n=12) indicaram possuir nível 3 de conhecimento, 29% (n=15) julgaram possuir nível 4 de conhecimento e 41% (n=21) assinalaram possuir o nível máximo de conhecimento sobre esse tema, o nível 5. Percebe-se que o tema “dados de pesquisa” foi o item que teve maior número de indicação de alto nível de conhecimento, com 21 sinalizações, representando um percentual de 41% das respostas nesse assunto.

Sobre o compartilhamento de dados de pesquisa, 12% (n=6) dos docentes indicaram possuir nível 2 de conhecimento, 31% (n=16) assinalaram possuir nível 3, 29% (n=15) apontaram ter nível 4 e 27% (n=14) registraram que tinham nível 5. Assim, pode-se perceber que, sobre essa temática, parte dos docentes apresentam um nível de conhecimento elevado, uma vez que as respostas variaram entre 3, 4 e 5 na escala.

Nesse sentido, pode-se inferir que o tema com maior nível de conhecimento (nível 5) entre os respondentes está relacionado aos dados de pesquisa com 41% (n=21) e compartilhamento de dados de pesquisa com 27% (n=14); e o menor é ciência aberta com 6% (n=3). Esses resultados ressaltam a importância da discussão sobre a ciência aberta, pois a maior familiaridade está relacionada a dados de pesquisa e compartilhamento de dados, temáticas inerentes as atividades relacionadas ao cotidiano do pesquisador, pode ser que isso justifique os maiores índices.

#### **4.1 Práticas sobre compartilhamento de dados de pesquisa**

Para o mapeamento das práticas de compartilhamento de dados de pesquisas dos docentes permanentes dos programas de pós-graduação do ICSA e do ITEC, estabeleceu-se as questões que vão de 2.1 até 2.13 (Apêndice A).

Foi solicitado aos participantes, que eles marcassem qual a melhor forma de descrever os tipos de dados de pesquisa produzidos por eles (Tabela 7). Vale ressaltar que as alternativas podiam ser julgadas conforme o respondente achasse necessário, podendo ter mais de uma opção de resposta. Desse modo, foram obtidas 186 respostas. Nota-se que na frequência total o item “dados documentais” possui maior indicação com 16% (n=29), seguido de “entrevistas” e “questionários” com 15% (n=27), cada. Em relação aos respondentes do ICSA esse percentual

se mantém entre dados documentais e entrevistas com 21% (n=16) e de questionários com 20% (n=15). Já para os respondentes do ITEC, esse percentual fica entre: “Dados coletados de sensores e equipamentos” e “Dados experimentais” com 18% (n=20), cada.

Tabela 7 - Termo que melhor descreve os tipos de dados de pesquisa por instituto

<b>Termos de pesquisa</b>	<b>ICSA</b>	<b>ITEC</b>	<b>TOTAL</b>
Dados automaticamente gerados por programas de computador	4 (5%)	10 (9%)	14 (8%)
Dados coletados de sensores e equipamentos	3 (4%)	20 (18%)	23 (12%)
Dados de simulação	5 (7%)	14 (13%)	19 (10%)
Dados documentais	16 (21%)	13 (12%)	29 (16%)
Dados experimentais	5 (7%)	20 (18%)	25 (14%)
Dados observacionais	10 (13%)	10 (9%)	20 (11%)
Entrevistas	16 (21%)	11 (10%)	27 (15%)
Questionários	15 (20%)	12 (11%)	27 (15%)
Outros	2 (3%)	0 (0%)	2 (1%)

Fonte: Dados da pesquisa (2021).

Para conhecer melhor quais os tipos de dados produzidos pelos docentes do ITEC e do ICSA, solicitou-se que os respondentes apontassem, entre 12 opções, aquelas que representassem o tipo dos dados de pesquisa mais produzidos pelos mesmos, marcando quantas alternativas desejassem. Assim, a Tabela 8 mostra a proporção das 208 indicações. Os resultados obtidos nessa questão revelam que, no geral, para os respondentes o tipo de dados mais produzidos trata-se de “textos” com 18% (n=37), seguido por “bancos de dados” e “planilhas” com 16% (n=33), cada. No que se refere aos dados alcançados no ICSA, em similaridade aos resultados gerais, os dados mais apontados foram: os “textos” com 23% (n=18) e “planilhas” com 19% (n=15). No entanto, para os docentes do ITEC os tipos de dados mais produzidos são: “bancos de dados” com 18% (n=23), “textos” com 15% (n=19) e planilhas, com 14% (n=18).

Tabela 8 - Tipos de dados de pesquisa por instituto

<b>Tipos de dados</b>	<b>ICSA</b>	<b>ITEC</b>	<b>Total</b>
Algoritmos	4 (5%)	8 (6%)	12 (6%)
Animações, modelos ou simulações	4 (5%)	14 (11%)	18 (9%)
Áudios	1 (1%)	4 (3%)	5 (2%)
Banco de dados	10 (13%)	23 (18%)	33 (16%)
Imagens	8 (10%)	14 (11%)	22 (11%)
Planilhas	15 (19%)	18 (14%)	33 (16%)
Slides	8 (10%)	13 (10%)	21 (10%)
Software	3 (4%)	7 (5%)	10 (5%)
Textos	18 (23%)	19 (15%)	37 (18%)
Vídeos	4 (5%)	7 (5%)	11 (5%)
Websites	1 (1%)	3 (2%)	4 (2%)
Outros	1 (1%)	1 (1%)	2 (1%)

Fonte: Dados da pesquisa (2021).

Além de buscar conhecer sobre os termos e tipos de pesquisa, indagou-se sobre o volume estimado que esses pesquisadores produzem em um ano. Como mostra a Tabela 9, foram sinalizadas 51 respostas sobre essa questão. Analisando o total de respostas, observa-se que a porcentagem do item “não sei responder” é alta, com 29% (n=15), similar ao que ocorre também entre os pesquisadores do ICSA com 29% (n=7) e do ITEC com 30% (n=8). A falta de conhecimento sobre o volume de dados produzidos pode ser um problema para instituição, visto a necessidade dessa mensuração para prever formas de armazenamentos, utilizando como exemplo o uso de um repositório de dados de pesquisa.

Tabela 9 - Volume estimado de dados produzidos em um ano

	<b>ICSA</b>	<b>ITEC</b>	<b>TOTAL</b>
< 1 GB	3 (13%)	1 (4%)	4 (8%)
1 - 50 GB	6 (25%)	7 (26%)	13 (25%)
50 - 100 GB	5 (21%)	1 (4%)	6 (12%)
100 - 500 GB	3 (13%)	3 (11%)	6 (12%)
500 GB - 1 TB	0 (0%)	5 (19%)	5 (10%)
1 - 50 TB's	0 (0%)	1 (4%)	1 (2%)
50 - 100 TB's	0 (0%)	0 (0%)	0 (0%)
> 100 TB's	0 (0%)	1 (4%)	1 (2%)
Não sei responder	7 (29%)	8 (30%)	15 (29%)
<b>Total</b>			<b>51 (100%)</b>

Fonte: Dados da pesquisa (2021).

Percebe-se que a maioria dos pesquisadores não conseguem mensurar os dados por eles produzidos. Ao comparar esse resultado com a pesquisa em âmbito nacional desenvolvida por Vanz *et al.* (2018) em que em média 32,90% não souberam responder, entende-se que esse resultado não trata-se de uma particularidade entre esses docentes.

Adicionalmente, buscou-se saber o quantitativo de pessoas envolvidas na produção dos dados de pesquisa, conforme mostra a Tabela 10. Observa-se que 47% (n=24) do volume de dados é produzido em grupos de até cinco pessoas. Verifica-se que a alternativa menos marcada é a que indica dados produzidos por grupo com mais de 12 pessoas, com um percentual de 8% (n=4). Em relação aos institutos, nota-se equiparidade em relação ao grupo de até cinco pessoas, mas para os respondentes do ITEC as pesquisas desenvolvidas com grupos de seis a doze pessoas, possuem incidência maior, com 48% (n=13).

Tabela 10 - Quantidade de envolvidos na produção de dados de pesquisa

	ICSA	ITEC	Total
De forma individual	4 (17%)	0 (0%)	4 (8%)
Em grupo de até 5 pessoas	12 (50%)	12 (44%)	24 (47%)
Em grupo de 6 até 12 pessoas	6 (25%)	13 (48%)	19 (37%)
Em grupo de mais de 12 pessoas	2 (8%)	2 (7%)	4 (8%)
<b>Total</b>			51 (100%)

Fonte: Dados da pesquisa (2021).

Com relação às formas de armazenamento dos dados produzidos por esses docentes (Tabela 11), pode-se perceber, que apenas o item “nuvem” foi assinalado pelo total de respondentes, 100% (n=51). As opções com menos respostas foram: “Servidor / *data center* da minha instituição” e “Arquivos em papel” com 69% (n=35) cada. Dentre as alternativas, “computador pessoal” (n=26) e “mídia externa” (n=25) foram as que apresentaram maior representatividade de armazenamento de todos os dados, seguido do “armazenamento em nuvem” para todos os seus dados (n=19) ou parte deles (n=28).

Tabela 11 - Locais de armazenamento

	<b>Todos os meus dados</b>	<b>Parte dos meus dados</b>	<b>Nenhum dos meus dados</b>	<b>Não sei responder</b>	<b>Total de respostas</b>
Computador pessoal	26	22	1	-	49 (96%)
Servidor / data center da minha instituição	1	17	16	1	35 (69%)
Arquivos em papel	-	12	22	1	35 (69%)
Nuvem (ex. Google drive / Dropbox)	19	28	4	-	51 (100%)
Mídia externa (ex. HD externo, USB/pendrive, CD/DVD)	25	17	3	-	45 (88%)

Fonte: Dados da pesquisa (2021).

Estes resultados demonstram que os pesquisadores têm a prática de manter os dados de pesquisa sob o seu controle pessoal, reforçando a conclusão do estudo de Vanz *et al.* (2018, p. 21) ao sinalizar a “falta de instrumentos e diretrizes uniformes para a preservação e segurança de dados de pesquisa - a matéria prima do empreendimento científico”.

Ao indagar os pesquisadores sobre a prática de compartilhamento de dados de pesquisa em acesso aberto (Tabela 12), onde os pesquisadores poderiam assinalar quantas alternativas julgassem necessárias, 59% (n=30) indicaram já terem praticado o compartilhamento e 41% (n=21) indicaram não terem essa prática.

Dos 30 que indicaram a prática de compartilhamento, percebeu-se que 36% (n=24) compartilharam dados de pesquisas para publicação em periódicos científicos ou em anais de evento. Mas o que chamou atenção foi o alto índice de respostas para aqueles que nunca compartilham os dados de pesquisa: 31% (n=21).

Em relação aos institutos, nota-se que os pesquisadores do ITEC indicaram mais prática em compartilhamento de dados, por meio de “publicações em periódicos” (n=16), “solicitação do principal financiador” (n=6) e por “iniciativa própria” (n=6). Já no ICSA a maior indicação foi “nunca compartilhei” (n=13).

Tabela 12 - Você já compartilhou dados de pesquisa em acesso aberto?

	ICSA	ITEC	Total
Sim, por solicitação do meu principal financiador	3	6	9 (13%)
Sim, por solicitação da instituição principal em que atuo	1	3	4 (6%)
Sim, para publicação em periódico ou em anais de evento	8	16	24 (36%)
Sim, por iniciativa própria	3	6	9 (13%)
Nunca compartilhei	13	8	21 (31%)

Fonte: Dados da pesquisa (2021).

Os resultados da Tabela 12, relacionados com o perfil dos participantes, demonstrou que a prática de compartilhamento de dados de pesquisa indicada pelos 30 participantes, se deu em maior frequência no sexo masculino 63% (n=19) em relação ao feminino 37% (n=11). Confrontando com os resultados de Veiga (2017), que realizou estudo sobre o compartilhamento de dados entre pesquisadores da área de neurociências, em que a predominância se deu no sexo feminino. Nesse sentido, percebe-se a relação de gênero entre as áreas do conhecimento, conforme apontam estudos de Vieira *et al.* (2019), Moreira, Mattos e Reis (2014) e Hausmann, Vogt e Hein (2018), ao demonstrarem que na área da saúde a incidência é sexo feminino, já na área da engenharia e ciências sociais aplicadas é o sexo masculino.

Relacionando a prática de compartilhamento de dados em comparação ao tempo de envolvimento em pesquisa, o maior percentual de compartilhamento 53% (n=16) ocorreu entre os que possuem >20 anos envolvidos em pesquisa.

Entre os programas dos dois institutos, pode-se identificar que no ITEC o Programa de Pós-Graduação em Engenharia Mecânica possui maior prática em compartilhamento de dados de pesquisa com representatividade de 20% (n=6), já no ICSA foi o Programa de Pós-Graduação em Ciência da Informação 10% (n=3). Esses resultados podem ser justificados pelo fato do PPGCI estar inserido nesse contexto de Ciência Aberta e compartilhamento de dados de pesquisa com o desenvolvimento de pesquisas sobre essas temáticas. E o PPGEM por possuir maior financiamento por parte de agências financiadoras.

Já a relação ao principal financiador percebeu-se maior relação de compartilhamento de dados de pesquisa entre a própria instituição e agências nacionais com 15 indicações cada. E entre a quantidade de envolvidos na produção de dados de pesquisa, notou-se que essa prática

ocorre com maior frequência entre os grupos de até 5 pessoas com 50% (n=15) seguido de grupos de 6 a 12 pessoas com 37% (n=11).

Concomitantemente, buscou saber se os dados produzidos por esses pesquisadores são compartilhados em acesso aberto (Tabela 13). Os resultados se mostraram bem divididos, pois enquanto 49% (n=25) dos pesquisadores compartilham parte dos seus dados, 47% (n=24) dos respondentes afirmaram que não compartilham nenhum dos seus dados.

Tabela 13 - Quanto dos seus dados de pesquisa você atualmente compartilha em acesso aberto?

	<b>ICSA</b>	<b>ITEC</b>	<b>Total</b>
Todos meus dados	2 (8%)	0 (0%)	2 (4%)
Parte dos meus dados	8 (33%)	17 (63%)	25 (49%)
Nenhum dos meus dados	14 (58%)	10 (37%)	24 (47%)

Fonte: Dados da pesquisa (2021).

