



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE HUMANIDADES**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**  
**MESTRADO EM CIÊNCIA DA INFORMAÇÃO**  
**JULIANA SOARES LIMA**

**GESTÃO DE DADOS DE PESQUISA NO CONTEXTO DA CIÊNCIA ABERTA:  
PERCEPÇÃO DOS PESQUISADORES DA UNIVERSIDADE FEDERAL DO  
CEARÁ**

**FORTALEZA**

**2020**

JULIANA SOARES LIMA

GESTÃO DE DADOS DE PESQUISA NO CONTEXTO DA CIÊNCIA ABERTA:  
PERCEPÇÃO DOS PESQUISADORES DA UNIVERSIDADE FEDERAL DO CEARÁ

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará, como requisito para a obtenção do título de mestre em Ciência da Informação.

**Área de concentração:** Representação e Mediação da Informação e do Conhecimento.

**Linha de Pesquisa:** Mediação e Gestão da Informação e do Conhecimento.

**Orientadora:** Profa. Dra. Maria Giovanna Guedes Farias.

FORTALEZA - CE

2020

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

L698g Lima, Juliana Soares.

Gestão de Dados de Pesquisa no contexto da Ciência Aberta : percepção dos pesquisadores da Universidade Federal do Ceará / Juliana Soares Lima. – 2020.  
347 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Ciência da Informação, Fortaleza, 2020.

Orientação: Profa. Dra. Maria Giovanna Guedes Farias.

1. Gestão de Dados de Pesquisa. 2. Dados de Pesquisa. 3. Ciência Aberta. 4. Compartilhamento de dados. I. Título.

CDD 020

---

JULIANA SOARES LIMA

GESTÃO DE DADOS DE PESQUISA NO CONTEXTO DA CIÊNCIA ABERTA:  
PERCEPÇÃO DOS PESQUISADORES DA UNIVERSIDADE FEDERAL DO CEARÁ

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará, como requisito para a obtenção do título de mestre em Ciência da Informação.

**Área de concentração:** Representação e Mediação da Informação e do Conhecimento.

**Linha de Pesquisa:** Mediação e Gestão da Informação e do Conhecimento.

Aprovada em: \_\_\_/\_\_\_/\_\_\_\_\_.

BANCA EXAMINADORA

---

Profa. Dra. Maria Giovanna Guedes Farias (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Luís Fernando Sayão (Membro Externo)  
Comissão Nacional de Energia Nuclear (CNEN)

---

Profa. Dra. Andréa Soares Rocha da Silva (Membro Interno)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Jefferson Veras Nunes (Membro Interno Suplente)  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Luana Farias Sales (Membro Externo Suplente)  
Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Aos Deuses.

Aos meus amores, Ravi e Tarcísio.

Aos meus pais, Maria Celes e Edival.

Aos meus irmãos, Júlio César e Paulo César.

À minha amada sogra, Laudi Mulato.

Aos meus avós, Cesarina, Antonio e Antonio Rufino, *in memoriam*.

Ao meu amado Thor, *in memoriam*.

À minha querida Vanderlândia Lima, um ser de luz que partiu desse planeta, *in memoriam*.

## **AGRADECIMENTOS**

A Deus, por me mostrar que tudo vem no tempo dele e não no meu, na hora certa e devida.

À Professora Maria Giovanna Guedes Farias, minha orientadora, por ser um anjo de luz em minha vida. Nenhuma palavra é suficiente para agradecer tudo o que ela representa.

Ao meu marido, José Tarcísio dos Santos Júnior, por ser um pai maravilhoso, por tanto amor e dedicação, por me apoiar e incentivar.

À minha família inteira: meus pais, irmãos, tios, tias, primos, primas e avó.

Ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará (PPGCI-UFC) e a todos os professores pelo aprendizado e apoio, especialmente, à Professora Lídia Eugênia Cavalcante (anteriormente Coordenadora do PPGCI-UFC), Professora Virgínia Bentes Pinto (Vice-Coordenadora anterior), Professor Luiz Tadeu Feitosa (atual Coordenador do PPGCI-UFC) e Professor Jefferson Veras Nunes (atualmente, Vice-Coordenador do PPGCI-UFC). Gratidão imensa à Veruska Maciel, Secretária do PPGCI-UFC, por ser um doce e amor de pessoa, uma profissional de excelência e ser humano exemplar.

Aos docentes participantes de minha Banca Examinadora, Andréa Soares, a quem tenho profunda admiração e respeito, Luís Fernando Sayão, ilustre pesquisador da Comissão Nacional de Energia Nuclear (CNEN) e referência em gestão de dados de pesquisa, Luana Sales, referência em gestão de dados de pesquisa, Jefferson Veras, sempre trazendo excelentes discussões para a sala de aula sobre a Ciência da Informação: agradeço por aceitarem o convite, dedicarem seu tempo para dar suas contribuições, tecer comentários, sugerir melhorias para essa dissertação.

Aos pesquisadores da UFC (docentes e discentes), que tanto contribuíram para esta pesquisa, a minha eterna gratidão: sem vocês, absolutamente nada seria possível. Agradeço pela disponibilidade e acolhida de cada um. Aos participantes do minicurso online, por terem participado e me ajudado tanto com seus comentários, opiniões, críticas e sugestões.

À querida Marlene Lopes Cidrack, pesquisadora da UFC, membro da Academia Cearense de Odontologia, por ser a simpatia em pessoa, bondosa e humilde, que tanto me ajudou a preparar a documentação dessa pesquisa para submeter ao Comitê de Ética da UFC via Plataforma Brasil.

Aos meus diletísimos colegas da turma do mestrado, admiro cada um: Ezequiel Sousa (Belchior Escobar), Thiago Pinheiro, Rafaela Pereira, Raquel Nascimento, Mayara Vasconcelos, Silvana Araújo, Nayeli Feitosa, Isabel Wanderley e Robson da Silva.

Aos bibliotecários e parceiros de trabalho da UFC, os quais eu tenho muito orgulho de conhecer: Francisco Edvander Pires Santos (amigo de todas as horas, minha emoção e meu podcaster preferido), Irlana Mendes Araújo (amiga querida e minha razão, poucas pessoas no mundo são sensatas como ela), Joana Páscoa (amiga super querida, excelente profissional e que tem minha total admiração), Michele Marinho (amiga linda, doce e com uma fé inabalável), Kalline Yasmin Soares Feitosa (amiga linda em todos os sentidos, por dentro e por fora, é só amor), Jairo Viana (amigo querido e maravilhoso que a vida me trouxe, agora está na UFPE), Ana Rafaela Sales de Araújo (Rainha da Catalogação), Socorro Lourenço (exemplo de resiliência e bondade, Bibliotecária aposentada da Biblioteca de Ciências Humanas – BCH).

Às bibliotecárias companheiras de estágio na Procuradoria Geral do Município (PGM) e amigas de longa data: Viviane de Holanda Cabral e Sandra Simões (Bibliotecária da PGM).

À Manuella Nobre Pitombeira e Andréa Acioly Maia Firmo, amigas queridas do coração, servidoras do Instituto Federal do Ceará (IFCE).

Às maravilhosas Professora Cristina Façanha, Sara Castro e Nágila Ruthe, ao finíssimo Claudenio Delfino, membros seletos do grupo de estudos e pesquisas Diálogos com a Abordagem Pikler, da Faculdade de Educação da UFC.

À Professora Diana Fortier, coordenadora do projeto Laboratório de Tradução, Revisão e Edição de Textos Acadêmicos (LETRARE) da UFC, à Professora Vlória Borges, Professora Márcia Gradvohl, Tadeu Azevedo e a todos os membros do LETRARE.

Às Professoras do curso de Letras Espanhol da UFC, Nadja Prata e Beatriz Furtado, pela simpatia, gentileza e apoio de sempre.

Ao curso de Psicologia: ao meu querido amigo Jésus Dias e sua esposa Juliana Borges, a maravilhosa Raquel Libório (Revista de Psicologia), Professora Verônica Ximenes, Professora Zulmira Bomfim, Professora Laéria Fontenele, Professor Caubi Tupinambá, Professora Karla Patrícia Holanda e ao estimado Allan Ratts, aluno da Pós-Graduação em Psicologia.

Aos cursos de Pós-Graduação em Economia, Letras, Medicina, Engenharia Civil, Estatística e Psicologia, por terem colaborado substancialmente com o estudo. Aos discentes Amisterdam Silva, Misael Torres e Roxanne Santos.

Ao Professor Gustavo Cabral e ao curso de Pós-Graduação em Direito da UFC, por apoiar esta pesquisa e por terem sido os primeiros a levarem a discussão do tema desta pesquisa para a sala de aula. Gratidão pelo convite!

Ao Jonathan Peter Tennant (*In Memoriam*), pela generosidade e por me dar a oportunidade de ser inserida em um genuíno projeto de colaboração que tanto contribuiu para o avanço da Ciência Aberta como o *Open Science MOOC*.

Aos docentes, discentes, servidores técnico-administrativos da UFC e comunidade externa por valorizarem o trabalho desenvolvido pela BCH, por participarem dos cursos e treinamentos ofertados (presenciais e EaD), por prestigiarem a Semana de Metodologia & Produção Científica, projeto liderado pela BCH e realizado em parceria com as Bibliotecas do Campus Benfica da UFC.

À vida, por cada dificuldade superada, pelas vitórias e graças alcançadas, pelas reviravoltas do destino que me surpreendem sempre.

Aos leitores deste trabalho, *lecturis salutem!*



*“Research is formalized curiosity. It is poking and prying with a purpose.”* (Zora Neale Hurston).

## RESUMO

Desde a antiguidade, com a introdução da aplicação do método científico para validar o conhecimento, sua produção e os resultados de pesquisas passaram a ser pautados na troca de ideias e sugestões entre os pares, no compartilhamento de informações que necessitavam passar pelo crivo dos membros da comunidade científica. A partir da evolução da ciência e dos aparatos tecnológicos que passaram a coletar maior quantidade de informações para as pesquisas científicas, os dados emergem como produto essencial para o avanço do conhecimento científico necessários para a validação dos resultados de qualquer estudo. Os dados de pesquisa se apresentam em várias formas e devem ser contextualizados dentro das disciplinas ou áreas as quais pertencem. Nesse sentido, esta pesquisa tem como objetivo investigar as práticas e necessidades informacionais dos pesquisadores (docentes, discentes e técnico-administrativos em Educação vinculados ao Mestrado ou ao Doutorado) dos cursos de Pós-Graduação da Universidade Federal do Ceará (UFC), concernentes ao gerenciamento de dados de pesquisa e à Ciência Aberta. Para tal, delinear-se os seguintes objetivos específicos: analisar a percepção dos pesquisadores sobre a gestão de dados de pesquisa e a Ciência Aberta; averiguar quais as práticas e as necessidades informacionais destes pesquisadores referentes a estas temáticas; propor um Programa de Gestão de Dados de Pesquisa (PGDP) para a UFC com o objetivo de sugerir uma Política de Gestão de Dados de Pesquisa; sugerir a criação de serviços de apoio e suporte ao pesquisador na UFC; desenvolver ações de educação e informação com vistas a testar um piloto de curso online como parte integrante do programa. Realizou-se um levantamento exaustivo a partir de buscas realizadas por meio do software Publish or Perish, Portal de Periódicos da CAPES, Wizdom.ai e Twitter. Como estratégia metodológica, utilizou-se a triangulação de métodos – Teoria Fundamentada em Dados e a Netnografia, além das técnicas de pesquisa análise documental e a observação participante. Para coletar os dados adotou-se o questionário e a entrevista, bem como o uso do diário de campo eletrônico e o caderno de laboratório eletrônico para as anotações, registros de notas de campo e na construção de memorandos. Os dados foram tratados por uma abordagem qualitativa com o uso do software Atlas.ti para a construção das categorias. Os resultados demonstram que em relação às práticas e estratégias de armazenamento dos pesquisadores, o computador pessoal e a nuvem são os mais utilizados para

manter seus arquivos e dados de pesquisas, embora a maioria tenha revelado não ter uma frequência de backup de seus arquivos por usar o serviço de sincronização automática da nuvem. Sobre as práticas de documentação da pesquisa com a elaboração de um Plano de Gestão de Dados (PGD), entre todos os respondentes do questionário apenas uma pessoa elaborou um PGD, enquanto no grupo de entrevistados nenhum chegou a usar o PGD para essa finalidade. Sobre o compartilhamento, os entrevistados afirmaram ter realizado algum tipo de compartilhamento, seja de informações ou dados de pesquisa, e, quando não compartilham, os motivos declarados foram: desconhecimento, por não saber como fazer ou por esbarrarem em questões éticas, legais e de integridade da pesquisa. Diante do exposto, conclui-se que o pesquisador tem um papel fundamental na Gestão dos Dados de Pesquisa, pois adotar essa postura representa garantia da qualidade e integridade da pesquisa, além de colaborar para as boas práticas na ciência. Ademais, a literatura mostra que o bibliotecário tem sido o profissional mais recomendado para auxiliar os pesquisadores nesse processo. Finalmente, esta pesquisa traz como contribuição a percepção dos pesquisadores sobre os dados de pesquisa e a Ciência Aberta, além da sugestão de uma proposta de Programa de Gestão de Dados de Pesquisa (PGDP) para a UFC que se concentra no desenvolvimento de políticas, diretrizes, ações de educação e informação, produtos, serviços e gestão dos dados de pesquisa na universidade.

**Palavras-chave:** Gestão de Dados de Pesquisa. Dados de Pesquisa. Ciência Aberta.

## **ABSTRACT**

Since antiquity, when the scientific method was introduced to validate knowledge, its production and research results began to be guided by the exchange of ideas and suggestions among peers, by the sharing of information which needed to pass through the scrutiny of the scientific community members. From the evolution of science and technological apparatuses that started to collect more information for scientific researches, the data emerge as an essential product for the advance of scientific knowledge necessary for the validation of the results of any study. The research data are presented in various forms and must be contextualized within the disciplines or areas to which they belong. In this sense, this research aims to investigate the practices and informational needs of researchers (teachers, students and technical-administrative in Education linked to the Master's or Doctorate) of the Postgraduate courses of the Federal University of Ceará (UFC), concerning the management of research data and Open Science. For this purpose, the following specific objectives were outlined: analyze researchers' perception of research data management and Open Science; investigate the practices and informational needs of these researchers regarding these themes; propose a Research Data Management Program (PGDP) for the UFC with the objective of suggesting a Research Data Management Policy; suggest the creation of support services for the researchers at the UFC; develop educational and information actions in order to test an online course pilot as part of the program. Exhaustive research was conducted using Publish or Perish software, CAPES Journal Portal, Wizdom.ai and Twitter. The methodological strategy used was the triangulation of methods - Grounded Theory and Netnography, in addition to the research techniques of documental analysis and participant observation. The data collection was based on the questionnaire and interview, as well as the use of the electronic field journal and the electronic laboratory notebook for the notes, field note records and in the construction of memos. The data were treated by a qualitative approach with the use of Atlas.ti software for the construction of the categories. The results demonstrate that in relation to researchers' storage practices and strategies, the personal computer and the cloud are the most widely used to maintain their research files and data, although most have revealed that they do not have a backup frequency of their files because they use the automatic synchronization service of the cloud. About survey documentation practices with the elaboration of a Data

Management Plan (PGD), among all survey respondents only one person elaborated a PGD, while in the group of interviewees none ever used the PGD for this purpose. Regarding the sharing, interviewees stated that they had shared some type of information or survey data, and when they did not share, the reasons stated were: lack of knowledge, for not knowing how to do it or for bumping into ethical, legal and integrity issues in the survey. Hence, the conclusion is that the researcher has a fundamental role in the Management of Research Data, since adopting this posture represents a guarantee of the quality and integrity of the research, in addition to collaborating for good practices in science. Furthermore, the literature shows that the librarian has been the most recommended professional to assist researchers in this process. Finally, the contribution of this investigation is based on researchers' perception of research data and Open Science, in addition to the suggestion of a proposal for a Research Data Management Program (PGDP) for the UFC, which focuses on the development of policies, guidelines, education and information actions, products, services and management of research data at the university.

**Keywords:** Research Data Management. Research Data. Open Science.

## RESUMEN

Desde la antigüedad, cuando se introdujo el método científico para validar los conocimientos, su producción y los resultados de las investigaciones comenzaron a guiarse por el intercambio de ideas y sugerencias entre pares, por el intercambio de información que debía pasar por el escrutinio de los miembros de la comunidad científica. A partir de la evolución de los aparatos científicos y tecnológicos que comenzaron a reunir más información para las investigaciones científicas, los datos surgen como un producto esencial para el avance del conocimiento científico necesario para la validación de los resultados de cualquier estudio. Los datos de las investigaciones se presentan de diversas formas y deben ser contextualizados dentro de las disciplinas o áreas a las que pertenecen. En este sentido, esta investigación tiene como objetivo investigar las prácticas y necesidades de información de los investigadores (profesores, estudiantes y técnicos-administrativos en Educación vinculados a la Maestría o Doctorado) de los cursos de Postgrado de la Universidad Federal de Ceará (UFC), en lo que respecta a la gestión de los datos de la investigación y la Ciencia Abierta. Para ello, se plantearon los siguientes objetivos específicos: analizar la percepción de los investigadores sobre la gestión de datos de investigación y Open Science; investigar las prácticas y necesidades de información de estos investigadores en relación con estos temas; proponer un Programa de Gestión de Datos de Investigación (PGDP) para la UFC con el objetivo de sugerir una Política de Gestión de Datos de Investigación; sugerir la creación de servicios de apoyo a los investigadores de la UFC; desarrollar acciones educativas e informativas para poner a prueba un curso piloto en línea como parte del programa. Se realizó una investigación exhaustiva utilizando el software Publish or Perish, el portal de revistas CAPES, Wizdom.ai y Twitter. La estrategia metodológica utilizada fue la triangulación de métodos - Teoría Fundamentada y Netnografía, además de las técnicas de investigación de análisis documental y observación de los participantes. La recopilación de datos se basó en el cuestionario y la entrevista, así como en el uso del diario electrónico de campo y el cuaderno electrónico de laboratorio para las notas, los registros de notas de campo y en la construcción de memorandos. Los datos se trataron con un enfoque cualitativo mediante la utilización del programa informático Atlas.ti para la construcción de las categorías. Los resultados demuestran que, en relación con las prácticas y estrategias de almacenamiento de los investigadores, la

computadora personal y la nube son las más utilizadas para mantener sus archivos y datos de investigación, aunque la mayoría ha revelado que no tienen una frecuencia de copia de seguridad de sus archivos porque utilizan el servicio de sincronización automática de la nube. En cuanto a las prácticas de documentación de las encuestas con la elaboración de un Plan de Gestión de Datos (PGD), entre todos los encuestados sólo una persona elaboró un PGD, mientras que en el grupo de entrevistados ninguno utilizó nunca el PGD para este fin. En cuanto al intercambio, los entrevistados declararon que habían compartido algún tipo de información o datos de la encuesta, y cuando no lo hicieron, las razones aducidas fueron: falta de conocimiento, por no saber cómo hacerlo o por tropezar con cuestiones éticas, legales y de integridad en la encuesta. De ahí que se concluya que el investigador tiene un papel fundamental en la gestión de los datos de la investigación, ya que adoptar esta postura representa una garantía de la calidad e integridad de la investigación, además de colaborar en las buenas prácticas de la ciencia. Además, la literatura muestra que el bibliotecario ha sido el profesional más recomendado para ayudar a los investigadores en este proceso. Por último, la contribución de esta investigación se basa en la percepción que tienen los investigadores de los datos de investigación y de la Ciencia Abierta, además de la sugerencia de una propuesta de Programa de Gestión de Datos de Investigación (PGDP) para la UFC, que se centra en el desarrollo de políticas, directrices, acciones de educación e información, productos, servicios y gestión de datos de investigación en la universidad.

**Palabras clave:** Gestión de datos de investigación. Datos de investigación. Ciencia Abierta.

## RÉSUMÉ

Depuis l'Antiquité, lorsque la méthode scientifique a été introduite pour valider les connaissances, sa production et ses résultats de recherche ont commencé à être guidés par l'échange d'idées et de suggestions entre pairs, par le partage d'informations qui devaient passer par l'examen minutieux des membres de la communauté scientifique. Avec l'évolution des appareils scientifiques et technologiques qui ont commencé à recueillir davantage d'informations pour les recherches scientifiques, les données apparaissent comme un produit essentiel pour le progrès des connaissances scientifiques nécessaires à la validation des résultats de toute étude. Les données de recherche sont présentées sous différentes formes et doivent être contextualisées dans les disciplines ou les domaines auxquels elles appartiennent. En ce sens, cette recherche vise à étudier les pratiques et les besoins informationnels des chercheurs (enseignants, étudiants et technico-administrateurs en éducation liés au Master ou au Doctorat) des cours de troisième cycle de l'Université fédérale du Ceará (UFC), concernant la gestion des données de recherche et la Science ouverte. À cette fin, les objectifs spécifiques suivants ont été définis : analyser la perception des chercheurs sur la gestion des données de recherche et la science ouverte ; étudier les pratiques et les besoins en information de ces chercheurs sur ces thèmes ; proposer un programme de gestion des données de recherche (PGDP) pour l'UFC dans le but de suggérer une politique de gestion des données de recherche ; suggérer la création de services de soutien pour les chercheurs de l'UFC ; développer des actions éducatives et d'information afin de tester un cours pilote en ligne dans le cadre du programme. Des recherches exhaustives ont été menées à l'aide des logiciels Publish ou Perish, du portail du journal CAPES, de Wizdom.ai et de Twitter. La stratégie méthodologique utilisée a été la triangulation des méthodes - Grounded Theory et Netnography, en plus des techniques de recherche de l'analyse documentaire et de l'observation des participants. La collecte de données a été basée sur le questionnaire et l'interview, ainsi que sur l'utilisation du journal de terrain électronique et du carnet de laboratoire électronique pour les notes, les enregistrements des notes de terrain et dans la construction des mémos. Les données ont été traitées par une approche qualitative avec l'utilisation du logiciel Atlas.ti pour la construction des catégories. Les résultats montrent qu'en ce qui concerne les pratiques et stratégies de stockage des chercheurs, l'ordinateur personnel et le nuage



sont les plus utilisés pour conserver leurs fichiers et données de recherche, bien que la plupart aient révélé qu'ils n'ont pas de fréquence de sauvegarde de leurs fichiers parce qu'ils utilisent le service de synchronisation automatique du nuage. En ce qui concerne les pratiques de documentation des enquêtes avec l'élaboration d'un plan de gestion des données (PGD), parmi toutes les personnes interrogées, une seule a élaboré un PGD, alors que dans le groupe des personnes interrogées, aucune n'a jamais utilisé le PGD à cette fin. En ce qui concerne le partage, les personnes interrogées ont déclaré qu'elles avaient partagé un certain type d'informations ou de données d'enquête, et lorsqu'elles ne l'ont pas fait, les raisons invoquées étaient les suivantes : manque de connaissances, ne pas savoir comment le faire ou se heurter à des problèmes éthiques, juridiques et d'intégrité dans l'enquête. La conclusion est donc que le chercheur a un rôle fondamental dans la gestion des données de recherche, puisque l'adoption de cette posture représente une garantie de la qualité et de l'intégrité de la recherche, en plus de la collaboration pour les bonnes pratiques scientifiques. En outre, la littérature montre que le bibliothécaire a été le professionnel le plus recommandé pour aider les chercheurs dans ce processus. Enfin, la contribution de cette enquête est basée sur la perception qu'ont les chercheurs des données de recherche et de la science ouverte, en plus de la suggestion d'une proposition de programme de gestion des données de recherche (PGDP) pour l'UFC, qui se concentre sur le développement de politiques, de directives, d'actions d'éducation et d'information, de produits, de services et de gestion des données de recherche à l'université.