Embora não seja o objetivo da pesquisa comparar os institutos, mas devido a importância das características intrínseca em cada área de atuação, verificou-se que entre os respondentes dos dois institutos, os pesquisadores do ITEC possuem maior prática em compartilhamento: 63% (n=17) assinalaram que compartilham parte dos seus dados, ao contrário do ICSA onde 58% (n=14) dos participantes sinalizaram não compartilhar nenhum dos seus dados em acesso aberto. Corroborando com o estudo de Vanz *et al.* (2018) que apontam não ser uma prática comum entre os pesquisadores brasileiros.

No que tange aos locais em que os pesquisadores compartilham seus dados, a Tabela 14 apresenta onde essa prática é realizada. Verifica-se que, em sua maioria, os dados são compartilhados em: “repositórios de revistas/editor” com 75% (n=18) das indicações, seguido dos “repositórios temáticos” com 50% (n=13); “repositório da minha instituição” com 55% (n=12) e os repositórios nacionais com 50% (n=11) indicações.

Tabela 14 - Se você respondeu que compartilha seus dados de pesquisa em acesso aberto, onde você compartilha?

	<b>Todos os meus dados</b>	<b>Parte dos meus dados</b>	<b>Nenhum dos meus dados</b>	<b>Não sei responder</b>	<b>Total de respostas</b>
Repositório temático	2 (8%)	13 (50%)	9 (35%)	2 (8%)	26 (51%)
Repositório de revista / editor	1 (4%)	18 (75%)	4 (17%)	1 (4%)	24 (47%)
Repositório da minha instituição	0 (0%)	12 (55%)	9 (41%)	1 (5%)	22 (43%)
Repositório nacional	2 (9%)	11 (50%)	7 (32%)	2 (9%)	22 (43%)

Fonte: Dados da pesquisa (2021).

De forma complementar, buscou-se saber quais as razões do não compartilhamento ou compartilhamento parcial dos dados de pesquisa. Assim, a Tabela 15 traz os resultados para esse questionamento.

Tabela 15 - Quais suas razões para não compartilhar ou compartilhar parcialmente os seus dados de pesquisa?

	<b>ICSA</b>	<b>ITEC</b>	<b>Total de respostas</b>
Falta de financiamento	3 (5%)	2 (3%)	5 (4%)
Falta de padronização dos dados	4 (6%)	3 (4%)	7 (5%)
As pessoas não precisam deles	0 (0%)	3 (4%)	3 (2%)
Não tenho tempo para compartilhá-los	1 (2%)	5 (7%)	6 (4%)
Não há infraestrutura para compartilhá-los	8 (13%)	5 (7%)	13 (10%)
Eles não podem ser compartilhados	2 (3%)	4 (6%)	6 (4%)
Não me foi exigido	10 (16%)	10 (14%)	20 (15%)
Não tenho permissão para torná-los públicos	7 (11%)	8 (11%)	15 (11%)
Perderia o controle dos dados	2 (3%)	1 (1%)	3 (2%)
Preciso publicar os resultados da pesquisa antes de compartilhá-los	12 (19%)	15 (21%)	27 (20%)
Não sei como compartilhar	7 (11%)	4 (6%)	11 (8%)
Os dados não estão totalmente documentados	2 (3%)	5 (7%)	7 (5%)
Os dados estão em um formato que não permite esse compartilhamento	3 (5%)	3 (4%)	6 (4%)
Outros	3 (5%)	2 (3%)	5 (4%)

Fonte: Dados da pesquisa (2021).

Conforme apresentado na Tabela 15, a causa mais frequente para o não compartilhamento de dados é a opção: “Preciso publicar os resultados da pesquisa antes de compartilhá-los” com 20% (n=27) seguido da não exigência desta prática 15% (n=20). Na sequência, as razões que se destacam são: “não tenho permissão para torná-los público” com 11% (n=15) e “não há infraestrutura para compartilhá-los” com 10% (n=13).

Estes percentuais são percebidos nos dois institutos investigados. Outros motivos mais sinalizados são: “não sei como compartilhar” com 8% (n=11) e “falta de padronização” com 5% (n=7), além da opção “outros” com 4% (n=5). Nesta última, cinco docentes descreveram ser: falta de conhecimento aprofundado sobre os aspectos da ação, falta de criar o site para o compartilhamento, não senti necessidade em compartilhar, falta de cultura de compartilhamento e falta de hábito.

Os resultados demonstram a necessidade do estabelecimento de política interna sobre o assunto, assim como suporte técnico e de pessoal, conforme aponta estudo de Silva e Silveira (2019a) uma das necessidades do desenvolvimento da Ciência Aberta é o investimento não apenas em infraestruturas, políticas e tecnológicas, mas principalmente em recursos humanos com competências fundamentais para a implementação de serviços com vistas a amparar o cientista em cada fase do ciclo de pesquisa. Isso porque essa ação não consiste apenas em transferir todas as responsabilidades ao pesquisador, mas sim dar ferramentas e condições para que ele construa esse novo modo de fazer ciência com mais transparência, ética e colaboração.

Os pesquisadores também foram indagados sobre atitudes relacionadas ao compartilhamento de dados de pesquisa, numa escala entre “sempre, às vezes e nunca”. A Tabela 16 apresenta os resultados obtidos, nos quais apenas duas alternativas foram respondidas pela totalidade de participantes. São elas: “Outros podem acessar meus dados de pesquisa livremente” e “Falta de acesso aos dados gerados por outros pesquisadores ou instituições tem dificultado a realização das minhas pesquisas”.

Tabela 16 - As seguintes afirmativas se referem ao compartilhamento e ao acesso a dados de pesquisa

	<b>Sempre</b>	<b>Às vezes</b>	<b>Nunca</b>	<b>Tota de respostas</b>
Outros precisam da minha permissão para acessar meus dados de pesquisa	13 (27%)	28 (57%)	8 (16%)	49 (96%)
Outros podem acessar meus dados de pesquisa livremente	3 (6%)	26 (51%)	22 (43%)	51 (100%)
Falta de acesso aos dados gerados por outros pesquisadores ou instituições tem dificultado a realização das minhas pesquisas.	4 (8%)	31 (61%)	16 (31%)	51 (100%)
Minha instituição disponibiliza pessoal especializado para preparar os dados de pesquisa para serem compartilhados.	1 (2%)	10 (21%)	37 (77%)	48 (94%)
Minha instituição oferece as orientações específicas para o compartilhamento de dados de pesquisa.	4 (8%)	12 (24%)	33 (67%)	49 (96%)
Minha instituição disponibiliza um repositório para compartilhamento de dados de pesquisa.	4 (8%)	16 (33%)	29 (59%)	49 (96%)
Na minha área, os pesquisadores acessam livremente repositórios de dados de pesquisa.	3 (6%)	31 (63%)	15 (31%)	49 (96%)

Fonte: Dados da pesquisa (2021).

Com uma porcentagem alta, 77% (n=37) dos docentes apontam que a instituição “nunca” disponibiliza pessoal especializado para preparar os dados de pesquisa para serem compartilhados e para 67% (n=33) dos participantes da pesquisa, a instituição “nunca” ofereceu orientações específicas para o compartilhamento de dados de pesquisa.

Estes dados permitem inferir que os pesquisadores ainda não receberam o apoio para compartilhamento por parte da instituição, uma vez que a mesma não tem uma política de compartilhamento de dados de pesquisa, e por conseguinte, não tem orientação acerca desse modelo de gestão de dados, e prática de fazer ciência sob a perspectiva da Ciência Aberta.

Outro aspecto importante é em relação à satisfação na organização dos dados de pesquisas desses docentes. A Tabela 17 aponta que no geral os pesquisadores estão parcialmente satisfeitos com percentuais que variam entre 53% e 65% de respostas.

Tabela 17 - Quão satisfeito você está com a organização dos seus dados de pesquisa nas seguintes situações

	Satisfeito	Parcialmente satisfeito	Insatisfeito	Não sei responder	Total de respostas
Coleta de dados	17 (33%)	31 (61%)	2 (4%)	1 (2%)	51 (100%)
Descrição dos dados	16 (31%)	33 (65%)	1 (2%)	1 (2%)	51 (100%)
Análise dos dados	23 (45%)	27 (53%)	0 (0%)	1 (2%)	51 (100%)
Armazenamento dos dados durante a execução do projeto	16 (31%)	30 (59%)	5 (10%)	0 (0%)	51 (100%)
Preparação dos relatórios de pesquisa	20 (39%)	27 (53%)	4 (8%)	0 (0%)	51 (100%)
Recuperação dos meus dados	11 (22%)	32 (63%)	7 (14%)	1 (2%)	51 (100%)
Preservação dos dados após a conclusão do projeto	12 (24%)	33 (65%)	5 (10%)	1 (2%)	51 (100%)

Fonte: Dados da pesquisa (2021).

Em relação à familiaridade com a gestão de dados de pesquisa (Tabela 18), os respondentes afirmam ter pouco conhecimento sobre o tema, com 71% (n=36) das respostas. Em relação aos institutos, o termo “desconheço” possui uma distinção: para o ITEC essa taxa é 22% (n=6) enquanto que para o ICOSA, o percentual é de 4% (n=1).

Tabela 18 - Qual sua familiaridade com a gestão de dados de pesquisa?

	ICOSA	ITEC	Total
Conheço bastante	4 (17%)	4 (15%)	8 (16%)
Conheço pouco	19 (79%)	17 (63%)	36 (71%)
Desconheço	1 (4%)	6 (22%)	7 (14%)
<b>Total</b>			51 (100%)

Fonte: Dados da pesquisa (2021).

Pode-se inferir que embora os pesquisadores tenham um certo nível de conhecimento sobre dados de pesquisa e compartilhamento de dados (Tabela 6), quando se trata de gestão de dados de pesquisa, que é mais complexo, para a maioria dos pesquisadores, esse tema ainda é pouco familiar. Esse resultado evidencia os estudos de Lima (2020) onde 91% entrevistados apontaram “pouco conhecimento” ou “desconhecimento”; e o de Vanz *et al.* (2018) que apontaram 87,8% dos respondentes “conhecem pouco” ou “desconhecem” sobre a gestão de

dados de pesquisa, demonstrando a incipiência sobre a prática de compartilhamento de dados de pesquisa e o desconhecimento por parte dos pesquisadores sobre a gestão de dados de pesquisa.

Baseando-se na questão anterior, buscou-se também conhecer a experiência desses docentes com relação ao plano de gestão de dados (Tabela 19), em que podiam marcar quantas opções julgassem necessárias. Verificou-se que 60% (n=37) dos pesquisadores apontam que nunca elaboraram um PDG e que o maior percentual de elaboração do PDG é para a publicação em periódico com 16% (n=10).

Tabela 19 - Você já elaborou um plano de gestão de dados

	ICSA	ITEC	Total
Sim, por solicitação do meu principal financiador	1 (17%)	5 (83%)	6 (10%)
Sim, por solicitação da instituição principal que eu atuo	1 (33%)	2 (67%)	3 (5%)
Sim, para publicação em periódico	3 (30%)	7 (70%)	10 (16%)
Sim, por outras razões	2 (33%)	4 (67%)	6 (10%)
Nunca elaborei	19 (51%)	18 (49%)	37 (60%)

Fonte: Dados da pesquisa (2021).

Analisando os resultados entre os institutos, o ITEC apresentou um maior número de respostas apontando que já elaboraram um plano de gestão de dados, somando 18 resultados positivos. Este resultado, que pode estar relacionados com sua área de atuação e o seu principal financiador, pois, conforme mencionado pelos pesquisadores, o ITEC possui como maior financiador as agências nacionais (n=16), que tendem a exigir o PDG. Infere-se que a indicação de 60% (n=37) não terem elaborado um PGD seja por falta de conhecimento, falta de solicitação seja por parte da instituição em que atuam, financiadores ou a inexistência de política de gestão de dados para instruírem quanto a esse documento.

Ao questionar se os participantes utilizam algum repositório para acessar dados de pesquisa de outros pesquisadores, as respostas se equiparam, conforme se apresenta na Tabela 20, em que 44% (n=22) responderam positivamente e 56% (n=28) responderam de forma negativa.

Tabela 20 - Você utiliza algum repositório para acessar dados de pesquisa de outros pesquisadores?

	ICSA	ITEC	Total
Sim	9 (18%)	13 (26%)	22 (44%)
Não	14 (28%)	14 (28%)	28 (56%)
<b>Total</b>			50 (98%)

Fonte: Dados da pesquisa (2021).

Para identificar quais repositórios de dados os docentes utilizam para acessar dados de pesquisa, fez-se uma pergunta aberta, respondida por 20 participantes, dos 22 que afirmaram positivamente na questão anterior. Percebe-se que entre os docentes do ITEC e ICSA há uma diversidade de fontes, e que poucos indicam realmente repositórios de dados, como o IPEA DATA e *Case Western Reserve University Bearing Data Center*. Isso denota o desconhecimento por parte dos docentes sobre o que realmente são repositórios de dados de pesquisa.

Quando perguntado se os docentes já utilizaram em suas pesquisas dados abertos compartilhados por outros grupos de pesquisa (Tabela 21), verifica-se que: 47% (n=24) responderam que “não”, 45% (n=23) “sim” e 8% (n=4) julgaram não saber responder essa questão, conforme os dados apresentados na Tabela 22.

Tabela 21 - Você já utilizou em suas pesquisas dados abertos compartilhados por outros grupos de pesquisa?

	ICSA	ITEC	Total
Sim	9 (39%)	14 (45%)	23 (45%)
Não	14 (58%)	10 (42%)	24 (47%)
Não sei responder	1 (25%)	3 (75%)	4 (8%)
<b>Total</b>			51 (100%)

Fonte: Dados da pesquisa (2021).

É válido destacar que, em comparação com o compartilhamento de dados de pesquisa (Tabela 12), percebe-se uma equiparidade entre a prática de compartilhamento e o uso de dados de pesquisa compartilhados por outros pesquisadores, onde os resultados apontam: 30 afirmativas para a prática de compartilhamento e 23 afirmativas para a utilização dados de pesquisa de outros pesquisadores.

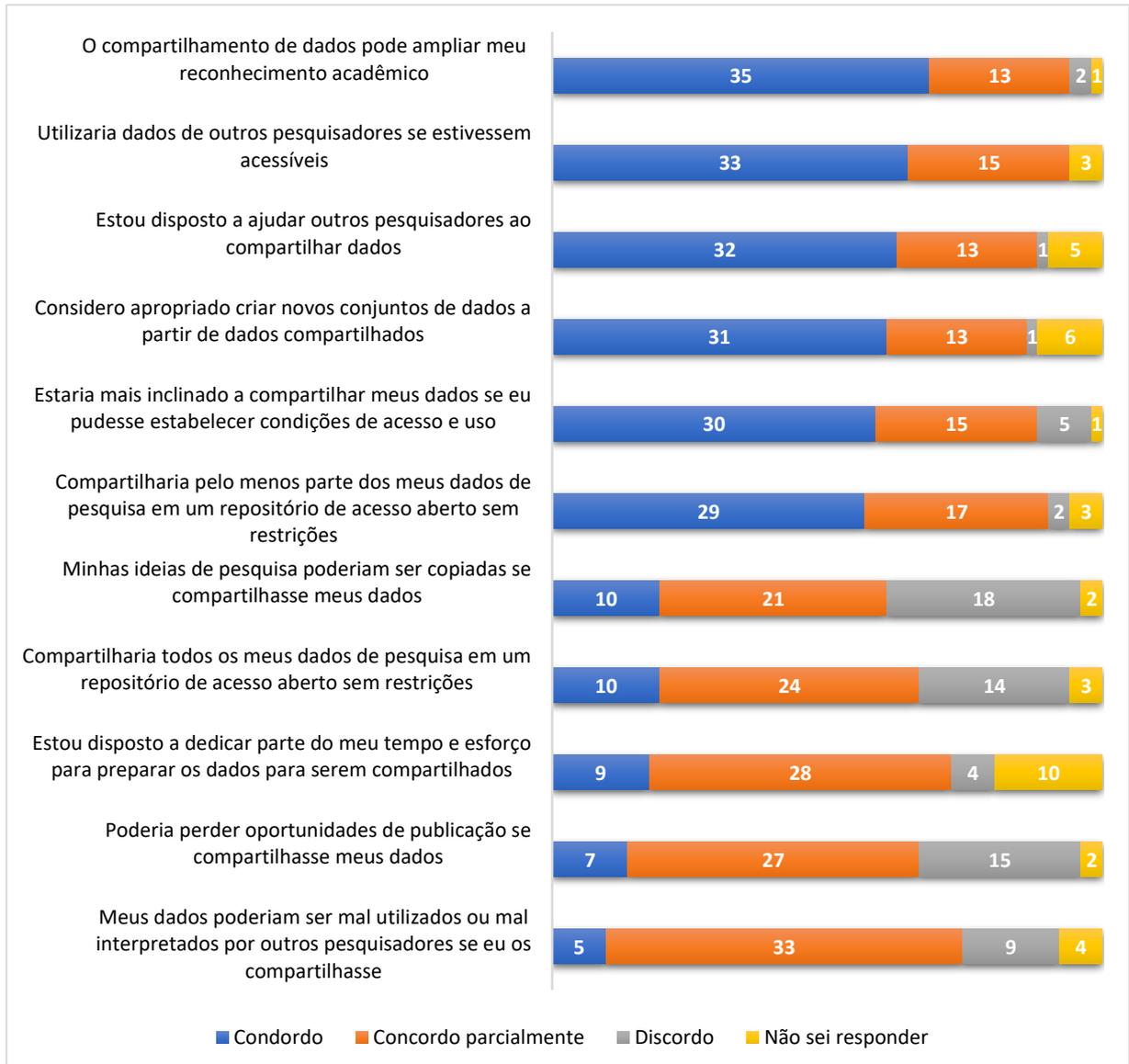
Somados em conjunto, esses resultados referentes ao compartilhamento de dados de pesquisa entre os docentes permanentes dos programas pós-graduação do ITEC e ICESA, mostram que os docentes possuem certa familiaridade com o tema, mas que falta uma orientação institucional sobre o assunto.

#### **4.2 Percepções em relação ao compartilhamento de dados de pesquisa**

Esta subseção foi reservada para apresentar os resultados acerca da percepção de compartilhamento de dados de pesquisa, correspondente às questões de 3.1 até 3.4 do questionário.

O Gráfico 1 apresenta o resultado de um conjunto de perguntas sobre compartilhamento de dados em uma escala de concordância de 1 a 4 (concordo, concordo parcialmente, discordo, não sei responder).

Gráfico 1 - Grau de concordância com cada uma das afirmativas sobre compartilhamento de dados



Fonte: Dados da pesquisa (2021).

Os níveis de concordância foram maiores com relação a alternativa “compartilhamento de dados pode ampliar meu reconhecimento acadêmico”, seguindo de “utilizaria dados de outros pesquisadores se estivessem acessíveis”, “estou disposto a ajudar outros pesquisadores ao compartilhar dados”, “considero apropriado criar novos conjuntos de dados a partir de dados compartilhados”, “estaria mais inclinado a compartilhar meus dados se eu pudesse estabelecer condições de acesso e uso” e “compartilharia pelo menos parte dos meus dados de pesquisa em um repositório de acesso aberto sem restrições”.

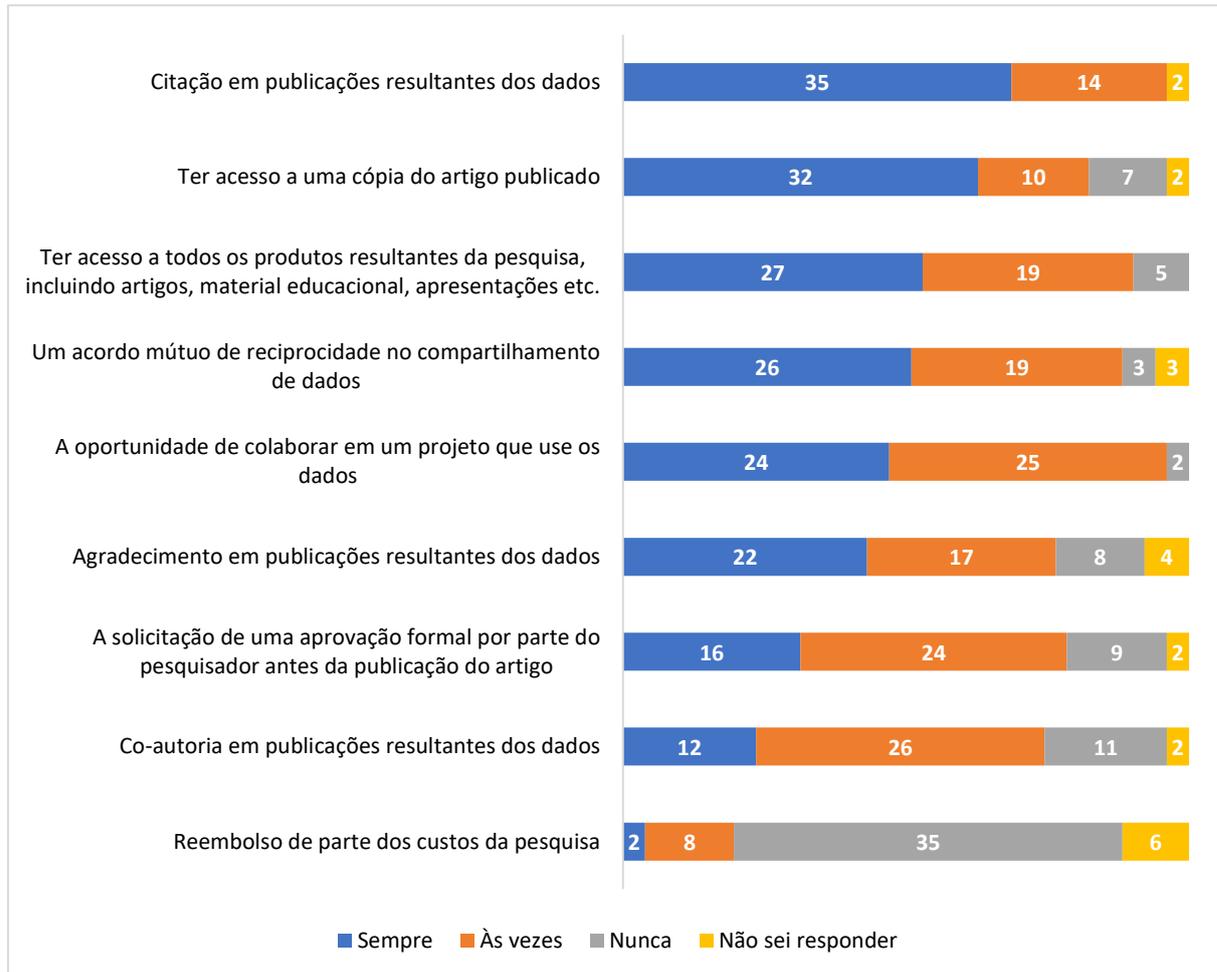
Os níveis de concordância parcial foram maiores com relação as alternativas: “meus dados poderiam ser mal utilizados ou mal interpretados por outros pesquisadores se eu os compartilhasse”, seguido de “estou disposto a dedicar parte do meu tempo e esforço para preparar os dados para serem compartilhados” e “poderia perder oportunidades de publicação se compartilhasse meus dados”.

Esse resultado mostra que os docentes não demonstram resistência em relação ao compartilhamento, corroborando com o apontado na pesquisa de Vanz *et al.*, (2018) referente a receptividade à ideia do compartilhamento e reuso de dados por parte dos pesquisadores brasileiros.

Destaca-se ainda um número significativo de “não sei responder” para a opção “Estou disposto a dedicar parte do meu tempo e esforço para preparar os dados para serem compartilhados”, ressaltando, desse modo, a importância de diretrizes para o gerenciamento de dados de pesquisa e pessoal capacitado para auxiliar o pesquisador nessa prática. Bem como oferta de treinamentos e disseminação das ideias de compartilhamento entre a comunidade universitária.

Outra questão abordada foi em relação a recompensa caso outros pesquisadores usem seus dados de pesquisa, conforme apresenta o Gráfico 2.

Gráfico 2 - Retorno ao usarem os dados de pesquisa



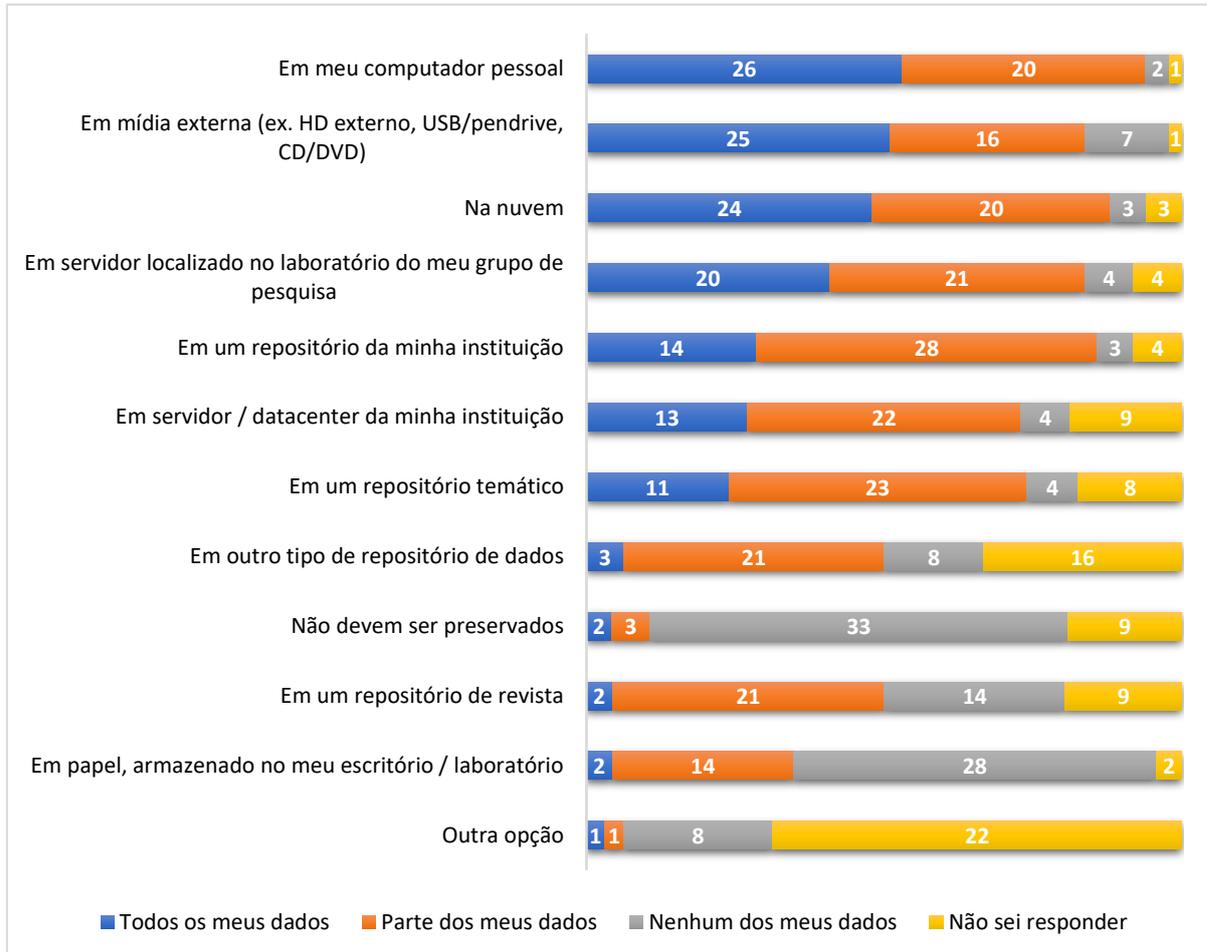
Fonte: Dados da pesquisa (2021).

No Gráfico 2, pode ser visualizado os resultados relacionados a forma de retorno que os docentes esperam receber em caso de compartilhamento de dados. Verifica-se que os mesmos indicam na escala, que “sempre” desejam que ao terem seus dados de pesquisa utilizados, sejam citados nas publicações resultantes dos dados (n=35), bem como ter acesso a uma cópia do artigo publicado (n=32). Em contrapartida, em relação ao reembolso de parte dos custos da pesquisa, a maioria (n=35) dos docentes indicaram que “nunca” esperam que este seja o retorno quando outros pesquisadores utilizarem os seus dados de pesquisa.