**Mots-clés:** Gestion des données de recherche. Recherche de données. La science ouverte.

## LISTA DE FIGURAS

Figura 1 - Pirâmide invertida da delimitação do tema.....	20
Figura 2 - Publicações sobre Data Management ao longo dos anos.....	21
Figura 3 - Tópicos relacionados mais pesquisados.....	22
Figura 4 - Interface do software Publish or Perish – PoP.....	25
Figura 5 - Pilares da Ciência Aberta.....	40
Figura 6 - Princípios da Ciência Aberta.....	41
Figura 7 - Taxonomia da Ciência Aberta.....	45
Figura 8 - Concepções sobre o conceito Aberto.....	47
Figura 9 - Iceberg dos dados científicos.....	60
Figura 10 - Taxonomia da Gestão de Dados de Pesquisa.....	63
Figura 11 - Mapa de repositórios por países.....	65
Figura 12 - Ciclo de vida da Gestão de Dados.....	70
Figura 13 - Research Data Life Cycle.....	71
Figura 14 - Modelo do ciclo de vida dos dados e curadoria (DCC).....	73
Figura 15 - Ciclo de vida dos dados (DataOne).....	75
Figura 16 - Ciclo de vida dos dados em Ciência da Informação de Sant’Ana (CDV-CI) .....	77
Figura 17 - Fases da Gestão de Dados de Pesquisa aplicadas ao PGD.....	80
Figura 18 - FAIR Principles.....	88
Figura 19 - Diagrama de Venn sobre Competência em dados.....	101
Figura 20 - Triangulação metodológica aplicada à pesquisa.....	112
Figura 21 - Ciclo da Teoria Fundamentada.....	114
Figura 22 - Processo de codificação na Teoria Fundamentada.....	118
Figura 23 - Etapas do DI 4.0.....	138
Figura 24 - Nuvem de palavras das primeiras coletas de dados.....	150
Figura 25 – Relação dos códigos do questionário.....	151
Figura 26 - Primeiros códigos extraídos das entrevistas.....	153
Figura 27 - Relação das categorias e seus códigos.....	154
Figura 28 – Categoria principal e categorias (Temas).....	155
Figura 29 - Exemplos dos computadores de alto desempenho do Cenapad-UFC .	167
Figura 30 – Categorias e temas levantados para compor o módulo online do curso .....	229

Figura 31 - Interface do Google Classroom.....	232
Figura 32 - Exemplo de slides com QR Code.....	232
Figura 33 - Exemplo de videoaula.....	233
Figura 34 - Exemplo dos dados sobre estatísticas de acesso aos materiais elaborados .....	234
Figura 35 – Exemplo de interação entre os participantes no mural .....	238
Figura 36 - Jamboard (Lousa Virtual).....	242
Figura 37 - Matriz de Feedback .....	243
Figura 38 - Estrutura do Programa de Gestão de Dados de Pesquisa .....	248
Figura 39 - Estrutura da Proposta da Política de Dados.....	249
Figura 40 - Tipos de ações de educação e informação .....	251
Figura 41 - Constituintes da comunicação científica agregada à gestão de dados de pesquisa.....	256

## LISTA DE GRÁFICOS

Gráfico 1 - Categoria de respondentes do questionário .....	156
Gráfico 2 - Porcentagem de participação por Área do Conhecimento .....	160
Gráfico 3 – Porcentagem de participação dos cursos .....	161
Gráfico 4 - Tipos de dados gerados em pesquisas na UFC .....	162
Gráfico 5 - Estratégias de armazenamento dos dados de pesquisa .....	164
Gráfico 6 - Frequência de backup dos dados .....	165
Gráfico 7 - Experiências de perda dos dados de pesquisa .....	166
Gráfico 8 - Porcentagem de pesquisadores que conhecem o CENAPAD-UFC.....	168
Gráfico 9 - Elaboração de um Plano de Gestão de Dados (PGD) .....	169
Gráfico 10 - Ferramenta utilizada na elaboração do PGD .....	169
Gráfico 11 - Sondagem sobre os tipos de revistas em que os pesquisadores publicam.....	172
Gráfico 12 - Produções da UFC em Acesso Aberto x Acesso Restrito .....	173
Gráfico 13 - Acesso e leitura de pré-prints por parte dos pesquisadores .....	173
Gráfico 14 - Valores e princípios da Ciência Aberta .....	176
Gráfico 15 - Formas de compartilhamento de dados .....	177
Gráfico 16 - Identificadores digitais para autores usados pelos pesquisadores.....	180
Gráfico 17 - Canais informais de divulgação científica usados pelos pesquisadores .....	181
Gráfico 18 - Opinião dos pesquisadores sobre a responsabilidade institucional para a criação de serviços de dados .....	182
Gráfico 19 - Opinião dos pesquisadores sobre quem ou qual setor está qualificado para oferecer serviços de apoio ao pesquisador .....	183
Gráfico 20 - Opinião dos pesquisadores sobre a importância do apoio bibliotecário em serviços de dados .....	183
Gráfico 21 - Categoria dos cursistas que avaliaram o minicurso .....	239
Gráfico 22 - Contribuição do curso para a aprendizagem dos cursistas .....	240
Gráfico 23 - Grau de satisfação com itens gerais do curso .....	241

## LISTA DE QUADROS

Quadro 1 – Levantamento inicial nos processos de busca para compor o Referencial Teórico .....	26
Quadro 2 - Justaposição dos conceitos de Linked Data (Dados conectados), Linked Open Data (Dados abertos conectados) e Open Data (Dados abertos) .....	51
Quadro 3 - Princípios FAIR .....	89
Quadro 4 - Competências do bibliotecário de dados .....	96
Quadro 5 - Competências do bibliotecário na gestão de dados de pesquisa .....	100
Quadro 6 - Competências do Bibliotecário na Ciência Aberta .....	103
Quadro 7 - Características das principais correntes da Teoria Fundamentada em Dados.....	117
Quadro 8 - Características da Teoria Fundamentada e da Netnografia .....	124
Quadro 9 - Objetivos específicos, Instrumentos e fontes de coleta .....	125
Quadro 10 - Características e funcionalidades do OneNote .....	132
Quadro 11 - Distribuição dos sujeitos selecionados para a entrevista .....	146
Quadro 12 - Exemplos de algumas das percepções e conceitos sobre dados de pesquisa informados pelos respondentes .....	163
Quadro 13 - Percepção sobre a Ciência Aberta .....	174
Quadro 14 - Distribuição dos entrevistados de acordo com a área de conhecimento .....	189
Quadro 15 - Estratégias de organização e armazenamento dos dados .....	191
Quadro 16 - Percepção sobre os Dados de Pesquisa .....	197
Quadro 17 - Tipos de instrumento ou estratégias de documentação das pesquisas .....	202
Quadro 18 - Percepção sobre a Ciência Aberta .....	210
Quadro 19 - Número de interações em comentários nas postagens e mural de cada aula .....	235
Quadro 20 - Exemplo de interação entre os participantes - Aula 1 .....	236

## LISTA DE TABELAS

Tabela 1 - Total de repositórios por países da América Latina .....	66
Tabela 2 - Síntese das informações sobre o perfil dos respondentes .....	157
Tabela 3 - Familiaridade com as práticas e processos da Ciência Aberta (Porcentagem).....	171
Tabela 4 - Sobre o compartilhamento .....	178
Tabela 5 - Média de satisfação e nota do curso .....	241

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>10</b>
<b>2 APORTE TEÓRICO: PROCESSO DE CONSTRUÇÃO</b> .....	<b>20</b>
<b>2.1 Comunicação Científica: Acesso Aberto, Ciência Aberta e Dados de Pesquisa</b> .....	<b>29</b>
<b>2.2 Ciência Aberta: definições</b> .....	<b>36</b>
<b>2.3 Dados: explorando definições dos elementos em forma bruta</b> .....	<b>48</b>
<b>2.4 Gestão de dados de pesquisa</b> .....	<b>59</b>
<b>2.4.1 Ciclos de vida da pesquisa e dos dados</b> .....	<b>69</b>
<b>2.4.2 Plano de Gestão de Dados (PGD)</b> .....	<b>77</b>
<b>2.4.3 O papel do bibliotecário na Ciência Aberta e gestão de dados</b> .....	<b>91</b>
<b>3 METODOLOGIA</b> .....	<b>106</b>
<b>3.1 Caracterização da Pesquisa e Processo de construção metodológica</b> .....	<b>106</b>
<b>3.2 Instrumentos de coleta de dados</b> .....	<b>125</b>
<b>3.2.1 Questionário</b> .....	<b>126</b>
<b>3.2.2 Entrevista Semiestruturada</b> .....	<b>130</b>
<b>3.2.3 Diário de Campo Eletrônico e Caderno de Laboratório Eletrônico</b> .....	<b>132</b>
<b>3.3 Ação de Informação: proposta de minicurso online</b> .....	<b>136</b>
<b>3.4 Campo de Pesquisa</b> .....	<b>141</b>
<b>3.5 Universo e amostra de pesquisa</b> .....	<b>143</b>
<b>4 ANÁLISE DOS DADOS E DISCUSSÃO DOS RESULTADOS</b> .....	<b>148</b>
<b>4.1 Identificação das categorias de análise</b> .....	<b>149</b>
<b>4.2 Sondagem e investigação das práticas e necessidades dos pesquisadores</b> .....	<b>156</b>
<b>4.3 Aspectos da pesquisa de campo, entrevistas e aprofundamento das investigações</b> .....	<b>184</b>
<b>4.3.1 Dados</b> .....	<b>191</b>
<b>4.3.2 Ciência Aberta</b> .....	<b>204</b>