Quanto ao conhecimento acerca dos locais em que os docentes consideram oportunos para a preservação dos dados de pesquisa, de acordo com Gráfico 3, observa-se que a maioria dos docentes consideram que todos os seus dados de pesquisa devem ser preservados em

computador pessoal (n=26), seguido de mídia externa (n=25), nuvem (n=24) e em servidor localizado no laboratório do próprio grupo de pesquisa (n=20).

Gráfico 3 - Onde você considera que seus dados de pesquisa devem ser preservados?



Fonte: Dados da pesquisa (2021).

No caso do compartilhamento de partes dos dados pesquisa, observa-se que os pesquisadores julgam ser preservados em repositório da instituição (n=28), repositório temático (n=23) e em repositório da própria instituição (n=22). Ainda sobre a questão de preservação dos dados, chamou atenção o fato da alta indicação para “não devem ser preservados”, com 33 respostas, ou seja, mais da metade dos participantes.

Esse resultado é diferente do recomendado na literatura, baseado em Uhler (2010) que estabelece três princípios para preservação de dados de pesquisa, sendo o primeiro: “o valor dos dados aumenta com seu uso”; o segundo “público (e, em grande medida, com financiamento público) “... a informação quer ser gratuita ...” e o terceiro “Os recursos digitais não

sobreviverão ou permanecerão acessíveis por acidente”. Desse modo, para o autor os dados de pesquisas produzidos, precisam ser preservados e disponibilizados o mais aberto possível.

Buscou-se saber, também, qual a percepção dos pesquisadores com relação a quem deve oferecer apoio a gestão de dados de pesquisa. Como mostra a Tabela 22, 70% (n=35) dos participantes da pesquisa, ou seja, a maioria dos docentes, considera que deve ser a instituição de atuação do pesquisador. No entanto, nota-se que 14% (n=7) dos participantes não souberam responder.

Tabela 22 - Quem deve oferecer apoio a gestão de dados de pesquisa

	ICSA	ITEC	Total
Agência de financiamento	3 (75%)	1 (25%)	4 (8%)
Agência não governamental	0 (0%)	0 (0%)	0 (0%)
Associação científica	0 (0%)	0 (0%)	0 (0%)
Ente governamental	1 (33%)	2 (67%)	3 (6%)
Instituição de atuação do pesquisador	17 (49%)	18 (51%)	35 (70%)
Não acho necessário apoio	1 (100%)	0 (0%)	1 (2%)
Não sei responder	1 (14%)	6 (86%)	7 (14%)

Fonte: Dados obtidos na resposta a questão 3.3 do questionário

Nesse sentido, a presente pesquisa mostra-se relevante no contexto institucional, que conforme respostas dos próprios docentes, para 70% (n=35) deles, a UFPA que deve oferecer apoio para a gestão de dados de pesquisa.

Juntos, esses resultados, obtidos por meio do questionário, possibilitou ter uma noção sobre o entendimento e as práticas dos docentes permanentes dos programas de pós-graduação do ITEC e ICSA sobre o compartilhamento de dados de pesquisa, sinalizando o papel da UFPA como instituição de pesquisa e a necessidade de diretrizes formais sobre o compartilhamento de dados de pesquisa nos preceitos da ciência aberta.

#### 4.3 Aprofundamento da investigação por meio da entrevista

Para o aprofundamento da temática deste estudo, adotou-se como procedimento a entrevista, visando esclarecer e/ou confirmar os dados coletados com a aplicação do

questionário, estratégia pensada para qualificar as questões do referido instrumento. Para esta etapa da pesquisa, destaca-se que ela foi eletiva, sendo assim, apenas 10 dos 51 participantes aceitaram prosseguir com a entrevista.

Visto que os dados de pesquisa são essenciais no contexto da pesquisa e inovação, e as universidades são grandes protagonistas dessa produção e disseminação do conhecimento, os entrevistados foram indagados quanto ao seu entendimento sobre o significado de acesso aberto aos dados de pesquisa. Em sua maioria, os participantes apresentaram entendimento sobre acesso aberto aos dados de pesquisa, conforme relata o Entrevistado 5 (grifo nosso):

Eu acredito que acesso aberto a dados de pesquisa seria tudo aquilo que nós coletamos dentro de um projeto de pesquisa, por exemplo, e depois construímos nossos relatórios e, por fim, preparamos nossos artigos científicos, *né?* Então **esses dados brutos foram coletados eles ficariam à disposição para outros pesquisadores em um certo repositório aberto**. É esse entendimento, apesar de que eu não tenha muita intimidade com esse conceito.

Como visto, trata-se de uma compreensão, que vai em direção à definição usada pela Comissão Europeia ([20--]) ao afirmar que ele refere-se ao direito de acessar e reutilizar dados de pesquisa de acordo com os termos e condições estabelecidas em sua submissão, o que foi identificado no relato de outro entrevistado:

Acesso aberto aos dados de pesquisa, significam conjunto de ferramentas, normas, resoluções e corpo técnico, que visem instrumentalizar que os **dados de uma determinada pesquisa fiquem acessíveis a um público irrestrito**. Para isso utiliza-se de resoluções e normas, por exemplo [...]. Nesse sentido então a questão de aberto [...] ela tem muita ligação com a ideia de *Open access*, que é um contexto maior que dados científicos. (ENTREVISTADO 3, grifo nosso)

Já o Entrevistado 8, entende que:

[...] é eu poder acessar o banco de dados de um pesquisador que publicou um artigo do meu interesse, não somente ter acesso ao dado bruto, mas também ter acesso ao banco de dados que ele produziu a partir daquele dado bruto.

De forma geral, as falas da maioria dos participantes denotam conhecimento empírico sobre acesso aberto a dados de pesquisa.

Na sequência buscou-se conhecer o quanto os entrevistados acham úteis os dados que produzem. Percebeu-se a importância entre os entrevistados, porém, no desenvolvimento de

suas pesquisas não existe tanta preocupação quanto a esse aspecto, conforme relata o entrevistado 5:

**Nunca tinha pensado a respeito disso antes dessa pesquisa**, na verdade. Então, a priori, eu não via muito uso para outros pesquisadores dos meus dados, então é algo que eu ainda preciso começar a pensar. De forma objetiva, eu acredito que ainda não seja muito útil porque quando eu faço as minhas pesquisas eu não penso nessa transferência dos dados em si, mas sim dos resultados da pesquisa, o que não quer dizer, que não possa ser pensado futuramente algo que realmente possa ter uma utilidade dos dados para comunidade [científica] (ENTREVISTADO 5, grifo nosso).

Em contrapartida, o entrevistado 8 relata: “Eu acho que eles sejam muito úteis para outros pesquisadores, eu só não sei medir exatamente qual utilidade, *né?* Mas eu diria que eles seriam muito úteis para produzir novas análises, para somar com outros dados que os pesquisadores já tenham”.

O Entrevistado 4 explica o porquê de talvez os dados produzidos não sejam tão úteis: “Eu penso que: primeiro cada pesquisador tem sua forma de coletar os dados, alguns são organizados e os outros não são, e isso pode ser um empecilho! Por que? Porque a lógica como eu organizo as coisas não é necessariamente a lógica de todo mundo, compreende?”

A questão da utilidade dos dados de pesquisa, conforme descrito na literatura, depende, em grande parte, da qualidade dos dados em si, que os geradores dos dados devem prestar especial atenção, visando garantir a conformidade com os padrões explícitos de qualidade (ORGANISATION..., 2007). Tais padrões ainda não existem. Instituições e associações de pesquisa devem se envolver com sua comunidade de pesquisa para constituir seu desenvolvimento. Embora todas as áreas de pesquisa possam se beneficiar da melhoria da qualidade dos dados, algumas exigem padrões muito mais rigorosos do que outras. Por esse motivo, padrões universais de qualidade de dados não são práticos. As normas devem ser desenvolvidas em consulta com pesquisadores, para garantir que o nível de qualidade e que a precisão atenda às necessidades das várias áreas.

Nessa perspectiva, os entrevistados foram sondados quanto a sua percepção em relação ao compartilhamento de dados em seus programas de pós-graduação. Constatou-se que é ínfima essa questão, visto que quando ocorre é apenas entre um grupo fechado ou nem existe, conforme destacam os entrevistados:

[...] entre professores dentro do programa, ocorre em alguns níveis, por exemplo: o compartilhamento não de resultado de pesquisa, mas por exemplo, compartilhamento de informações sobre revistas, isso é bem compartilhado dentro do programa, mas em termo de dados e resultado do professor, ainda não é feito. Eu acho importante, mas ainda não é feito. Não existe uma plataforma para que a gente possa se conectar (ENTREVISTADO 7)

Já o Entrevistado 8 relata:

[...] eu sou da linha de pesquisa em urbanismo. A gente trabalha muito articulado. Nossos professores... a gente é do mesmo laboratório e do mesmo grupo de pesquisa. Então de certa maneira os dados... eles são compartilhados, mas não há ainda uma centralização desses dados e tampouco um incentivo do programa de pós para que os dados das linhas de modo geral assumam um caráter mais compartilhado. Então, **existe esse compartilhamento entre nós** da linha de Urbano, mas ainda, também, **sem nenhuma organização**. É como se assim: eu perguntar para o professor x se ele me empresta a *tal*, assim eu sei que eu terei, mas eu não tenho acesso ao servidor dele, tampouco ao e-mail dele, tão pouco ao Gmail, não tenho nada centralizado do grupo para fazer isso (ENTREVISTADO 8, grifo nosso).

Na visão do Entrevistado 9 “Ainda é uma prática pouco existente de compartilhamentos”, assim como para o Entrevistado 10: “Pouca prática, acho que quase nenhuma prática”. Para o Entrevistado 3, a questão do compartilhamento de dados de pesquisa ainda é:

Muito fraco, um discurso muito teórico ainda, importante, diga-se de passagem [...] eu percebo dentro da área como um todo, é ainda um discurso mais no marco teórico e sensibilização do que de fato você encontrar pesquisas na nossa área publicada em consonância com os conjuntos de dados abertos. Ainda são iniciativas muito tímidas na prática

Os entrevistados expuseram os motivos os quais estimularia o compartilhamento de dados de pesquisa, em suas visões. Para o Entrevistado 7 o “Primeiro motivo é o desenvolvimento da área. Segundo o desenvolvimento institucional, o aprimoramento, a busca por novos pesquisadores na área, então... não só alunos, mas para que outros pesquisadores possam aprimorar as pesquisas que já estão em andamento”. Por outro lado, o Entrevistado 5 sinalizou: “Eu acredito que não precisa de muito para estimular. Na verdade, eu acho que se houver um movimento que possa explicar como os dados dos pesquisadores servem para outros pesquisadores [...] não seria difícil de conseguir adesão para esse tipo de coisa [...]”. Já o Entrevistado 1 diz que: “Os motivos que estimulariam, não só o programa que atuo, mas acho que qualquer outro programa no Brasil, é ter dados para validação de metodologias, *né?*”.

Dessa forma, também foi sondado quais motivos inibiriam o compartilhamento de dados. Segundo o Entrevistado 2 (pesquisador do ITEC na área de engenharia): “No nosso caso é o financiador, regra do financiador, certo? Do governo, da [iniciativa] privada ou, logicamente, questões éticas, *né?*”. Para o Entrevistado 1:

O que inibe são os acordos contratuais de empresas privadas [...]. Outro bloqueio: eu não vejo que seja um impedimento, por que os pesquisadores têm uma certa maneira, um cunho muito aberto, *né?* Você recebe um e-mail, por exemplo, de um colega, seja nacional ou estrangeiro em busca de dados, é muito comum fornecimento dados.

No que tange ao “compartilhamento de dados por exigência do financiador”, os entrevistados concordaram que não veem objeção nessa solicitação, conforme relata o Entrevistado 6:

Eu não teria problema com isso, a menos que esses dados envolvessem informações pessoais, alguma coisa nesse sentido. Mas aí a gente teria que **levar em consideração as legislações**, por exemplo, a lei geral de proteção de dados e assim por diante, e eu acredito que se uma agência pressupor isso para financiar o seu trabalho, ela já vai levar em consideração esses aspectos também (ENTREVISTADO 6, grifo nosso).

Para o Entrevistado 4: “Se uma condição é eu ter que cumprir? Acabou! Você sabe que no contrato desses financiamentos, ou você faz ou você não recebe o dinheiro.” O Entrevistado 9 considera que poderia ser um processo natural “[...] se fosse uma exigência da agência, seria um procedimento natural e os pesquisadores já teriam esse hábito [...]”

Como visto, quando perguntado se a instituição tornasse obrigatório o compartilhamento de dados de suas pesquisas, os docentes também não viram objeção. O único ponto contestado, foi o termo “obrigatório”, como observado na resposta do Entrevistado 1:

Não. Aí qualquer **imposição ela não é bem-vinda**. Obrigatoriamente, primeiro que iria ferir algum contrato privado, então o pesquisador ficaria... a sua universidade obriga o compartilhamento de dados, enquanto seu financiador obriga a não compartilhar, *como é que sai?* Em uma universidade livre acho que com certeza não haveria isso. Isso não estou de acordo (ENTREVISTADO 1, grifo nosso).

Mas, para o Entrevistado 9: “Seria natural. Se fosse algo da cultura organizacional terminaria sendo natural para os pesquisadores compartilhar esses dados, *né?* Como uma

política de compartilhamento, por exemplo, uma política de disseminação dos dados.” Nessa mesma perspectiva, o Entrevistado 8 relatou:

Eu **ia achar perfeito se ela tornasse obrigatório**, ela [a Universidade] ia obrigatoriamente estimular os programas, os programas iam começar a discutir sobre isso e ela, certamente, a instituição, daria o suporte que a gente precisa para fazer o compartilhamento, não só no ponto vista técnico, né? Nuvens, *essas coisas* [...] mas acima de tudo de formação e apoio técnico de gente que entenda de banco de dados para ajudar no processo (ENTREVISTADO 8, grifo nosso).

Para conhecimento de como reagiriam caso os editores de periódicos também exigissem os dados de pesquisa na submissão do artigo, os entrevistados dizem que “isso já é uma tendência” (ENTREVISTADO 2). Para o Entrevistado 5:

Alguns já exigem, mas como eu falei exigem, mas não para divulgação. Exigem muito mais como um processo de avaliação do periódico e eu não vejo problema, mesmo que seja para divulgação, uma vez que seja respeitado um prazo aí de que o autor possa trabalhar os seus próprios resultados.

Assim como para o Entrevistado 7:

Acho válido, acho que é importante, muito importante para conseguir chegar à validade do que foi publicado. Algumas revistas de alto impacto fazem isso. Você pode enviar isso como pesquisador, depende muito do pesquisador. Quando você submete o trabalho tem uma guia lá “materiais suplementares” você pode tornar disponível na sua pesquisa.

Com a entrevista, também buscou-se saber mais sobre a experiência no compartilhamento de dados de pesquisa desses pesquisadores, para mapear as práticas desses participantes. Em seguida, são apresentados os relatos dos Entrevistados 5, 7 e 9:

Quando a gente apresenta os resultados, eventualmente a revista pode requerer a base de dados da coleta dos dados que foram realizados e foi o que eu fiz, **disponibilizei numa planilha essa base de dados**, assim como instrumento de pesquisa, né? E foi isso né, isso deve ter sido aberto para os avaliadores e para os avaliadores fazerem qualquer observação que fosse necessário, mas não obtive nenhum feedback a respeito dos dados em si, somente do artigo (ENTREVISTADO 5, grifo nosso).

Compartilhei uma vez [...] eu fiz o compartilhamento... na verdade foi o compartilhamento de um **instrumento de avaliação que utilizam questionário** e aí teve um outro pesquisador de fora do país que solicitou o questionário, ele solicitou o questionário completo, então nós fizemos esse compartilhamento (ENTREVISTADO 7, grifo nosso).

Eu já tive uma experiência enquanto participava do grupo de pesquisa na Unesp onde nós temos vários pesquisadores no mesmo grupo e cada um em nível, né? Tinham algumas pessoas da graduação, outras do mestrado, outras

do doutorado, outros professores. Então os **dados gerados eram compartilhados entre o grupo todo**, mesmo porque para todo mundo entender como estava acontecendo o procedimento de análise desses dados, do tratamento de dados num todo. Serviu como uma experiência positiva (ENTREVISTADO 9, grifo nosso).

Em seguida, buscou saber dos entrevistados se a capacitação e a sensibilização poderiam ajudar na adoção da prática de compartilhamento de dados de pesquisa. Percebeu-se algumas divergências. Por exemplo, para Entrevistado 1 “[...] não é isso o fundamental, não é o treinamento. O fundamental é ter uma área que necessita de compartilhamento de dados.” Já para o Entrevistado 8 “Eu acho que a capacitação e a sensibilização são fundamentais para essa prática de compartilhamento, mas eu acho que mais do que a sensibilização a gente precisa avançar na capacitação de coleta de dados, das bases *open data* que a gente tem disponível”, assim como para o Entrevistado 3:

Sim, eu acredito que é um conjunto de processos, né? Se a gente for olhar mais amplo. Na verdade, que tem que se criar uma política de informação diferente, né? Nos mais variados entes, que trabalham nesse tema então: agência de fomento, universidades, institutos de pesquisa, pesquisadores, bibliotecários arquivistas e técnicos administrativos [...] se tudo isso for “maestrado”, no sentido de dar um direcionamento mais coeso eu vejo que esses procedimentos de capacitação são importantes. [...] Eu creio que essas capacitações de chão, locais, elas podem contribuir bastante e eu já vejo que essas iniciativas já existem, a questão de fato é criar uma política de informação e alocar recursos, não só recursos financeiros, mas também recursos de tempo para esse servidor que ampara as pesquisas junto aos pesquisadores para eles se capacitarem para poder auxiliar os pesquisadores na normalização, na estética, na forma desses conjuntos dados, para ajudá-los a publicarem, visto que os pesquisadores, às vezes, eles até se sensibilizam mas não têm tempo devido às demandas que eles têm pra organizar tudo isso e fazer um envio mais coeso, por exemplo, ao repositório institucional da universidade.

Esses resultados subsidiaram a elaboração de um panorama das percepções e práticas por área do conhecimento/por programa de pós-graduação. O Quadro 7 apresenta os resultados do ICSA.

Quadro 7 - Panorama das percepções e práticas de compartilhamento de dados de pesquisa (ICSA)

ICSA	
<b>Percepções</b>	Para parte desses docentes existe o entendimento empírico sobre o compartilhamento, porém, faltam mais diretrizes e infraestrutura para que essas práticas ocorram de fato, visto que ainda está muito no campo teórico.

<b>Práticas</b>	<p>Pode-se identificar que 52% dos entrevistados nunca compartilharam dados de pesquisa. Já os que praticaram, foi por meio de publicações em periódicos ou anais de eventos. Além disso, demonstraram que também compartilham seus dados de pesquisa por meio repositórios temáticos, revista, própria instituição e nacional.</p> <p>Quanto às razões para o não compartilhamento, destacamos a falta de infraestrutura, a não exigência, precisam publicar os resultados da pesquisa antes de compartilha-los e não sabem como partilhar.</p> <p>Ressaltamos que entre esses docentes 73% desses docentes nunca elaboraram um plano de gestão de dados e 61% não utilizam repositórios para acessar dados de pesquisa de outros pesquisadores, assim como 58% nunca utilizou dados de pesquisa gerados por outros pesquisadores.</p>
-----------------	---

Fonte: Dados da pesquisa (2021)

Elaborou-se, ainda, o panorama baseado nas respostas dos participantes do ITEC, conforme Quadro 8.

Quadro 8 - Panorama das percepções e práticas de compartilhamento de dados de pesquisa (ITEC)

<b>ITEC</b>	
<b>Percepção</b>	<p>Parte desses docentes não demonstraram resistência quanto ao compartilhamento, mas sim a preocupação em como fazer de forma correta. Outra questão está relacionada a credibilidade que essa abertura poderá gerar às pesquisas científicas e quanto a autorização para torna-los público, visto que para os pesquisadores do ITEC o número de financiamento por empresas privadas é maior do que para os ICSA.</p>
<b>Prática</b>	<p>Dentre esses pesquisadores, 79% disseram já ter compartilhado dados de pesquisa, seja por solicitação do financiador, iniciativa própria e em periódico ou anais de evento. Os locais citados como indicação para compartilhar parte dos seus dados foram: repositório temático, de revista, da própria instituição e nacional. As razões para o não compartilhamento se dão pelo fato de precisarem publicar os resultados antes de compartilha-los, por não ser exigido e não possuir permissão para torna-los público. Apesar de possuírem maior</p>

	prática, cerca de 78% desconhecem ou possuem pouco conhecimento sobre gestão de dados, como também 46% apontam que nunca elaboraram um plano de gestão de dados.
--	--

Fonte: Dados da pesquisa (2021)

Mediante os resultados percebe-se que entre esses docentes há conhecimento sobre a temática, mas carece de maiores discussões sobre o assunto nos fundamentos da ciência aberta e em âmbito institucional. Ademais, os participantes da presente pesquisa não demonstraram resistência em relação ao compartilhamento de dados, mas apontaram falta de diretrizes, infraestrutura e suporte para execução dessa prática.

## 5 CONCLUSÕES

Esta pesquisa foi realizada tendo em vista o cenário da Ciência Aberta, envolvendo discussões que englobam os dados de pesquisa e diante de sua importância e impacto na construção do conhecimento e comunicação científica. Percebendo a contemporaneidade do tema, teve-se a preocupação com o gerenciamento dos dados de pesquisa produzidos no âmbito da UFPA, por ser responsável em gerar, usar e disseminar conhecimento. Desse modo, buscou-se conhecer a percepção e prática de compartilhamento de dados dos docentes permanentes do ITEC e ICSA.

Entende-se que cada área do conhecimento possui diferentes formas na geração de dados de pesquisa. Assim, esta pesquisa não buscou fazer esse comparativo, mas foi importante a escolha de áreas distintas para compreender os seguintes objetivos específicos: (a) Identificar as percepções dos docentes permanentes dos programas de pós-graduação ITEC e do ICSA sobre compartilhamento de dados de pesquisa; (b) Mapear as práticas de compartilhamento de dados de pesquisa; (c) Apresentar o panorama das percepções e práticas por área do conhecimento/por programa de pós-graduação.

No que tange ao objetivo de identificar as percepções dos docentes permanentes desses dois programas, mediante questionário e entrevista, pode-se inferir que para esses docentes: o compartilhamento de dados possibilita ampliar seu conhecimento; auxilia na construção de novos dados, a partir de dados primários; a recompensa que esperam é serem citados nas publicações que utilizarem seus dados e o compartilhamento traria mais credibilidade e transparência às pesquisas.

Por meio da entrevista constatou-se que esses docentes possuem um bom entendimento em relação à temática, embora os resultados do questionário tenham apontado que eles desconhecem a Ciência Aberta.

Salienta-se, também, que no entendimento dos participantes, os dados por eles produzidos podem ser úteis para outros pesquisadores, embora desconheçam como proceder. Além disso, apontam a importância do compartilhamento de dados de pesquisa, seja por solicitação de agência de fomento, instituição a qual faz parte ou para publicação em periódicos. Assim como, o apoio da instituição para a gestão desses dados. Considera-se a abordagem do

compartilhamento de dados de pesquisa nos princípios da Ciência Aberta, que ainda é embrionária na UFPA. Dessa forma, esta pesquisa mostrou a incipiência sobre essa temática na universidade, a falta de infraestrutura para armazenamento desses dados e a ausência de política e diretrizes também são pontos inexistentes.

Já em relação ao mapeamento das práticas de compartilhamento de dados de pesquisa, conforme mencionado anteriormente, essa prática quase não existe nos programas e quando ocorrem é de forma centralizada, por exemplo: grupos fechados ou em publicação de artigos científicos. Vale ressaltar que apesar do envio dos dados, é publicado apenas o artigo em si.

Os docentes não demonstram barreiras para a questão do compartilhamento, porém, algumas questões são levantadas, por exemplo: quando os dados forem oriundos de pesquisa financiada, questões éticas, assim como uma política para padronização desses dados, visto que cada pesquisador possui uma forma de organizar suas informações e para haver compartilhamento esses dados precisam ser entendidos por outros pesquisadores, assim como, os erros que esses dados podem possuir.

Além disso, por meio desse grupo amostral, notou-se que os docentes utilizam seu computador pessoal para armazenamento dos seus dados de pesquisa, mostrando assim o desconhecimento deles em relação a ferramentas específicas com a finalidade de preservação de dados de pesquisa.

Mediante os resultados obtidos no presente estudo, pode-se elaborar um panorama das percepções e práticas, em que se denota conhecimento por parte desses pesquisadores, assim como práticas de compartilhamento de dados, mas não de modo formal, com instrumentos mencionados pela Ciência Aberta.

Por fim, esta pesquisa pode ser continuada considerando sua relevância quanto aos dados de pesquisa para a instituição. Destaca-se que novos estudos podem ser realizados, cujos resultados inferem em proposta de uma política de compartilhamento de dados, bem como, estudos sobre implementação do compartilhamento de dados na instituição.

## 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. Disponível em: [http://livroaberto.ibict.br/bitstream/1/1060/1/Ciencia%20aberta\\_questoes%20abertas\\_PORTUGUES\\_DIGITAL%20\(5\).pdf](http://livroaberto.ibict.br/bitstream/1/1060/1/Ciencia%20aberta_questoes%20abertas_PORTUGUES_DIGITAL%20(5).pdf). Acesso em: 07 out. 2021.
- ALBAGLI, Sarita; APPEL, Andre Luiz; MACIEL, Maria Lucia. E-Science, ciência aberta e o regime de informação em ciência e tecnologia. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.7, n.1, jan./jun. 2014. Disponível em: <https://ridi.ibict.br/bitstream/123456789/854/1/124-540-1-PB.pdf>. Acesso em: 07 out. 2021.
- ANGLADA, Lluís; ABADAL, Ernest. ¿Qué es la ciencia abierta? **Anuario ThinkEPI**, v. 12, p. 292–298, 2018. DOI: [doi.org/10.3145/thinkepi.2018.43](https://doi.org/10.3145/thinkepi.2018.43). Disponível em: <https://recyt.fecyt.es/index.php/ThinkEPI/article/view/thinkepi.2018.43>. Acesso em: 25 maio 2021.
- AVENTURIER, Pascal. **Dados de pesquisa**: descrição do ciclo de vida de gestão de dados A publicação científica, 2016. Disponível em: <https://publicient.hypotheses.org/1359>. Acesso em: 27 maio. 2021.
- BOAVENTURA, E. M. **Metodologia da Pesquisa**: monografia, dissertação e tese. São Paulo: Atlas, 2004.
- BORGMAN, Christine L. **Research Data**: Who Will Share What, with Whom, When, and Why? Rochester, NY: Social Science Research Network, 2010. DOI: [10.2139/ssrn.1714427](https://doi.org/10.2139/ssrn.1714427). Disponível em: <https://papers.ssrn.com/abstract=1714427>. Acesso em: 16 ago. 2021.
- BORGMAN, Christine L. The conundrum of sharing research data. **Journal of the American Society for Information Science and Technology**, v. 63, n. 6, p. 1059–1078, 2012. DOI: <https://doi.org/10.1002/asi.22634>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.22634>. Acesso em: 30 maio 2021
- BUDAPEST OPEN ACCESS INITIATIVE. Dez anos da Iniciativa de Budapeste em Acesso Aberto: a abertura como caminho a seguir. 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese-brazilian-translation>. Acesso em: 9 out. 2020.
- CAREGNATO, Sonia Elisa *et al.* Práticas e percepções dos pesquisadores brasileiros sobre serviços de acesso aberto a dados de pesquisa. **Liinc em Revista**, Rio de Janeiro, v. 15, n. 2, p. 121-141, 2019. DOI: [10.18617/liinc.v15i2.4771](https://doi.org/10.18617/liinc.v15i2.4771). Disponível em: <http://revista.ibict.br/liinc/article/view/4771>. Acesso em: 2 nov. 2020.
- CARVALHO, Érika Rayanne Silva. **Diferenças na produção, compartilhamento e (re)uso de dados**: a percepção de pesquisadores da Química, Antropologia e Educação. 2018. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2018. Disponível em: <https://repositorio.unb.br/handle/10482/32559>. Acesso em: 08 mar. 2021.