<b>4.3.3 Pesquisa .....</b>	<b>215</b>
<b>4.3.4 Biblioteca.....</b>	<b>219</b>
<b>4.3.5 Compartilhamento.....</b>	<b>222</b>
<b>4.4 Ação de Informação Educacional: Minicurso online .....</b>	<b>228</b>
<b>4.4.1 Avaliação do Minicurso Online .....</b>	<b>239</b>
<b>4.4.2 Proposta de Programa Para a Gestão de Dados de Pesquisa na UFC....</b>	<b>247</b>
<b>5 CONSIDERAÇÕES FINAIS.....</b>	<b>255</b>
<b>REFERÊNCIAS .....</b>	<b>261</b>
<b>APÊNDICE A - QUESTIONÁRIO .....</b>	<b>287</b>
<b>APÊNDICE B - ROTEIRO DE ENTREVISTA .....</b>	<b>301</b>
<b>APÊNDICE C – MODELO DO MINICURSO ONLINE .....</b>	<b>303</b>
<b>APÊNDICE D – QUESTIONÁRIO DE AVALIAÇÃO DO MINICURSO ONLINE</b>	<b>318</b>
<b>APÊNDICE E - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.....</b>	<b>320</b>
<b>APÊNDICE F – PLANO DE GESTÃO DE DADOS .....</b>	<b>323</b>
<b>APÊNDICE G – LISTA DE CURSOS SOBRE CIÊNCIA ABERTA E GESTÃO DE DADOS DE PESQUISA.....</b>	<b>327</b>
<b>APÊNDICE H – PROPOSTA DE PROGRAMA DE GESTÃO DE DADOS.....</b>	<b>329</b>
<b>APÊNDICE I – MINUTA DA PROPOSTA DE POLÍTICA DE GESTÃO DE DADOS DE PESQUISA DA UNIVERSIDADE FEDERAL DO CEARÁ.....</b>	<b>339</b>
<b>ANEXO A – PARECER CONSUBSTANCIADO DO COMITÊ DE ÉTICA EM PESQUISA .....</b>	<b>345</b>



## 1 INTRODUÇÃO

Desde os tempos remotos, considera-se o conhecimento científico primordialmente como um produto de colaboração humana. Na Antiguidade, por exemplo, credita-se aos gregos o estabelecimento da comunicação científica por meio da fala – em discussões filosóficas e reuniões realizadas na periferia de Atenas – e da escrita – com a conservação dos manuscritos copiados repetidas vezes (MEADOWS, 1999). Com o advento da ciência moderna, Galileu tornou-se conhecido por ser o responsável pela aplicação do método científico para validar o conhecimento produzido, no qual se baseou na lógica da observação, experimentação e explicação do fenômeno estudado por meio de teorias e processos da pesquisa. Assim como Galileu Galilei, também são creditados pela criação dos métodos científicos Francis Bacon e René Descartes. A partir de então, com a introdução do método, a produção de conhecimento e os resultados de pesquisas passaram a ser pautados na troca de ideias, em sugestões que agregassem valor ao estudo, no compartilhamento de informações e resultados que necessitam passar pelo crivo dos pares, membros da comunidade científica.

Fundamentalmente, a organização social da ciência contemporânea traz em seu bojo a colaboração como ponto-chave, ao mesmo tempo em que desvela um cenário de virtualização e internacionalização da ciência, além de incentivar a abertura e cooperação entre os pares, a fim de contribuir com o avanço do conhecimento científico, sendo este um pressuposto básico, especialmente pela tendência atual da valorização da produção coletiva do conhecimento, da livre circulação das informações, e mais recentemente, do compartilhamento de dados.

Diante disso, a ciência no século XX apresenta em seu processo evolutivo a presença de aprimoramento do aparato instrumental e tecnológico na coleta dos dados de pesquisa científica, por meio de uma série de equipamentos, tais como telescópios, satélites, sensores etc., ao passo que esses dados eram disponibilizados online, comprovando assim o seu caráter colaborativo em que muitos estudos foram desenvolvidos ao redor do planeta Terra. (BELL, 2011; COSTA, 2017; FOX; HENDLER, 2011; GREEN, 2011).

Esse panorama também recebeu influências da comunicação científica ao evidenciar o estado caótico das produções científicas, a exemplo da fragmentação das pesquisas, o aumento exponencial de publicações e a falta de integração da

ciência com a indústria e outros setores da sociedade já na década de 1930, conforme reportado por Bernal (c1935, 1939). Além disso, foi igualmente influenciada pelo contexto histórico da explosão informacional já prevista pioneiramente por Vannevar Bush (1945), com o crescimento da quantidade de relatórios e documentos, aliado ao desenvolvimento tecnológico e o uso dos computadores no período da Segunda Guerra Mundial (CAPURRO; HJØRLAND, 2003; WERSIG, 1993), cenário que marca o início da Ciência da Informação.

A partir da década de 90, a internet foi marcada pela explosão e evolução das Tecnologias da Informação e Comunicação (TIC) e possibilitou a mediatização das comunicações por computador (LOGAN, 2012), fazendo com que a informação passasse a ter um lugar de destaque no contexto social, da mesma forma em que se inseriu no cotidiano dos cidadãos (ALVES, 2017), isto é, na 'Era da Informação', do conhecimento e da 'sociedade em rede' (CASTELLS, 1999). Então, a informação passou a ser analisada como objeto valioso, tal como atualmente ocorre com os dados, considerados como o novo petróleo (PERKINS, 2017) na esfera social e acadêmica.

Por seu turno, a explosão informacional culminou com o fenômeno do dilúvio de dados científicos coletados e disponíveis online, marcando a Era do *Big Data* e da *e-Science*, denotando assim a relação da preocupação inicial da Ciência da Informação com a informação científica e tecnológica concernentes ao tratamento dos dados oriundos da pesquisa científica. Segundo Davenport (2014), *Big Data* é um conceito revolucionário e caracteriza-se por um considerável volume de dados desestruturados, provenientes de diferentes fontes e que demandam uma análise contínua dos mesmos, denominada de *streaming data*. Sobre o termo *e-Science*, na literatura científica também é conhecido como *data-driven science*, isto é, a ciência orientada para dados e está relacionada aos conceitos de ciberinfraestrutura, o quarto paradigma da ciência e o dilúvio de dados (COSTA, 2017; GRAY, 2009; HEY, TREFETHEN, 2005; MARCUM; GEORGE, 2010; MEDEIROS; CAREGNATO, 2012).

Outrossim, o crescimento da ciência e sua influência na comunicação científica tem demonstrado que o volume de pesquisas e da literatura científica seguiu o mesmo ritmo substancial de produção de publicações especializadas em várias áreas, conforme alertou Solla-Price (1976), que o exponencial aumento do número de publicações científicas duplicaria a cada dez ou quinze anos após a Segunda Guerra Mundial. Corroborando essa perspectiva, Fuller (1981) formulou a teoria da 'Curva de

Duplicação do Conhecimento' para explicar o vertiginoso crescimento do conhecimento humano e como este dobraria a cada século. Portanto, da mesma forma em que ocorre a produção maciça de novas pesquisas, também são gerados consideráveis volumes de dados, desvelando dessa forma o panorama da Era do *Big Data* e da *eScience*.

Emergem ainda nesse cenário do desenvolvimento científico e tecnológico novas formas de produção do conhecimento na ciência contemporânea, como a colaboração entre os cientistas por meio da internet, o surgimento de publicações ampliadas (*enhanced publications*), a criação de diversos recursos e plataformas de infraestrutura tecnológica com vistas a dar suporte ao volume de dados produzidos nas pesquisas e aos trabalhos online desenvolvidos pelos cientistas. Em outras palavras, a era da informação corresponde à explosão de dados (WURMAN, 1992), trata-se de uma ciência fortemente embasada na computação intensiva de dados, ou como costumam denominar alguns autores, o quarto paradigma (GRAY, 2009; HEY, TREFETHEN, 2005).

Inclusive, a Royal Society of London (2012) incentiva que as pesquisas sejam conduzidas de forma aberta com o intuito de permitir a publicação de teorias científicas, a disponibilização de dados experimentais e observacionais que permitam aos pares avaliar esses estudos, identificar erros, apoiar, rejeitar ou aprimorar teorias, corroborando o pensamento de Popper (1999) sobre o pesquisador estar pronto para colocar as suas teorias à prova e ao escrutínio da comunidade científica.

Por esse prisma, a adoção desse modelo cooperativo e aberto, compatível com as novas práticas de produção e difusão do conhecimento tem sido denominado como **Ciência Aberta** (*Open Science* ou Ciência 2.0). A *Open Science* se utiliza de tecnologias e ferramentas digitais colaborativas, preferencialmente *open source* (código-fonte aberto), inclusive com o foco no expressivo volume de dados, uso de métodos e outros recursos computacionais que vêm sendo empregadas nos últimos anos. Mais do que isso, a Ciência Aberta impacta diretamente no modo como as pesquisas são feitas e nos princípios de abertura em todo o ciclo de vida da pesquisa. Promove uma mudança sistêmica e incentiva o compartilhamento e a colaboração entre os pares. (FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH - FOSTER, 2015).

Atualmente, o tema Ciência Aberta está atrelado aos Dados de Pesquisa, o que torna difícil tratar de um desses sem abordar o outro. Portanto, nesse estudo, a

Comunicação Científica, a Ciência Aberta e os Dados de Pesquisas são temas que estão entrelaçados na trama de fios que envolvem a Gestão de Dados Científicos.

Dando seguimento ao liame que envolvem as temáticas supracitadas, a fim de promover a abertura, disponibilização e o acesso aos **dados de pesquisa** é preciso gerenciá-los, uma vez que, sem os dados, não é possível verificar e reproduzir experimentos, assegurar a integridade da pesquisa, colaborar com o fluxo e circulação das informações entre os pares e acelerar as etapas de descobertas em prol da ciência, contribuindo assim para a efetivação da Agenda 2030<sup>1</sup>, a qual tem como objetivo promover o desenvolvimento de sociedades mais justas, igualitárias e sustentáveis.

Para tanto, é imprescindível o gerenciamento dos dados de pesquisa, assim como a elaboração da documentação que facilite o processo de organização e tratamento para o depósito e futura disponibilização, conforme será apresentado neste trabalho. Sobre o gerenciamento de dados, trata-se de um conjunto de atividades que orientam a coleta, o armazenamento, o gerenciamento, a preservação e o compartilhamento dos dados provenientes de uma pesquisa. Por se apresentar em variadas formas, para serem efetivamente entendidos, os dados precisam estar ligados a um contexto. (SAYÃO; SALES, 2015; UNIVERSIDADE DE SÃO PAULO, 2019).

A literatura científica evidencia que a adoção dos princípios da Ciência Aberta, assim como o gerenciamento de dados científicos tem sido um desafio para os pesquisadores de qualquer área do conhecimento, e, por esse motivo, eles precisam do auxílio de um bibliotecário, profissional capaz de prestar serviços de apoio e suporte ao pesquisador por causa de suas competências e conhecimentos em padrões, normas e políticas de gestão de dados; estudos métricos; citações; ciclo de vida dos dados e da pesquisa, gestão da informação etc., especialmente com foco nas práticas e processos para a implantação de serviços em gerenciamento de dados científicos oriundos de pesquisas. (BORGMAN, 2015; COSTA, 2017; COX; KENNAN; LYON; PINFIELD, 2017; MORGAN; DUFFIELD; HALL, 2017; NIELSEN; HJØRLAND, 2013; SALES, 2014; SAYÃO; SALES, 2015; SILVA, 2016; TRIPATHI; SHUKLA; SONKER, 2017; VAN DEN EYNDEN; CORTI, 2017; XIA; WANG, 2014; YOON; SCHULTZ, 2017).

---

<sup>1</sup> Agenda 2030: <http://www.agenda2030.com.br/>.

Seguindo essa tendência, é preciso despertar nos pesquisadores a conscientização (*awareness*), além de promover o letramento por meio de programas de capacitação para ajudá-los a adquirir as competências necessárias para compreender o ciclo de vida da pesquisa e o ciclo de vida dos dados, para que possam gerenciar seus dados de pesquisa, elaborar seus Planos de Gestão de Dados (PGD's), escolher os formatos ideais para seus arquivos, conhecer as estratégias de versionamento, segurança e licenciamento dos dados, seguir padrões e protocolos, entre outras atividades concernentes a esse processo.

Ademais, trabalha-se com a hipótese de que todas as ações supracitadas podem surgir antes mesmo da oferta de infraestrutura, afinal, de nada adianta disponibilizar recursos e ferramentas se os pesquisadores não forem informados do que eles têm à sua disposição, das possibilidades existentes para o gerenciamento dos seus dados de pesquisa, assim como pode contribuir para uma maior sensibilização sobre a importância de gerenciar os próprios dados de pesquisa, quer seja visando o cumprimento de exigências institucionais ou de financiadores, para contribuir com a replicabilidade e reprodutibilidade da pesquisa, para a garantia de integridade da pesquisa científica, para a preservação digital dos dados coletados, para uso, disponibilização e reuso, entre outras ações de igual envergadura.

Além disso, por meio da experiência na prática profissional diária e durante a pesquisa de campo notou-se que os pesquisadores gerenciam seus dados de pesquisa estabelecendo critérios e regras próprias, ou seja, conforme o gosto pessoal, provavelmente, pelo desconhecimento e falta de políticas, diretrizes ou padrões estabelecidos que os orientem nesse aspecto, a exemplo da divulgação da existência do Plano de Gestão de Dados (PGD) e a necessidade de preencher um documento como este são ações simples que contribuem para que os pesquisadores passem a adotá-los como algo comum e parte do ciclo de vida da pesquisa.

A falta de políticas, diretrizes e padrões resulta no fenômeno 'cauda longa da ciência' (ANDERSON, 2004, 2006; SAYÃO, SALES, 2019), descrito por alguns autores como literatura cinzenta (resultados e dados de pesquisa) trancada em armários e gavetas, ou seja, são arquivos (documentos físicos ou digitais) que residem exclusivamente em laboratórios de pesquisa, ou em *pendrives*, computadores e HD externos, sem a garantia de *backup* frequente e que não foram publicados e divulgados. Muitos desses documentos são importantes para o avanço da ciência e tem a capacidade de divulgar o potencial dos estudos desenvolvidos pelos

pesquisadores e seus laboratórios, dentro e fora dos muros das instituições, porém, estão restritos, prestes à destruição, dano ou extravio, quando poderiam ser tratados, gerenciados, curados e preservados.

Diante do exposto, essa pesquisa justifica-se por mostrar a necessidade de se apresentar aos pesquisadores os conceitos e as práticas relacionadas ao gerenciamento de dados de pesquisa através da mediação em ações de informação que promovam a competência em informação e a competência em dados, uma vez que a partir da conscientização e conhecimento dessas práticas os pesquisadores de diversas áreas se sentirão mais capazes e confiantes para cumprir as etapas de coleta, armazenamento, gerenciamento e compartilhamento dos dados provenientes da pesquisa científica.

Diversas instituições internacionais, pesquisadores e bibliotecas têm investido fortemente em projetos de capacitação presenciais e online empreendidos ao redor do mundo, em iniciativas individuais ou coletivas de pesquisadores, a exemplo dos *workshops* promovidos pelas bibliotecas do *Massachusetts Institute of Technology* (MIT); dos cursos da iniciativa *Open Science MOOC*<sup>2</sup>, coordenado por Jonathan Tennant e colaboradores; em universidades e com o envolvimento das bibliotecas que lidam com dados de pesquisa, como é o caso pioneiro do *MANTRA: Research Data Management Training*<sup>3</sup>, produzido pela Universidade de Edimburgo; e parcerias com universidades e entidades como a União Europeia no financiamento do projeto *Facilitate Open Science Training For European Research* (FOSTER)<sup>4</sup>; os cursos de capacitação sobre Ciência Aberta promovidos pela Fiocruz (Brasil)<sup>5</sup>; o curso em Ciência Aberta oferecido pelo Ministério da Ciência, Tecnologia e Ensino Superior de Portugal<sup>6</sup>, entre outros que oferecem certificação mediante o pagamento de uma taxa<sup>7</sup>.

Dessa forma, acreditamos que a partir do desenvolvimento de ações de informação, capacitação, educação de usuários e promoção da competência em informação e em dados (*data literacy*) é possível sensibilizar todos os *stakeholders*

---

<sup>2</sup> Open Science MOOC: <https://opensciencemooc.eu/>.

<sup>3</sup> MANTRA: <https://mantra.edina.ac.uk/>.

<sup>4</sup> FOSTER: <https://www.fosteropenscience.eu/about#theprevproject>.

<sup>5</sup> Ciência Aberta (Fiocruz): <https://portal.fiocruz.br/ciencia-aberta>.

<sup>6</sup> Curso sobre Ciência Aberta (Portugal): <https://www.ciencia-aberta.pt/mooc>.

<sup>7</sup> Exemplos de curso sobre Gestão e compartilhamento de dados científicos disponível no Coursera e oriundo do MANTRA: <https://www.coursera.org/learn/data-management> e curso sobre Ciência Aberta da Delft University of Technology disponível no EDX: <https://www.edx.org/course/open-science-sharing-your-research-with-the-world-2>. Confira também a lista de cursos sobre gestão de dados de pesquisa e Ciência Aberta disponível em: <https://is.gd/cursosrdmos>.

(instituições, bibliotecários, pesquisadores e usuários em geral) sobre a importância dos dados, sua influência diária na vida das pessoas e do governo, e, principalmente, na adoção de práticas da Ciência Aberta, como a abertura e o compartilhamento de dados de pesquisa, contribuindo para convencer os usuários na adoção desses princípios.

Para tanto, torna-se imprescindível, antes de tudo, investigar questões mais elementares, o que conduz aos seguintes questionamentos: Qual a percepção dos pesquisadores (Professores e alunos de Pós-Graduação *Stricto Sensu* na Universidade Federal do Ceará) em relação aos dados de pesquisa e aos princípios da Ciência Aberta? A questão principal supracitada suscita ainda outra pergunta: Quais são os desafios e dificuldades encontradas pelos pesquisadores no gerenciamento dos dados, no compartilhamento de dados e na adoção das práticas da Ciência Aberta?

O objetivo geral se concentra em investigar as práticas e necessidades informacionais dos pesquisadores dos cursos de Pós-Graduação da Universidade Federal do Ceará (UFC) concernentes ao gerenciamento de dados científicos e à Ciência Aberta. Para tal, delineamos os seguintes objetivos específicos:

- Analisar a percepção dos pesquisadores sobre os princípios da Ciência Aberta e o gerenciamento de dados de pesquisa;
- Averiguar quais as práticas e as necessidades informacionais destes pesquisadores referentes a estas temáticas;
- Elaborar uma proposta de Política de Gestão de Dados com a finalidade de incentivar a adoção de práticas em gerenciamento de dados de pesquisa;
- Propor Programa de Gestão de Dados (PGDP) com o objetivo de sugerir a criação de serviços de dados na UFC;
- Desenvolver ações de informação com vistas a testar um piloto de curso online como parte integrante do programa.

Para os fins desta pesquisa, a ideia de programa concentra-se na mediação da informação, baseada na explicação de Silva e Farias (2017, p. 115): “A mediação atua como um programa de ação, intervenção e interferência, e os serviços de informação se constituem como uma das principais atividades técnica, pedagógica, institucional e pragmática da mediação.” Assim, volta-se para os estudos de usuários, busca uma aproximação da comunidade atendida e incentiva a autonomia dos

usuários, por exemplo, por meio de capacitações e treinamentos a fim de que tenham condições de escolher, apreender e apropriar-se da informação.

Nesse caso específico, compõe a proposta do programa uma sugestão de criação de serviços de informação voltado para os dados de pesquisa, uma minuta de diretrizes e a oferta de capacitações sobre o gerenciamento dos dados de pesquisa.

A título de esclarecimento, conforme a visão de Rozados (2004), serviço de informação é toda atividade voltada para a identificação, aquisição, processamento, transmissão e fornecimento da informação ou na satisfação das necessidades informacionais dos usuários. Nesse sentido, é natural que a biblioteca funcione como o canal de transmissão do conhecimento e como aparato educativo-cultural por desempenhar a mediação em diversos serviços, funções, atividades-meio e atividades-fim, quer seja reunindo, organizando e disseminando a informação ou dialogando com os seus usuários presencialmente ou virtualmente (SILVA; FARIAS, 2017; VARELA; BARBOSA; FARIAS, 2014).

Portanto, diante da emergência do tema nos últimos anos e no contexto atual, torna-se cada vez mais importante diminuir as incertezas e consolidar as habilidades voltadas para a busca de informações, englobando a compreensão dos aspectos afetivos envolvidos na busca e uso da informação, conforme preconizado por Carol Kuhlthau (1991). Nesse aspecto, de acordo com a abordagem de Kuhlthau (1993), a mediação entre sujeitos permite que as pessoas façam conexões, movam-se do concreto ao abstrato de tal forma que reconheçam a necessidade de saber mais sobre os assuntos que despertam a sua curiosidade, fazendo com que esses usuários estudem e atinjam maior nível de aprofundamento e compreensão.

Sobre o gerenciamento dos dados de pesquisa vale ressaltar que se torna um círculo vicioso exigir dos pesquisadores a abertura e compartilhamento de dados de pesquisa sem que existam políticas nacionais, infraestrutura consistente de apoio à ciência e principalmente sem a capacitação dos pesquisadores para que estes possam desenvolver competências em informação e em dados. Portanto, para transformar esse aspecto em um círculo virtuoso, o papel do bibliotecário nesse contexto macro tem relevância e uma grande agenda de questões a serem discutidas, especialmente no investimento em ações de informação e mediação, no desenvolvimento de políticas nacionais, de diretrizes e modelos capazes de contribuir para o aprendizado dos pesquisadores e na transformação dessa realidade. Exatamente por isso, estabeleceu-se como ponto de partida investigar qual a



percepção dos pesquisadores para traçar os planos, estabelecer estratégias e aplicar as propostas deste estudo.