CARVALHO, Érika Rayanne Silva; LEITE, Fernando César Lima. Diferenças na produção, no compartilhamento e no (re)uso de dados de pesquisa: a percepção de pesquisadores de Química, Antropologia e Educação. **Em Questão**, Porto Alegre, v. 25, n. 3, p. 321–347, set./dez. 2019. DOI: 10.19132/1808-5245253.321-347. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/view/89449>. Acesso em: 08 mar. 2021.

COMISSÃO EUROPEIA. **H2020 Online Manual: Open access & Data management**. [20--]. Disponível em: [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm). Acesso em: 17 jul. 2021.

CÓRDULA, Flavio Ribeiro; ARAUJO, Wagner Junqueira. O compartilhamento de dados científicos na era do e-science. *In*: DIAS, Guilherme Ataíde (org.). **Dados científicos: perspectivas e desafios**. João pessoa: EDUFPA, 2019. p. 189–207. Disponível em: [O-compartilhamento-de-dados-cientificos-na-era-do-e-science.pdf](https://www.researchgate.net/publication/351111111) (researchgate.net). Acesso em: 30 maio 2020.

COSTA, Sely M. S. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 39–50, maio/ago. 2006. DOI: 10.1590/S0100-19652006000200005. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652006000200005&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000200005&lng=pt&tlng=pt). Acesso em: 14 out. 2020.

COSTA, Michelli Pereira; LEITE, Fernando César Lima. Open access in the world and Latin America: a review since the Budapest Open Access Initiative. **Transinformação**, Campinas, v. 28, n. 1, p. 33-46, abr. 2016. DOI: 10.1590/2318-08892016002800003. Disponível em: <https://repositorio.unb.br/handle/10482/19433>. Acesso em: 30 ago. 2021.

DATAONE. **Data life cycle**. DataONE. [2019]. Disponível em: <https://old.dataone.org/data-life-cycle>. Acesso em: 27 maio. 2021.

DECLARAÇÃO de Berlim sobre Acesso Livre ao Conhecimento nas Ciências e Humanidades. [2003]. Disponível em: [https://openaccess.mpg.de/67693/BerlinDeclaration\\_pt.pdf](https://openaccess.mpg.de/67693/BerlinDeclaration_pt.pdf). Acesso em: 26 out. 2020.

DESLAURIERS, Jean-Pierre; KÉRISIT, Michèle. O delineamento de pesquisa qualitativa. *In*: POUPART, Jean et al. **A pesquisa qualitativa: enfoques epistemológicos e metodológicos**. Petrópolis, RJ: Vozes, 2008. 464 p. (Coleção Sociologia).

EUROPEAN COMMISSION; DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION. **H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Version 3.0**. Luxembourg: European Commission, Directorate-General for Research & Innovation. 2016. DOI: <http://dx.doi.org/10.25607/OBP-774>. Disponível em: <https://repository.oceanbestpractices.org/handle/11329/1259>. Acesso em: 27 maio 2021.

FAUSTO, Sibeles. Evolução do Acesso Aberto: breve histórico. **SciELO em Perspectiva**, 2013. Disponível em: <https://blog.scielo.org/blog/2013/10/21/evolucao-do-acesso-aberto-breve-historico/>. Acesso em: 30 ago. 2021.

FIOCRUZ. **Dados Abertos**. Fiocruz: Rio de Janeiro, [20--]. (Aula 5: Plano de Gestão de Dados). Disponível em: Curso Ciência Aberta - Série 3: Curso 2 - Aula 5: Planos de Gestão de Dados (PGD) (fiocruz.br). Acesso em: 27 maio 2021.

FOSTER. **Open Science Definition**. [20-?]. Disponível em: <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>. Acesso em: 10 out. 2020.

FOSTER. **Open science taxonomy**. 2018. Disponível em: [https://www.fosteropenscience.eu/themes/fosterstrap/images/taxonomies/os\\_taxonomy.png](https://www.fosteropenscience.eu/themes/fosterstrap/images/taxonomies/os_taxonomy.png). Acesso em: 26 out. 2020.

GALVINO, Claudio Cesar Temoteo; ROSA, Maria Nilza Barbosa; OLIVEIRA, Bernardina Maria Juvenal Freire. O movimento de Acesso Aberto e a Ciência Aberta: uma proposta de repositório de dados e memória na Universidade Federal de Alagoas. **Ciência da Informação em Revista**, Maceió, v. 7, n. 1, p. 34 – 45, 2020. DOI: 10.28998/cirev.2020v7n1c. Disponível em: <https://www.seer.ufal.br/index.php/cir/article/view/9521>. Acesso em: 06 out. 2020.

GO FAIR. **FAIR Principles**. [20--]. Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 27 dez. 2020.

HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (ed). Jim Gray on escience: a transformed scientific method. *In*: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **The Fourth Paradigm: Data-Intensive Scientific Discovery**. Redmond, Washington: Microsoft Research, 2016, 2009. p. xvii-xxxii. Disponível em: <https://digital.library.unt.edu/ark:/67531/metadc31516/>. Acesso em: 07 out. 2021.

HANSON, Brooks; SUGDEN, Andrew; ALBERTS, Bruce. Making Data Maximally Available. **Science**, v. 331, n. 6018, p. 649–649, 2011. DOI: 10.1126/science.1203354. Disponível em: <https://science-sciencemag-org.ez3.periodicos.capes.gov.br/content/331/6018/649>. Acesso em: 08 jun. 2021.

HAUSSMANN, Darclê Costa Silva; VOGT, Mara; HEIN, Nelson. Desigualdades de Gênero no Mercado de Trabalho Brasileiro: uma análise na área de Ciências Sociais Aplicadas. **Enfoque: reflexão contábil**, Universidade Estadual de Maringá, v. 37, n. 3, p. 129-148, 2018. Disponível em: <https://www.redalyc.org/journal/3071/307160598009/>. Acesso em: 22 out. 2021.

HIGGINS, Sarah. The DCC Curation Lifecycle Model. **International Journal of Digital Curation**, v. 3, n. 1, p. 134–140, 2008. DOI: 10.2218/ijdc.v3i1.48. Disponível em: <http://www.ijdc.net/article/view/69>. Acesso em: 27 maio 2021.

KIM, Youngseek. Institutional and Individual Influences on Scientists' Data Sharing Behaviors. 2013. 304 f. Tese (Doutorado em Filosofia em Ciência e Tecnologia da Informação) - Universidade de Syracuse, Syracuse, 2013. Disponível em: [https://surface.syr.edu/it\\_etd/85](https://surface.syr.edu/it_etd/85). Acesso em: 16 ago. 2021.

KURATA, Keiko; MATSUBAYASHI, Mamiko; MINE, Shinji. Identifying the Complex Position of Research Data and Data Sharing Among Researchers in Natural Science.  **Sage**

**Open**, p. 1-12, jul.set. 2017. Disponível em: <https://doi.org/10.1177/2158244017717301>. Acesso em: 17 out. 2020.

IMPERIAL COLLEGE LONDON. **What is research data?** c2020. Disponível em: <http://www.imperial.ac.uk/research-and-innovation/support-for-staff/scholarly-communication/research-data-management/introduction-to-research-data-management/what-is-research-data/>. Acesso em: 22 dez. 2020.

LEITE, Fernando César Lima. *Gestão do conhecimento científico no contexto acadêmico: proposta de um modelo conceitual*. 2006. 240 f. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2006. Disponível em: <https://repositorio.unb.br/handle/10482/3975>. Acesso em: 24 maio 2021.

LIMA, Juliana Soares. **Gestão de Dados de Pesquisa no contexto da Ciência Aberta: percepção dos pesquisadores da Universidade Federal do Ceará**. 2020. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal do Ceará, Fortaleza, 2020. Disponível em: <http://www.repositorio.ufc.br/handle/riufc/51572>. Acesso em: 8 mar. 2021.

LU, Yi-Ching; KE, Hao-Ren. A Study on Scholars' Perceptions and Practices of Research Data Management. *Journal of library and Information Studies*, v. 18, n. 2, 2020. DOI: 10.6182/jlis.202012\_18(2).103. Disponível em: <https://scholar.lib.ntnu.edu.tw/en/publications/a-study-on-scholars-perceptions-and-practices-of-research-data-ma>. Acesso em: 15 mar. 2021.

MACIEL, André Filipe da Costa. *Plano de gestão de dados para a produção de dados FAIR: O caso de uso FRAILSURVEY*. 2020. 122 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Engenharia e Faculdade de Letras, Universidade do Porto. 2020. Disponível em: <https://repositorio-aberto.up.pt/handle/10216/128564>. Acesso em: 30 maio. 2021.

MARÍN-ARRAIZA, Paloma; PUERTA-DÍAZ, Mirelys; GREGORIO-VIDOTTI, Silvana. *Gestión de datos de investigación y bibliotecas: preservando los nuevos bienes científicos*. **Hipertext.net**, n. 19, p. 13-31, nov. 2019. DOI:10.31009/hipertext.net.2019.i19. Disponível em: <https://raco.cat/index.php/Hipertext/article/view/360098/455750>. Acesso em: 07 out. 2021.

MEADOWS, Arthur Jack. **A comunicação científica**. 1999. Disponível em: <https://docero.com.br/doc/x11s1e>. Acesso em: 14 mar. 2021.

MEDEIROS, Claudia Bauzer. *Gestão de dados científicos: da coleta à preservação*. **SciELO em Perspectiva**, 2018. Disponível em: <https://blog.scielo.org/blog/2018/06/22/gestao-de-dados-cientificos-da-coleta-a-preservacao/>. Acesso em: 23 dez. 2020.

MIKSA, Tomasz *et al.* Dez princípios para planos de gerenciamento de dados acionáveis por máquina. **PLoS Computational Biology**, v. 15, n. 3, 2019. Disponível em: <https://doi.org/10.1371/journal.pcbi.1006750>. Acesso em: 07 out. 2021.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR (Portugal). **Ciência Aberta**: sobre Ciência Aberta. Lisboa: MCTES, c2016. Disponível em: <http://www.ciencia-aberta.gov.pt/>

aberta.pt/sobre-ciencia-aberta. Acesso em 09 out. 2020. Disponível em: [https://70f8fba8-2c4e-4a79-ad0c-4f2ab4770372.filesusr.com/ugd/a8bd7c\\_503bd4fcaa3d45f79a1a9702acad94f7.pdf](https://70f8fba8-2c4e-4a79-ad0c-4f2ab4770372.filesusr.com/ugd/a8bd7c_503bd4fcaa3d45f79a1a9702acad94f7.pdf). Acesso em: 9 out. 2020.

MONTEIRO, Gabriela; LUCAS, Elaine R. de Oliveira; LECARDELLI, Jane. Financiadores da pesquisa brasileira e os dados científicos abertos. *In: ENCONTRO IBÉRICO EDICIC*, 9., 2019, Barcelona. **Anais** [...]. Barcelona, 2019. p. Disponível em: <http://eprints.rclis.org/38555/>. Acesso em: 28 jun. 2021.

MOREIRA, Josilene Aires; MATTOS, Giorgia de Oliveira; REIS, Luana Silva. Um panorama da presença feminina em Ciência da Computação. *In: ENCONTRO INTERNACIONAL DA REDE FEMINISTA NORTE E NORDESTE DE ESTUDOS E PESQUISA SOBRE A MULHER E RELAÇÕES DE GÊNERO*, 18., 2014, Recife. **Anais** [...]. Recife: UFRP, 2014, p. 3527- 3542. Disponível em: <http://www.ufpb.br/evento/index.php/18redor/18redor/paper/viewFile/1935/853>. Acesso em: 22 out. 2021.

MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1138>. Acesso em: 14 out. 2020.

NATIONAL SCIENCE BOARD. **Long-lived digital data collections: enabling research and education in the 21st century**. Washington: National Science Foundation, 2005. Disponível em: <https://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf>. Acesso em: 05 nov. 2020.

OLIVEIRA, Adriana Carla Silva. **Desvendando a autoralidade colaborativa na e-science sob a ótica dos direitos de propriedade intelectual**. 2016. 297 f. Tese (Doutorado em Ciência da Informação) - Universidade Federal da Paraíba, João Pessoa, 2016. Disponível em: [https://repositorio.ufpb.br/jspui/handle/tede/8849?locale=pt\\_BR](https://repositorio.ufpb.br/jspui/handle/tede/8849?locale=pt_BR). Acesso em: 26 maio. 2021.

OLIVEIRA, Adriana Carla Silva. Ciência aberta e quarto paradigma científico: a multidimensionalidade da ciência contemporânea. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 20., 2019, Florianópolis **Anais** [...]. Florianópolis, 2019. p. 1-21. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/994>. Acesso em: 26 maio. 2021.

OLIVEIRA, Adriana Carla Silva; SILVA, Edilene Maria. Ciência aberta: dimensões para um novo fazer científico. **Informação & Informação**, Londrina, v. 21, n. 2, p. 5–39, maio/ago. 2016. DOI: 10.5433/1981-8920.2016v21n2p5. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27666>. Acesso em: 26 out. 2020.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding**. Paris: OECD, 2007. Disponível em: <http://www.oecd.org/dataoecd/9/61/38500813.pdf>. Acesso em: 26 nov. 2017.

PINHEIRO, Lena Vania Ribeiro. Do acesso livre à ciência aberta: conceitos e implicações na comunicação científica. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 8, n. 2, p. 153-165, 2014. DOI: 10.3395/reciis.v8i2.629. Disponível em: <https://homologacao-reciis.icict.fiocruz.br/index.php/reciis/article/view/629>. Acesso em: 14 out. 2020.

RIBEIRO, Nivaldo Calixto; SILVEIRA, Lúcia da; SANTOS, Sarah Rúbia de Oliveira. **Taxonomia da Ciência Aberta**. [Traduzido e adaptado de] Nancy Pontika e Petr Knoth. 2020. Título original: Open Science Taxonomy. Disponível em: <http://repositorio.ufla.br/jspui/handle/1/39498>. Acesso em: 16 out. 2021.