Complementando o desenho da pesquisa, esta tem relação com o paradigma social da Ciência da Informação que, segundo Capurro (2003), considera a informação como um fenômeno social, está em suportes físicos e ao mesmo tempo é percebida como algo impalpável, presente no cerne dos processos cotidianos e das práticas histórico-sociais. Por esse prisma, não é por acaso que Shera (1973) afirma ser impossível conhecer os processos intelectuais de uma sociedade baseado no estudo de um indivíduo isoladamente. Há que ser levado em consideração o contexto social desse indivíduo, seu comportamento e práticas a fim de identificar suas necessidades informacionais sem dissociá-lo de sua cultura.

Além disso, Borko (1968) caracteriza a Ciência da Informação como uma ciência interdisciplinar e destaca que é uma área atuante em pesquisas sobre a definição de padrões, fluxos de informação e as formas de processá-las, comportamento e estudos de usuários, além de se preocupar com uma série de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação, representações da informação em sistemas naturais e artificiais etc., logo, tem componentes de uma ciência pura e aplicada que cria serviços e produtos. Sobre a interdisciplinaridade da Ciência da Informação, Corrêa e Rozados (2017, p. 2), afirmam que ela “[...] está aberta ao diálogo com outras disciplinas e ao emprego de métodos de pesquisa que possam contribuir para a resolução dos problemas práticos e teóricos que emergem em seu campo de atuação.”

Para tanto, após o delineamento das considerações iniciais, o desenvolvimento do estudo foi organizado conforme o que se segue, além desta seção introdutória. Na seção intitulada ‘Aporte teórico’ apresenta-se a construção do processo de busca e sistematização da escolha das fontes de informação para compor a abordagem utilizada nas etapas seguidas e proceder a revisão de literatura, o que foi fundamental para a escolha das publicações selecionadas e para embasar teoricamente este estudo. Em seguida, são discutidas as relações da comunicação científica, dos dados de pesquisa e da Ciência Aberta. São apresentadas ainda as definições de Ciência Aberta e de Dados de Pesquisa até chegar na exposição de diversos conceitos sobre o que são dados de pesquisa, quais são os seus tipos, natureza etc. Discorreremos ainda sobre a gestão de dados de pesquisa à luz de Sayão

e Sales (2015), Silva (2016), da iniciativa Foster (2015, 2019) e muitas outras entidades importantes relacionadas aos dados de pesquisa para abordar o Plano de Gerenciamento de Dados (PGD).

Encerrando a seção do aporte teórico, discorreremos acerca do papel do bibliotecário na Ciência Aberta e na gestão de dados de pesquisa. Expusemos uma série de casos baseados na literatura científica que demonstram a importância do bibliotecário nesse cenário, quais são as competências exigidas, quais as características que estes profissionais devem buscar obter, além de mostrar que o bibliotecário ainda agrega outras questões adicionais relacionadas à Ciência Aberta e a gestão de dados.

A seção da metodologia se inicia com uma breve caracterização da pesquisa, descrição dos métodos empregados e por meio da triangulação metodológica adicionada a este estudo com a escolha da *Grounded Theory* (Teoria Fundamentada em dados) e da Netnografia. Como técnicas de pesquisa apresenta-se a análise documental e a observação participante, ambas já são naturalmente incorporadas e fazem parte dos processos de coleta e análise de dados na Teoria Fundamentada e na Netnografia. Como instrumentos de coletas de dados utilizamos o questionário, entrevista e diário de campo eletrônico (*One Note*) e *ResearchSpace* (RSpace) para as anotações, registros de notas de campo e na construção de memorandos. Discorreremos ainda sobre a proposta de ação de informação a ser aplicada nesta investigação, que é um minicurso online com carga horária de 15h para professores, alunos e técnico-administrativos em educação que estejam cursando ou tenham concluído uma pós-graduação *stricto sensu* (mestrado ou doutorado) na Universidade Federal do Ceará (UFC). Finalmente, encerramos a seção da metodologia apresentando o campo de estudo, o universo e amostra da pesquisa.

Na quarta seção dessa dissertação traz-se a análise dos dados e discussão dos resultados, onde serão esmiuçadas cada etapa da pesquisa, os dados coletados, a formação das categorias e os apontamentos devidos relacionados. Na última seção, foram tecidas as considerações finais a respeito da investigação desenvolvida, assim como a apresentação de uma agenda de pesquisa como sugestão de estudos futuros.

## 2 APORTE TEÓRICO: PROCESSO DE CONSTRUÇÃO

Com a finalidade de compor o referencial teórico e a elaboração da revisão de literatura inicial deste estudo, de maneira condizente com o contexto atual e por ser um tema emergente, alguns passos essenciais para a construção do aporte teórico da pesquisa foram traçados. O primeiro deles foi a delimitação do tema por meio da técnica da pirâmide invertida (HOHENDORFF, 2014), conforme a Figura 1:

Figura 1 - Pirâmide invertida da delimitação do tema



Fonte: Elaborada pela autora (2019), baseado em Hohendorff (2014, p. 42).

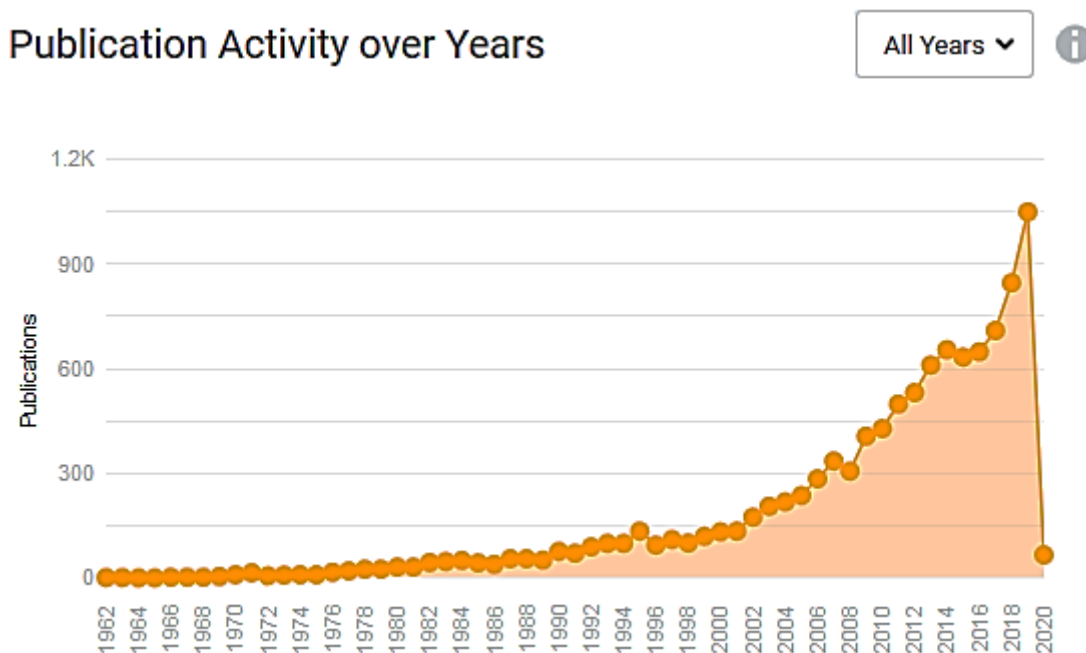
Com efeito, a aplicação da técnica da pirâmide invertida antes de realizar o levantamento bibliográfico auxilia na definição e delimitação do tema de pesquisa; na busca e organização dos materiais para a revisão de literatura; e para a elaboração do roteiro a fim de compor o referencial teórico.

Ressalta-se que, na Teoria Fundamentada em Dados (*Grounded Theory*) – método principal adotado para este estudo –, de acordo com a visão de Strauss e Corbin (2008), a revisão de literatura não faz parte da fase inicial da pesquisa, pois

somente a partir da coleta e análise dos dados é que as informações obtidas direcionarão o pesquisador para o que se deve buscar na literatura científica, pois esta é usada como fonte secundária de dados. Naturalmente, o pesquisador pode consultar a literatura antes de iniciar o seu projeto de pesquisa para ajudar a formular suas primeiras questões e ponto de partida durante a realização de entrevistas e observações, mas, em seguida o pesquisador irá voltar-se para as questões e conceitos que surjam da análise de dados.

Nas palavras de Dantas *et al.* (2009, p. 3), “O método é circular e, por isso, permite ao pesquisador mudar o foco de atenção e buscar outras direções, reveladas pelos dados que vão entrando em cena.” Por ter essa característica, por si só isso já se constituiu em um grande desafio para o planejamento e delineamento da pesquisa. Inclusive, o procedimento de busca foi uma tarefa efetuada no início, durante e na redação final desta pesquisa, tendo em vista o crescimento exponencial e interesse pelo tema em estudo, especialmente nos últimos 10 anos, como se pode comprovar a partir dos dados obtidos através do Wizdom.ai<sup>8</sup> em uma busca livre realizada correspondente ao termo “Gestão de dados” (*Data Management*)<sup>9</sup> (Figura 2).

Figura 2 - Publicações sobre Data Management ao longo dos anos



Fonte: Wizdom.ai (2020).

<sup>8</sup> Wizdom.ai: <https://www.wizdom.ai/>.

<sup>9</sup> A análise completa está disponível em: [https://www.wizdom.ai/topic/data\\_management/759312](https://www.wizdom.ai/topic/data_management/759312).



verificar os dados bibliográficos salvos no gerenciador e encontrar documentos semelhantes ou compatíveis que são sugeridos de acordo com as pesquisas realizadas e referências salvas por meio do uso de inteligência artificial.

Dando prosseguimento, a revisão de literatura consiste na contextualização de um problema e análise de aspectos já descritos e presentes na literatura científica a fim de construir o referencial teórico de uma pesquisa e buscar referências sobre o tema (ALVES-MAZZOTTI, 2002; CERVO; BERVIAN, 2002). Então, o segundo passo foi a pesquisa no Portal de Periódicos da Capes com o intuito de identificar os principais conceitos e definições em gestão de dados de pesquisa, principalmente visando a exploração das características, dimensões e tendências mundiais, particularmente, com foco na Ciência Aberta por ser um tema correlato e considerado como um termo guarda-chuva que engloba também a *gestão de dados científicos* ou *gestão de dados científicos*, conforme se identificou durante as buscas o uso desses dois termos para se referir ao mesmo assunto na Língua Portuguesa.

O Portal de Periódicos da Capes tem como vantagem principal ser um metabuscador, isto é, reunir várias ferramentas de busca em um único espaço, além de disponibilizar um acervo com mais de 48 mil títulos de periódicos com texto completo, 130 bases de dados referenciais e de resumos, 41 bases de dados estatísticas, 64 bases de dissertações e teses, 48 obras de referências, 15 bases de conteúdos audiovisuais, 14 bases de arquivos abertos e redes de e-prints, 12 bases de patentes, 2 bases de normas técnicas e mais de 275 mil documentos (anais de eventos, relatórios, livros, anuários, manuais etc.).

As buscas foram realizadas utilizando os termos em inglês: “*Research Data Management*” (Gestão de Dados de Pesquisa), “*Open Science*” (Ciência Aberta) e “*Data Sharing*” (Compartilhamento de Dados), contudo, é necessário ressaltar que, em Língua Inglesa, outros termos correlatos ou relacionados ao termo “*Research Data Management*” surgiram durante o levantamento bibliográfico, tais como “*Data Management*” (Gestão de dados), “*Data Stewardship*” (Curadoria de dados), “*Data Steward*” (Curador de dados). Destaca-se ainda que os termos usados nos textos de Língua Portuguesa sobre o assunto costumam variar para se referir à Gestão de Dados de Pesquisa, são eles: Dados científicos, Gestão de Dados Científicos, Dados de investigação, Gestão de Dados de Investigação. Neste estudo, adotaremos os termos “Dados de pesquisa” e “Gestão de Dados de Pesquisa”, podendo usar uma vez ou outra ao longo deste trabalho o termo “Dados científicos” como sinônimo.

Para que o retrato atual e o contexto da pesquisa sobre a temática seja o mais fidedigno possível ao estado da arte e que oportunize a construção de um amplo quadro teórico, estabeleceu-se como critério principal a procedência das fontes científicas formais, tais como livros, artigos científicos, dissertações e teses, além disso, apostou-se na variedade e atualidade dos materiais consultados por meio de fontes tais como os sites de entidades oficiais que recomendam normas, boas práticas e critérios para a Ciência Aberta e gestão de dados de pesquisa.

Novas buscas foram realizadas e combinadas a outras fontes de pesquisa após a primeira coleta de dados (aplicação do questionário e das primeiras entrevistas). Assim, recorreu-se à rede social *Twitter* de forma direta e também por intermédio do recurso de busca presente no software de análise qualitativa *Atlas.ti* na extração de tweets por meio do uso de hashtags, e conseqüentemente, salvaram-se vários *tweets* e materiais indicados nas postagens. O software mencionado é o mesmo que foi utilizado para a análise qualitativa deste estudo, conforme será visto em seção destinada para essa finalidade. Retomando a descrição dos procedimentos sobre as postagens do *Twitter* salvas, estas foram inseridas em listas contidas no perfil pessoal da pesquisadora (respectivamente denominadas *RDM – Research Data Management, Open Science, Reproducibility, Protocols*), recurso disponível no próprio *Twitter*, além de terem sido salvos os registros bibliográficos das fontes no gerenciador de referências *Mendeley* por meio do *plugin* disponível no aplicativo para celulares *Android*, pois a maioria dos acessos à referida rede social se deu via telefone celular.

O *Twitter* foi escolhido e empregado dessa forma porque viabiliza o acompanhamento das notícias, atualizações, *webinars*, oferta de materiais e cursos, dicas de artigos e outras publicações ligadas aos temas. Ademais, costuma ser uma rede bastante usada por pesquisadores e instituições nacionais e internacionais ligadas à Gestão de Dados de Pesquisa e Ciência Aberta. Também foram fundamentais para a composição do referencial os materiais de cursos sobre gestão de dados de pesquisa e Ciência Aberta de várias instituições nacionais e internacionais.

Nessa perspectiva, Elias *et al.* (2012) afirma que esse tipo de investigação permite identificar novas temáticas, apontar perspectivas, estabelecer relações com a literatura existente, o que contribui para a consolidação de uma área do conhecimento. Bentes Pinto e Cavalcante (2015, p. 32) enfatizam que “a construção de um aporte teórico reflexivo e crítico sobre o estado da arte de uma determinada área temática,

desde as suas concepções históricas, culturais e ideológicas até o que há de mais atual sobre o estudo” são fundamentais para a elaboração da revisão bibliográfica e análise documental.

O terceiro passo para subsidiar a referida revisão foi a realização de um levantamento bibliográfico automatizado através do software bibliométrico *Publish or Perish* (PoP), sendo este utilizado como uma ferramenta auxiliar para complementar o processo de busca com o intuito de identificar os autores mais citados nas temáticas supracitadas, constituindo-se em um critério adicional para a seleção dos documentos e fontes de pesquisa.

O PoP recupera e analisa citações através de seis fontes de dados: *Crossref*, *Google Scholar*, *Google Scholar Profile*, *Microsoft Academic*, *Scopus* e *Web of Science*. O PoP foi criado em 2006 e idealizado pela bibliotecária Ann-Wil Harzing com o intuito de calcular o índice H referente às métricas de citações de pesquisadores e suas publicações. Além disso, o PoP permite apresentar o impacto de uma pesquisa por meio da pesquisa bibliométrica, serve como ferramenta para escolher em qual periódico publicar, como recurso para buscar manuscritos a fim de fazer um levantamento bibliográfico automatizado, sendo ainda bastante utilizado como instrumento de auxílio em uma revisão sistemática, entre outras funções (Figura 4).

Figura 4 - Interface do software Publish or Perish – PoP

The screenshot displays the Publish or Perish software interface. At the top, there is a menu bar (File, Edit, Query, View, Help) and a toolbar. Below this is a table of search results with columns for Query, Source, Papers, Cites, Cites/y..., h, g, h/kno..., h/ann..., \*C..., Query date, Cache date, and Las... The table lists several queries from Google Scholar, Microsoft Academic, and other sources, with their respective metrics.

Below the table, there is a section for "Google Scholar query" with search filters for Authors, Publication/Journal, ISSN, and Title words. There are also buttons for "LookUp", "Clear All", "Revert", "Copy", and "New".

The bottom section shows "Metrics" and "Results". The "Metrics" section includes a table with columns for Cites, Per year, Rank, Authors, and Title. The "Results" section shows a list of publications with columns for Year, Publication, and Publisher.

Query	Source	Papers	Cites	Cites/y...	h	g	h/kno...	h/ann...	*C...	Query date	Cache date	Las...
"Open Science" AND "Research Data Management" AND "Data sharing"	Google Sc...	819	4570	21.06	30	59	19	0.09	18	24/05/2019	24/05/2019	0
Open Access OR Acceso Aberto	Microsoft ...	3000	81073	2026.83	127	219	92	2.30	199	20/04/2019	20/04/2019	0
Authorship AND Contributorship AND Acknowledge AND Integrity	Google Sc...	837	17978	122.30	56	122	45	0.31	44	09/04/2019	09/04/2019	0
"Information literacy" AND "Data literacy" AND "Open Science" AND "Research Data M...	Google Sc...	68	1157	128.56	11	34	8	0.89	4	31/03/2019	31/03/2019	0
"Open Science" AND "Research Data Management" AND "Training" OR "Course" OR "C..."	Google Sc...	206	1284	55.83	19	34	13	0.57	6	31/03/2019	31/03/2019	0
"Open Science" AND "Data Literacy" AND "Research Data Management"	Google Sc...	141	1324	147.11	13	35	11	1.22	4	31/03/2019	31/03/2019	0

Metrics	Cites	Per year	Rank	Authors	Title	Year	Publication	Publishe...
Publication years: 1802-2019	607	86.71*	3	CL Borgman	The conundrum of sharing research data	2012	Journal of the American S...	Wiley Or
Citation years: 217 (1802-2019)	409	18.59*	78	WK Michener, JW ...	Nongeospatial metadata for the ecological sciences	1997	Ecological ...	Wiley Or
Papers: 819	311	34.56*	92	C Yang, R Raskin, ...	Geospatial cyberinfrastructure: past, present and future	2010	..., Environment and Urban...	Elsevier
Citations: 4570	152	30.40*	28	L Corti, V Van den ...	Managing and sharing research data: a guide to good practice	2014	books.g...	books.g...
Cites/year: 21.06	130	21.67*	19	JC Prado, MÁ Mar...	Incorporating data literacy into information literacy programs: Core competencies and contents	2013	Libri	degruyte
Cites/paper: 5.98	95	15.83*	5	H Pampel, P Vierk...	Making research data repositories visible: The re3data. org registry	2013	PloS one	journals
2310.36	91	30.33*	8	RD Morey, CD Ch...	The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review	2016	... Open Science	royalsoci
476.23	88	14.67*	4	KG Akers, J Doty	Disciplinary differences in faculty research data management practices and perspectives	2013	International Journal of Di...	ijdc.net
2.52	69	17.25*	40	L Candela, D Cast...	Data journals: A survey	2015	Journal of the Association...	Wiley Or
30	66	9.43	34	S Corral	Roles and responsibilities: Libraries, librarians and data	2012	Managing research data	books.g...
g-index: 59	64	9.14	62	C Willis, J Greenb...	Analysis and synthesis of metadata goals for scientific data	2012	Journal of the American ...	Wiley Or
h-index: 19	59	11.80*	13	P Andreoli-Versba...	Open access to data: An ideal professed but not practised	2014	Research Policy	Elsevier
h1norm: 0.09	55	27.50*	83	SIC Janssen, CH P...	Towards a new generation of agricultural system data, models and knowledge products: Information and com...	2017	Agricultural systems	Elsevier
*Count: 18	55	13.75*	150	E Afgan, S Slogge...	Genomics virtual laboratory: a practical bioinformatics workbench for the cloud	2015	PloS one	journals
	54	18.00*	67	PA Meyers, S Socia...	Data publication with the structural biology data grid supports five analysis	2016	Nature...	nature.c...
	54	18.00*	153	D Lakens, J Hilgar...	On the reproducibility of meta-analyses: Six practical recommendations	2016	BMC psychology	bmcpsy...

Fonte: Publish or Perish (2019).



Assim, além de utilizar o PoP, em maio de 2019 realizamos buscas exaustivas em fontes distintas sobre o tema e refinamos esse processo a fim de encontrar relações da temática principal com outros temas relacionados. Os termos foram pesquisados em inglês a fim de ampliar o número de materiais recuperados, compondo a seguinte *string*: “Open Science” AND “Research Data Management” AND “Data sharing”. A partir dos resultados recuperados (819 registros) baseados nos termos de busca principais buscamos identificar materiais que tivessem relação com mediação e competência em informação; competência em dados; alguns relatos ou experiências em cursos e treinamentos; educação de usuários; desenvolvimento de programas de capacitação e conscientização visando a etapa de aplicação da pesquisa e a realização do minicurso online. Alguns artigos em chinês, japonês, francês, alemão, russo e outros idiomas semelhantes foram excluídos, selecionando-se apenas os manuscritos em inglês, português e espanhol. Esses procedimentos ajudaram a inter-relacionar os assuntos, aprofundar o referencial teórico, auxiliar na proposta de construção de um programa de capacitação dos usuários (pesquisadores) e inclusive contribuiu para a criação das diretrizes em Ciência Aberta, gestão de dados e entre outras ações necessárias para a universidade. Alguns dos resultados recuperados e selecionados durante os diferentes processos de busca para aprofundamento sobre o tema em estudo estão expressos no quadro 1 a seguir:

Quadro 1 – Levantamento inicial nos processos de busca para compor o Referencial Teórico

(continua)

<b>CIÊNCIA ABERTA</b>	<i>Definições</i>	Albagli (2015); Albagli, Appel e Maciel (2013); Bezjak <i>et al.</i> (2018); David (2008); Fecher e Friesike (2014); Fiocruz (2019); Foster (2015); Griffith University (2019); Murray-Rust <i>et al.</i> (2010); Nielsen (2011a, 2011b); OECD (2015); Orion (2018); Tennant (2018); Tennant <i>et al.</i> (2019)
	<i>Colaboração e Pesquisa aberta</i>	Albagli, Appel e Maciel (2013); Albagli; Clinio; Raychtock (2014); Arrow (1962); Dasgupta e David (1987); Egbert (2018); Griffith University (2019); Hagstrom (1965, 1972); Kraker <i>et al.</i> (2016); McKiernan <i>et al.</i> (2016); Merton (1973a; 1973b); Nelson (1959); Olson; Zimmerman e Bos (2008); Polanyi, Ziman e Fuller (2000); Popper (1999); Royal Society of London (2012); Stephan (1996); Stephan (1996); Tennant <i>et al.</i> (2019); Walstad (2002)