ROA-MARTÍNEZ, Sandra M.; VIDOTTI, Silvana A. B.; SANTANA, Ricardo C. Estructura propuesta del artículo de datos como publicación científica. **Revista Española de Documentación Científica**, [S. l.], v. 40, n. 1, p. e167–e167, 2017. DOI: 10.3989/redc.2017.1.1375. Disponível em: <http://redc.revistas.csic.es/index.php/redc/article/view/971>. Acesso em: 03 jun. 2021.

SALES, Luana Farias; SAYÃO, Luís Fernando. A ciência invisível: revelando os dados da cauda longa da pesquisa. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais [...]**. Londrina: UEL, 2018, p. 4180- 4199. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/103678>. Acesso em: 08 set. 2021.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116–142, 2016. DOI: 10.5433/1981-8920.2016v21n2p116. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27940/20124>. Acesso em: 18 dez. 2020.

SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, v. 21, n. 2, p. 90-115, 2016. DOI: 10.5433/1981-8920.2016v21n2p90. Disponível em: <https://brapci.inf.br/index.php/res/download/46190>. Acesso em: 16 nov. 2020.

SAYÃO, Luís Fernando; SALES, Luana Farias. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **RECIIS – Revista Eletrônica de Comunicação Informação e Inovação em Saúde**, Rio de Janeiro, v. 8, n. 2, p. 76-92, jun. 2014. Disponível em: <https://www.recis.icict.fiocruz.br/index.php/reciis/article/view/611>. Acesso em: 20 out. 2020.

SAYÃO, Luís Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://www.aben.com.br/Arquivos/420/420.pdf>. Acesso em: 5 dez. 2019.

SCIENCE EUROPE. Practical Guide to the International Alignment of Research Data Management. California: Science Europe, 2018. Disponível em: Practical Guide to the International Alignment of Research Data Management - Extended Edition - Science Europe. Acesso em: 28 dez. 2020.

SILVA, Fabiano Couto Corrêa ; SILVEIRA, Lúcia. O ecossistema da Ciência Aberta. **Transinformação**, [S. l.], v. 31, 2019a. DOI: 10.1590/2318-0889201931e190001. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_abstract&pid=S0103-37862019000100302&lng=en&nrm=iso&tlng=pt](http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0103-37862019000100302&lng=en&nrm=iso&tlng=pt). Acesso em: 9 out. 2020.

SILVA, Fabiano Couto Corrêa; SILVEIRA, Lúcia. O que é e qual a importância da Ciência Aberta? **SciELO em Perspectiva: Humanas**, 2019b. Disponível em: <https://humanas.blog.scielo.org/blog/2019/12/02/o-que-e-e-qual-a-importancia-da-ciencia-aberta/>. Acesso em: 01 set. 2021.

SCHRAMM, Wilbur. **Notes on Case Studies of Instructional Media Projects**. [S. l.], 1971. Disponível em: <https://eric.ed.gov/?id=ED092145>. Acesso em: 14 mar. 2021.

SWAN, Alma. Macroscopic: Open Access and the Progress of Science. **American Scientist**, v. 3, p. 197-199, 2007. Disponível em: <http://www.jstor.org/stable/27858953>. Acesso em: 16 out. 2020

TENOPIR, Carol *et al.* Data Sharing by Scientists: practices and perceptions. **PLoS ONE**, [S. l.], v. 6, n. 6, p. e21101, 2011. DOI: 10.1371/journal.pone.0021101. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.00229003>. Acesso em: 10 nov. 2020.

THE STATE of open data. Londres: Springer Nature Wiley, 2017. Disponível em: [https://digitalscience.figshare.com/articles/report/The\\_State\\_of\\_Open\\_Data\\_Report\\_2017/5481187](https://digitalscience.figshare.com/articles/report/The_State_of_Open_Data_Report_2017/5481187). Acesso em: 27 maio 2021.

UHLIR, Paul F. Information gulags, intellectual straightjackets, and memory holes: three principles to guide the preservation of scientific data. **Data Science Journal**, [S. l.], v. 9, p. ES1-ES5, 2010. DOI: 10.2481/dsj.Essay-001-Uhlir. Disponível em: [https://www.researchgate.net/publication/273224347\\_Information\\_Gulags\\_Intellectual\\_Straightjackets\\_and\\_Memory\\_Holes](https://www.researchgate.net/publication/273224347_Information_Gulags_Intellectual_Straightjackets_and_Memory_Holes). Acesso em: 18 ago. 2021.

UHLIR, Paul F.; SCHRÖDER, Peter. Open data for global Science. **Data Science Journal**, v. 6, n. 17, jun. 2007, Disponível em: <http://datascience.codata.org/articles/abstract/10.2481/dsj.6.OD36/>. Acesso em: 09 nov. 2020.

UK DATA SERVICE. Research data management. Reino Unido, c2020. Disponível em: <https://www.digital-science.com/resources/portfolio-reports/state-open-data-2017/>. Acesso em: 27 maio 2021.

UNIVERSIDADE DE SÃO PAULO. **Gestão de Dados Científicos - Pró Reitoria de Pesquisa**, [2019]. Disponível em: <http://prp.usp.br/gestao-de-dados-cientificos/>. Acesso em: 20 dez. 2020.

UNIVERSIDADE DE SALAMANCA. Biblioteca da Universidade de Salamanca. Las 5 rutas para llegar al acceso abierto: verde, dorada, bronce, híbrida y diamante. **Blog Universo Abierto**. Salamanca, 2019. Disponível em: <https://universoabierto.org/2019/03/13/las-5-rutas-para-llegar-al-acceso-abierto-verde-dorada-bronce-hibrida-y->

diamante/?fbclid=IwAR0SIXcWTEBjYgJYv41sQGKGSCJBG672P8HjPMlt9EC84MmGv-gz7QoSZSE. Acesso em: 07 out. 2021.

VAN DEN EYNDEN *et al.* Managing and sharing data. Reino Unido: UK Data Archive. 2011. Disponível em: <https://dam.ukdataservice.ac.uk/media/622417/managingsharing.pdf>. Acesso em: 15 mar. 2021.

VANZ, Samile Andrea de Souza *et al.* **Acesso aberto a dados de pesquisa no Brasil: práticas e percepções dos pesquisadores: relatório 2018.** Porto Alegre: UFRGS, 2018. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/185195/001082283.pdf?sequence=1&isAllowed=y>. Acesso em: 17 dez. 2020.

VEIGA, Viviane Santos de Oliveira. **Percepção dos pesquisadores portugueses e brasileiros da área de Neurociências quanto ao compartilhamento de artigos científicos e dados de pesquisa no acesso aberto verde: custos, benefícios e fatores contextuais.** 2017. 291 f. Tese (Doutorado em Informação, comunicação e saúde) – Fundação Oswaldo Cruz, Rio de Janeiro, 2017. Disponível em: <https://www.arca.fiocruz.br/handle/icict/26842>. Acesso em: 6 out. 2020.

VEIGA, Viviane *et al.* Plano de gestão de dados FAIR: uma proposta para a Fiocruz. **Liinc em Revista**, Rio de Janeiro, v.15, n.2, p. 275-286, nov. 2019. Disponível em: <http://revista.ibict.br/liinc/article/view/5030>. Acesso em: 30 maio 2021.

VIEIRA, Adriane *et al.* Um estudo das relações entre gênero e âncoras de carreira. **Cadernos EBAPE.BR** [online]. v. 17, n. 3, p. 577-589, 2019. Disponível em: <https://www.scielo.br/j/cebape/a/WPJBgkq5ByTv7y9ZVsjBp5w/?lang=pt#>. Acesso em: 22 out. 2021

VIEIRA, Luciana Corrêa. Organização e disseminação da produção científica dos docentes do cesh/ufsm em um repositório digital. 2012. 130 f. Dissertação (Mestrado em Administração) – Universidade Federal de Santa Maria, Santa Maria, RS, 2013. Disponível em: <https://repositorio.ufsm.br/bitstream/handle/1/4622/VIEIRA%2C%20LUCIANA%20CORREA.pdf?sequence=1&isAllowed=y>. Acesso em: 12 jul. 2021.

WILKINSON, Mark, *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, v. 3, n. 160018, 2016. Disponível em: <https://doi.org/10.1038/sdata.2016.18>. Acesso em: 20 dez. 2020.

YIN, Robert K. **Estudo de Caso: planejamento e métodos.** 5. ed. Santana, Porto Alegre: Bookman Editora, 2015.

## APÊNDICE A – INTERPRETAÇÃO DO QUESTIONÁRIO

Temática 1: Questões referente ao perfil do pesquisador: com a finalidade de conhecer o perfil do discente com informações particulares como idade, sexo, nível de instrução, área de atuação, programa de pós-graduação, financiador de pesquisa e tempo de pesquisa.						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
1.1 Qual sua idade?	Identificar grupo etário com maior incidência	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) 20-25 2) 26-35 3) 36-45 4) 46-55 5) 56-65 6) Mais de 65	Associação de grupo etário com o compartilhamento de dados de pesquisa
1.2 Qual é o seu gênero?	Identificar maior incidência de gênero	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) Feminino 2) Masculino 3) Não informar	Associação de gênero em relação ao compartilhamento de dados de pesquisa
1.3 Qual seu nível de escolaridade?	Identificar o nível de escolaridade	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) Pós-doutorado 2) Doutorado	Associação de nível de instrução em relação ao compartilhamento de dados de pesquisa

<b>Temática 1: Questões referente ao perfil do pesquisador:</b> com a finalidade de conhecer o perfil do discente com informações particulares como idade, sexo, nível de instrução, área de atuação, programa de pós-graduação, financiador de pesquisa e tempo de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
1.4 Pertence a qual programa de pós-graduação?	Identificar a qual programa pertence	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	<b>Google Formulário:</b> escrevendo o nome do programa por extenso	Associação de programas de pós-graduação em relação ao compartilhamento de dados de pesquisa
1.5 Qual sua principal área de atuação?	Identificar a área do pesquisador	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	<b>Google Formulário:</b> escrevendo a área de atuação por extenso	Associação de áreas de atuação com o compartilhamento de dados de pesquisa
1.6 Qual das alternativas melhor descreve o principal financiador da sua pesquisa?	Identificar financiador de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) Agência nacional 2) Agência internacional 3) Agência regional 4) Empresa privada 5) A própria instituição 6) Outro	Associação de financiador de pesquisa e identificar qual o financiador com maior ocorrência
1.7 Há quanto tempo está envolvido em pesquisas?	Identificar o tempo de envolvimento em pesquisas	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) <5 anos 2) 5-10 anos 3) 11-15 anos 4) 16-20 anos 5) >20 anos 6) Nunca estive envolvido em pesquisa	Maior tempo em pesquisa implica possível familiaridade no compartilhamento de dados

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.1 Que termos melhor descrevem os tipos de dados de pesquisa que você produz? (marque quantas alternativas julgar necessário)	Identificar os termos de dados de pesquisa mais produzidos	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) Dados automaticamente gerados por programas de computador 2) Dados coletados de sensores e equipamentos 3) Dados de simulação 4) Dados documentais 5) Dados experimentais 6) Dados observacionais 7) Entrevistas 8) Questionários 9) Outros	O item com maior frequência demonstrará o termo mais utilizado pelos discentes.
2.2 Qual o tipo dos dados de pesquisa que você produz? (marque quantas alternativas julgar necessário)	Identificar os tipos de dados de pesquisa mais produzidos	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) Algoritmos 2) Animações, modelos ou simulações 3) Áudios 4) Banco de dados 5) Imagens 6) Planilhas 7) Slides 8) Software 9) Textos 10) Vídeos 11) Websites 12) Outros	O item com maior frequência demonstrará o tipo de dado mais utilizado pelos discentes.

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.3 Qual o volume estimado dos dados de pesquisa que você produz em um ano?	Identificar o volume de dados produzido por docente	Programas de Pós-graduação do ICESA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) < 1 GB 2) 1 - 50 GB 3) 50 - 100 GB 4) 100 - 500 GB 5) 500 GB - 1 TB 6) 1 - 50 TB's 7) 50 - 100 TB's 8) >100 TB's 9) Não sei responder	O item com maior frequência demonstrará o volume de dados do pesquisador
2.4 Seus dados de pesquisa são produzidos preponderantemente	Identificar de que forma é produzido os dados de pesquisa do docente.	Programas de Pós-graduação do ICESA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções: 1) De forma individual 2) Em grupo de até 5 pessoas 3) Em grupo de 6 até 12 pessoas 4) Em grupo de mais de 12 pessoas	O item com maior frequência demonstrará de forma é produzido os dados de pesquisa em termos quantitativos de pessoas.
2.5 Quanto dos seus dados de pesquisa você atualmente armazena nos locais a seguir?	Identificar os locais que são armazenados os dados de pesquisa dos docentes	Programas de Pós-graduação do ICESA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando entre: Todos os meus dados; Parte dos meus dados; Nenhum dos meus dados e Não sei responder para cada uma das opções abaixo: 1) Computador pessoal	O item com maior frequência demonstrará o local que os discentes mais armazenam seus dados.

Temática 2: Questões referentes as suas práticas sobre dados de pesquisa: conhecer as práticas sobre compartilhamento de dados de pesquisa.						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					2) Servidor localizado no laboratório do meu grupo de pesquisa 3) Servidor / data center da minha instituição 4) Arquivos em papel 5) Nuvem (ex. Google drive / Dropbox) 6) Mídia externa (ex. HD externo, USB/pendrive, CD/DVD)	
2.6 Você já compartilhou seus dados de pesquisa em acesso aberto? (marque quantas alternativas julgar necessário)	Identificar a experiência do docente no compartilhamento de dados de pesquisa	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando as opções que julgar necessário:  1) Sim, por solicitação do meu principal financiador 2) Sim, por solicitação da instituição principal em que atuo 3) Sim, para publicação em periódico ou em anais de evento 4) Sim, por iniciativa própria 5) Nunca compartilhei	O item com maior frequência demonstrará se o pesquisador teve experiência no compartilhamento de dados de pesquisa.

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.7 Quanto dos seus dados de pesquisa você atualmente compartilha em acesso aberto?	Identificar a proporção no compartilhamento de dados de pesquisa	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções:  1) Todos meus dados 2) Parte dos meus dados 3) Nenhum dos meus dados	O item com maior frequência demonstrará a experiência do pesquisador em relação ao compartilhamento de dados de pesquisa.
2.7a Se você respondeu que compartilha seus dados de pesquisa em acesso aberto, onde você compartilha?	identificar onde foi realizado o compartilhamento de dados de pesquisa do docente	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando entre: Todos os meus dados; Parte dos meus dados; nenhum dos meus dados e não sei responder para cada opção abaixo:  1) Repositório temático 2) Repositório de revista / editor 3) Repositório da minha instituição 4) Repositório nacional	O item com maior frequência demonstrará o local mais utilizado pelos docentes para compartilhar seus dados de pesquisa.
Caso você tenha respondido que compartilha todos seus dados de pesquisa, ignore a próxima questão						
2.7b Quais suas razões para não compartilhar ou compartilhar parcialmente os seus dados de pesquisa? (marque quantas alternativas julgar necessário)	Identificar os motivos pelo qual o pesquisador não compartilha os dados de pesquisa	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções:  1) Falta de financiamento 2) Falta de padronização dos dados 3) As pessoas não precisam deles	O item com maior frequência demonstrará o motivo pelo qual o pesquisador não compartilha seus dados de pesquisa

Temática 2: Questões referentes as suas práticas sobre dados de pesquisa: conhecer as práticas sobre compartilhamento de dados de pesquisa.						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					4) Não tenho tempo para compartilhá-los 5) Não há infraestrutura para compartilhá-los 6) Eles não podem ser compartilhados 7) Não me foi exigido 8) Não tenho permissão para torná-los públicos 9) Perderia o controle dos dados 10) Preciso publicar os resultados da pesquisa antes de compartilhá-los 11) Não sei como compartilhar 12) Os dados não estão totalmente documentados 13) Os dados estão em um formato que não permite esse compartilhamento 14) Outro:	
2.8 As seguintes afirmativas se referem ao compartilhamento e ao acesso a dados de pesquisa. Assinale conforme sua experiência.	Identificar condições de compartilhamento e acesso a dados de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando entre: sempre, as vezes e nunca para cada opção abaixo:	O item com maior frequência entre a escala sempre e nunca, resultará num panorama sobre compartilhamento e acesso a dados de pesquisa.

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What?</b> <b>(O que?)</b>	<b>Why?</b> <b>(Por que?)</b>	<b>Where?</b> <b>(Onde?)</b>	<b>When?</b> <b>(Quando?)</b>	<b>Who?</b> <b>(Quem?)</b>	<b>How?</b> <b>(Como?)</b>	<b>Interpretação</b>
					<ol style="list-style-type: none"> <li>1) Outros precisam da minha permissão para acessar meus dados de pesquisa</li> <li>2) Outros podem acessar meus dados de pesquisa livremente</li> <li>3) Falta de acesso aos dados gerados por outros pesquisadores ou instituições tem dificultado a realização das minhas pesquisas</li> <li>4) Minha instituição disponibiliza pessoal especializado para preparar os dados de pesquisa para serem compartilhados</li> <li>5) Minha instituição oferece as orientações específicas para o compartilhamento de dados de pesquisa</li> <li>6) Minha instituição disponibiliza um repositório para compartilhamento de dados de pesquisa</li> <li>7) Na minha área, os pesquisadores acessam livremente repositórios de dados de pesquisa</li> </ol>	

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.9 Quão satisfeito você está com a organização dos seus dados de pesquisa nas seguintes situações?	Apresentar o grau de satisfação quanto a organização dos dados de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando entre: satisfeito, parcialmente satisfeito, insatisfeito e não sei responder para cada opção abaixo:  1) Coleta de dados 2) Descrição dos dados 3) Análise dos dados 4) Armazenamento dos dados durante a execução do projeto 5) Preparação dos relatórios de pesquisa 6) Recuperação dos meus dados 7) Preparação dos relatórios de pesquisa 8) Preservação dos dados após a conclusão do projeto	O item com maior frequência entre a satisfeito e não sei responder resultará num panorama sobre organização dos dados de pesquisa.
2.10 Qual sua familiaridade com a gestão de dados de pesquisa?	Identificar a familiaridade dos docentes com a gestão de dados de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções:  1) Conheço bastante 2) Conheço pouco 3) Desconheço	O item com maior frequência demonstrará a familiaridade em gestão de dados de pesquisa

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.11 Você já elaborou um plano de gestão de dados? (marque quantas alternativas julgar necessário)	Identificar quantos já utilizaram esse instrumento	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando as opções que julgar necessárias:  1) Sim, por solicitação do meu principal financiador 2) Sim, por solicitação da instituição principal que eu atuo 3) Sim, para publicação em periódico 4) Sim, por outras razões 5) Nunca elaborei	O item com maior frequência demonstrará a familiaridade com plano de gestão de dados de pesquisa
2.12 Você utiliza algum repositório para acessar dados de pesquisa de outros pesquisadores?	Identificar se os pesquisadores utilizam repositórios para acessar dados de pesquisas	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções abaixo:  1) Não 2) Sim	O item com maior frequência demonstrará se o acesso a dados de pesquisa é por meio de repositórios
2.12a Se você respondeu sim na questão anterior, por favor, especifique aqui o repositório que você utiliza:	Mapear os repositórios mais utilizados pelos docentes	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , escrevendo por extenso o repositório que utilizam	Apresentar um panorama de repositórios utilizados pelos docentes do ITEC e ICSA

<b>Temática 2: Questões referentes as suas práticas sobre dados de pesquisa:</b> conhecer as práticas sobre compartilhamento de dados de pesquisa.						
<b>What? (O que?)</b>	<b>Why? (Por que?)</b>	<b>Where? (Onde?)</b>	<b>When? (Quando?)</b>	<b>Who? (Quem?)</b>	<b>How? (Como?)</b>	<b>Interpretação</b>
2.13 Você já utilizou em suas pesquisas dados abertos compartilhados por outros grupos de pesquisa?	Identificar se os docentes utilizam dados de pesquisa de outros grupos	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções abaixo: 1) Sim 2) Não 3) Não sei responder	O item com maior frequência demonstrará se os docentes utilizam dados compartilhados por outros grupos de pesquisa

**Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados:** conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa

What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
3.1 Especifique a sua concordância com cada uma das afirmativas a seguir:	Identificar a aceitação dos docentes em relação ao compartilhamento de dados de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	<p>Via <b>Google Formulário</b>, selecionando entre: concordo, concordo parcialmente, discordo e não sei responder, para as opções abaixo:</p> <ol style="list-style-type: none"> <li>1) Utilizaria dados de outros pesquisadores se estivessem acessíveis.</li> <li>2) Compartilharia pelo menos parte dos meus dados de pesquisa em um repositório de acesso aberto sem restrições</li> <li>3) Compartilharia todos os meus dados de pesquisa em um repositório de acesso aberto sem restrições</li> <li>4) Estaria mais inclinado a compartilhar meus dados se eu pudesse estabelecer condições de acesso e uso</li> <li>5) Considero apropriado criar novos conjuntos de dados a partir de dados compartilhados</li> </ol>	Frequência entre concordo e não sei responder em relação ao compartilhamento de dados de pesquisa.

**Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados:** conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa

What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					<p>6) Estou disposto a ajudar outros pesquisadores ao compartilhar dados</p> <p>7) O compartilhamento de dados pode ampliar meu reconhecimento acadêmico (por exemplo: divulgação do meu trabalho, citações recebidas).</p> <p>8) Poderia perder oportunidades de publicação se compartilhasse meus dados.</p> <p>9) Minhas ideias de pesquisa poderiam ser copiadas se compartilhasse meus dados.</p> <p>10) Meus dados poderiam ser mal utilizados ou mal interpretados por outros pesquisadores se eu os compartilhasse</p> <p>11) Estou disposto a dedicar parte do meu tempo e esforço para preparar os dados para serem compartilhados (por</p>	

Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados: conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					exemplo: organizar, anotar e anonimizar os dados).	
3.2 Para outros pesquisadores usarem meus dados de pesquisa, espero como retorno:	Identificar o que os docentes esperam de quem compartilham seus dados de pesquisa	Programas de Pós-graduação do ICSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	<p>Via <b>Google Formulário</b>, selecionando entre: Sempre, Às vezes, Nunca e Não sei responder, para as opções abaixo:</p> <ol style="list-style-type: none"> <li>1) Um acordo mútuo de reciprocidade no compartilhamento de dados</li> <li>2) Co-autoria em publicações resultantes dos dados</li> <li>3) Agradecimento em publicações resultantes dos dados</li> <li>4) Citação em publicações resultantes dos dados</li> <li>5) A oportunidade de colaborar em um projeto que use os dados.</li> <li>6) A solicitação de uma aprovação formal por parte do pesquisador antes da publicação do artigo</li> </ol>	Frequência entre sempre, às vezes, nunca e não sei responder, em relação a quem compartilhamento seus dados de pesquisa.

Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados: conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					7) Reembolso de parte dos custos da pesquisa 8) Ter acesso a uma cópia do artigo publicado 9) Ter acesso a todos os produtos resultantes da pesquisa, incluindo artigos, material educacional, apresentações etc.	
3.3 Onde você considera que seus dados de pesquisa devem ser preservados?	Identificar a percepção dos pesquisadores sobre o armazenamento dos dados de pesquisa	Programas de Pós-graduação do ICISA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando entre: Parte dos meus dados; Todos os meus dados; Nenhum dos meus dados e Não sei responder, para as opções abaixo:  1) Em servidor / datacenter da minha instituição 2) Em servidor localizado no laboratório do meu grupo de pesquisa 3) Em meu computador pessoal 4) Em papel, armazenado no meu escritório / laboratório	Frequência entre Todos os meus dados; Nenhum dos meus dados e Não sei responder, em relação a preservação de seus dados de pesquisa.

Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados: conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					5) Em um repositório temático 6) Em um repositório de revista 7) Em um repositório da minha instituição 8) Em outro tipo de repositório de dados 9) Na nuvem 10) Em mídia externa (ex. HD externo, USB/pendrive, CD/DVD) 11) Não devem ser preservados 12) Outra opção	
3.4 Quem você acha que deve oferecer serviços de apoio a gestão dos dados de pesquisa?	Identificar na percepção dos docentes sobre apoio na gestão de dados de pesquisa	Programas de Pós-graduação do ICOSA e o ITEC	01/07/2021 a 07/07/2021	Docentes	Via <b>Google Formulário</b> , selecionando uma das opções abaixo:  1) Agência de financiamento 2) Agência não governamental 3) Associação científica 4) Ente governamental (ex. instituição subordinada ao MCTIC) 5) Instituição de atuação do pesquisador 6) Não acho necessário apoio	O item com maior frequência representará, na opinião dos docentes, quem deve oferecer serviços em apoio a gestão dos dados de pesquisa

Temática 3 - Questões referentes às suas percepções em relação ao compartilhamento e uso de dados: conhecer as percepções sobre o compartilhamento e uso de dados de pesquisa						
What? (O que?)	Why? (Por que?)	Where? (Onde?)	When? (Quando?)	Who? (Quem?)	How? (Como?)	Interpretação
					7) Não sei responder 8) Outro:	

## APÊNDICE B – ROTEIRO DE ENTREVISTA

### ABERTURA

---

Prezado(a),

Agradeço a disponibilidade do entrevistado em fornecer as informações.

Saliento que os resultados da pesquisa estarão à disposição dele e, se tiver interesse, deverá entrar em contato com o pesquisador.

Este instrumento visa coletar dados para a pesquisa "Compartilhamento de dados científicos: percepção e práticas dos docentes dos programas de pós-graduação do ITEC e ICSA".

O objetivo é investigar as percepções e práticas dos docentes permanentes de programas de pós-graduação do ITEC e ICSA quanto ao compartilhamento de dados pesquisa.

Sua participação é de grande importância para o desenvolvimento desta pesquisa.

O roteiro de entrevista contém 2 blocos: (1) Percepção sobre o compartilhamento de dados de pesquisa; e, (2) Práticas de compartilhamento de dados de pesquisa.

### CARACTERIZAÇÃO DO RESPONDENTE

---

Nome: \_\_\_\_\_

Categoria: ( ) Docente do ITEC ( ) Docente do ICSA

### PERCEPÇÃO SOBRE O COMPARTILHAMENTO DE DADOS DE PESQUISA

---

- 1 Descreva, em poucas palavras, no seu entendimento o que significa acesso aberto aos dados de pesquisa?
- 2 Considerando os dados de pesquisa que você produz e/ou obtém durante a realização das pesquisas, na sua percepção: quão úteis eles seriam para outros pesquisadores?
- 3 O compartilhamento de dados de pesquisa tem sido uma prática em algumas áreas. Como você percebe essa questão no programa de pós-graduação que você atua?

- 4 Considerando o programa de pós-graduação em que atua, quais motivos estimulariam o compartilhamento de dados de pesquisa?
- 5 Considerando o programa de pós-graduação em que atua, quais motivos inibiriam o compartilhamento de dados de pesquisa?
- 6 Como você reagiria caso a agência que financia seus projetos tornasse o compartilhamento de dados de suas pesquisas uma condição para o financiamento?
- 7 Como você reagiria se sua instituição tornasse obrigatório o compartilhamento de dados de suas pesquisas?
- 8 Como você reagiria caso editores dos periódicos nos quais você publica exigissem que o artigo submetido ou aceito para publicação fosse acompanhado dos dados da pesquisa que gerou?

### **PRÁTICAS DE COMPARTILHAMENTO DE DADOS DE PESQUISA**

---

- 9 Você já teve alguma experiência de compartilhamento de dados de pesquisa? Se sim, comente onde, como e em que condições? Se não, por quê? Caso não compartilhe, quais os motivos? Compartilharia apenas se fosse uma exigência?
- 10 Você já publicou em revistas científicas que exigem o envio e compartilhamento dos dados referentes ao manuscrito submetido? Comente
- 11 Você acredita que a capacitação e a sensibilização poderiam ajudar na adoção dessa prática de compartilhamento de dados de pesquisa?
- 12 Quando você finaliza sua pesquisa, qual é o destino dos dados que foram produzidos?

### **CONSIDERAÇÕES FINAIS**

---

- 13 Há algo mais que gostaria de acrescentar em relação a essa entrevista?

### **FINALIZAÇÃO E AGRADECIMENTO**

---

Agradeço a disponibilidade do entrevistado em fornecer as informações.