Quadro 1 – Levantamento inicial nos processos de busca para compor o Referencial Teórico

(continuação)

	<i>Conhecimento como bem público</i>	Agenda 2030; Araújo (2019); Bourdieu (1975); Cruz (2019); Graham (2013); Mauss (2003); Polanyi, Ziman e Fuller (2000); Santos (2018); Suber (2007); Swartz (2008)
	<i>Ecologia de saberes</i>	Boaventura de Sousa Santos (2007, 2008, 2010, 2014)
CIÊNCIA ABERTA	<i>Comunicação científica</i>	Acosta-Hoyos (1980); Bernal (c1935, 1939); Costa (2006); Crane (1988); David (2008); Garvey (1979); Meadows (1999); Merton (1973b); Mueller (1994); Oliveira (2009)
	<i>Sociologia da Ciência</i>	Bourdieu (1975); Dasgupta e David (1987, 1994); Egbert (2018); Merton (1973a, 1973b); Hagstrom (1965); Stephan (1996); Walstad (2002); Ziman (2000); Ziman, Etzkowitz e Leydesdorff (1997, 2000)
	<i>Ciência, Tecnologia e Sociedade da Informação</i>	Albagli (2015); Bresser-Pereira (2018); Castells (1999); Drucker (1968); Galbraith (1979); May (2000); Moulner Boutang (2010); Schiller (2011)
GESTÃO DE DADOS DE PESQUISA	<i>Definições</i>	Ball (2012); Bécard <i>et al.</i> (2016); Borgman (2007, 2010, 2019); Corti; Van den Eynden; Bishop; Morgan-Brett (2011); CESSDA (2017); Costa (2017); European Commission (2016); Foster (2019); Green; Macdonald; Rice (2009); ICPSR (2012); Lehmkuhl; Gavron, Alves, Mesguich <i>et al.</i> (2012); Moraes de Bem (2018); LERU (2018); Sales (2014); Sayão e Sales (2015, 2019); USP (2019)

Quadro 1 – Levantamento inicial nos processos de busca para compor o Referencial Teórico

(conclusão)

	<i>Quarto Paradigma e e-Science</i>	Albagli, Appel e Maciel (2013); Bell (2011); Ferreira (2018); Fox e Hendler (2011); Gray (2009); Green (2011); Hey e Hey (2016); Hey e Trefethen (2005); Moura (2018); Sales (2014); Sayão e Sales (2015)
	<i>Compartilhamento de dados</i>	Borgman (2007, 2012), Berners-Lee (2006); Cox, Kennan, Lyon, Pinfield (2017); Digital Science e Figshare (2017); FAIR Principles; Morgan, Duffield, Hall (2017); Nielsen; Hjørland (2013); Sales (2014); Sayão e Sales (2015, 2019); Silva (2016); Tripathi, Shukla, Sonker (2017); Van den Eynden e Corti (2017); Xia e Wang (2014); Yoon e Schultz (2017)
<b>DADOS DE PESQUISA</b>	<i>Definições (Dados)</i>	Brasil (2019); Davenport (1998); Nielsen e Hjørland (2014); Oliveira (2002); ONU (2014); Open Knowledge International Foundation (2017); Padoveze (2000); Setzer (1999); Toffler (1995); Zins (2007, 2011)
	<i>Definições (Dados de pesquisa)</i>	Aventurier (2016); Corti, Van den Eynden, Bishop e Morgan-Brett (2015); Murray-Rust <i>et al.</i> (2010); NEH (2018); NIH (2018); NSF (2018); OECD (2007); Sales (2014); Sayão e Sales (2015); Universidade de Leeds (2019)
	<i>Tipologia</i>	Green, Macdonald e Rice (2009); Corti, Van den Eynden, Bishop e Morgan-Brett (2015); Sales (2014); Sayão e Sales (2015); Universidade de Leeds (2019)
	<i>Plano de Gestão de Dados (PGD)</i>	Erway <i>et al.</i> (2016); European Commission (2016); Foster (2019); ICPSR (2012); Sayão e Sales (2015); Silva (2016)
	<i>Ciclo de Vida da Pesquisa e dos Dados</i>	DataOne (2019); DCC (2019); Foster (2019); UK Data Archive (2014); Sant'Ana (2016); Sayão e Sales (2015); University of California, Santa Cruz (2015)

Fonte: Elaborado pela autora (2019).

Nesta investigação, o rigoroso processo de pesquisa bibliográfica em fontes formais (livros, artigos, periódicos, bases de dados, Portal da Capes, sites de entidades oficiais etc.) para a construção do referencial teórico e o estudo bibliométrico utilizando-se o software *Publish or Perish* se deu exclusivamente para identificar os autores mais citados à frente do tema investigado. Esses critérios foram adotados como uma maneira adicional de selecionar esses documentos e buscar outras fontes de informação úteis para este estudo.

Diante do exposto, a busca no Portal da Capes e no PoP complementaram uma a outra, especialmente reforçando as forças e fraquezas na recuperação das informações e dos dados de cada uma, conforme os estudos bibliométricos e da cobertura de indexação realizados por Prins *et al.* (2016) e Martín-Martín *et al.* (2018)<sup>11</sup> sobre o número de citações encontradas no *Google Scholar*, *Web of Science* e *Scopus*, sendo estas duas bases de dados presentes no Portal de Periódicos da Capes.

## **2.1 Comunicação Científica: Acesso Aberto, Ciência Aberta e Dados de Pesquisa**

Esta seção visa refletir a respeito do crescimento da produção científica, traçando um breve panorama sobre a comunicação científica até chegar nas mudanças que ocorreram no *modus operandi* de se fazer pesquisas científicas na contemporaneidade. Com efeito, percebe-se que o funcionamento da ciência é abordado por diferentes campos científicos, por exemplo, a Sociologia, a Filosofia, a Economia, entre outros. (BOURDIEU, 1975; DASGUPTA; DAVID, 1987, 1994; EGBERT, 2018; HAGSTROM, 1965; MERTON, 1973b; STEPHAN, 1996; WALSTAD, 2002).

Os aspectos históricos e a evolução da comunicação científica se iniciam na Grécia Antiga, por volta do ano 700 a.C., em reuniões dedicadas ao debate de questões filosóficas realizadas na periferia de Atenas. (MEADOWS, 1999). Nesse período, prevalecia a cultura da oralidade, e, muitos dos manuscritos e registros feitos eram escritos e copiados manualmente. Mais tarde, no século XV, surge a invenção dos tipos móveis de Gutenberg, revolucionando o processo de produção do livro. A

---

<sup>11</sup> Os referidos estudos foram comentados por Ernesto Spinak (2019) no Blog SciELO em Perspectiva: <http://twixar.me/DVzT>.

partir daí, surgiram as primeiras universidades na Idade Média, ampliando o público leitor, o que fomentou ainda mais o comércio de livros.

As universidades também tiveram papel importante nessa época, porque foi responsável pela produção e disseminação de conhecimento, além de ter sido um terreno fértil para o aparecimento dos colégios invisíveis e dos primeiros periódicos científicos em 1665: o *Journal des Sçavans* (França), editado por Dennis de Sallo; e o *Philosophical Transactions* (Inglaterra), vinculado à *Royal Society of London* e sendo administrado por Henry Oldenburg. (MEADOWS, 1999).

Nesse ínterim, as sociedades científicas e os colégios invisíveis se estabeleceram. Colégios invisíveis, são definidos como um grupo de pesquisadores que trabalham juntos em determinado problema ou área de pesquisa, mesmo que geograficamente dispersos, trocam informações e mantêm contato frequente sobre o andamento das pesquisas que desenvolvem em comum. Esta expressão foi cunhada por Robert Boyle, um cientista inglês que viveu no século XVII. (ACOSTA-HOYOS, 1980; CRANE, 1988; LE COADIC, 2004, MEADOWS, 1999; MERTON, 1973b; MUELLER, 1994).

Em 1939, o físico John Bernal, responsável por cunhar o termo ‘comunicação científica’, afirma que o termo se relaciona com um amplo processo de geração, transferência e uso de informação científica. Por seu turno, Meadows (1999, p. vii) afirma que a comunicação científica se situa “no próprio coração da ciência. É para ela tão vital quanto à própria pesquisa, pois a esta não cabe reivindicar com legitimidade este nome enquanto não houver sido analisada e aceita pelos pares.”

Por sua vez, Garvey (1979) considera a comunicação científica como uma troca de informações entre os membros da comunidade científica, assim como as atividades associadas à produção, disseminação e uso da informação, inclusive desde o momento em que o cientista concebe uma ideia até que esta tenha aceitação entre os pares. Naturalmente, um dos objetivos da ciência é a divulgação e publicação de seus resultados, conforme corroboram diversos autores (GARVEY, 1979; MEADOWS, 1999; MULLER, 1995; SOLLA-PRICE, 1976; ZIMAN, 1984).

Relacionando a história da ciência e a comunicação científica, o trabalho seminal de Paul David (2008) intitulado “*The Historical Origins of Open Science*” mostra que, a partir da criação das sociedades científicas, dos colégios invisíveis e do

mecenato,<sup>12</sup> na ciência nem sempre havia compartilhamento, colaboração e troca de informações e conhecimento entre os cientistas. Antes disso, a ciência era considerada segredo.

Existiam pessoas que achavam estranho que eu tivesse publicado segredos que não deveriam ser revelados. É realmente certo que nossas descobertas sejam tão fortes para nós que o público não tem direito a elas, que não pertence a eles de alguma forma? Haveria muitas circunstâncias, onde nós éramos absolutamente mestres de nossas descobertas? Nós nos devemos primeiro à nossa pátria, mas devemos ao resto do mundo, àqueles que trabalham para aperfeiçoar as ciências e as artes, até mesmo para se enxergarem como cidadãos de todo o mundo. (RÉAUMUR, [19--?] *apud* BERNAL, 1939, p. 150-151).

Diante do exposto, Bernal (1939) reforça que um dos objetivos da ciência é a divulgação e publicação de seus resultados. O autor também explica que o desenvolvimento da ciência moderna coincidiu com a rejeição definitiva do segredo, afinal, há a obrigação moral de publicar as pesquisas e que está explicitamente ligada a outros elementos do *ethos científico*. Na mesma linha de pensamento, Oliveira (2009, p. 291) afirma que “[...] o principal objetivo de um cientista é produzir novos conhecimentos e que a divulgação de suas descobertas aumenta as oportunidades de circulação desses novos conhecimentos [...]” Para a autora é fundamental contribuir com esse processo porque atende ao caráter cumulativo da ciência e suas regras fundamentais para assegurar a originalidade e prioridade.

Robert King Merton, sociólogo norte-americano, considerado como o pai da Sociologia da Ciência e que nos apresenta o *ethos* da ciência: um complexo de valores e normas, uma espécie de prescrições, proscricções, preferências e permissões que são legitimadas com os valores institucionais, isto é, um estudo comparativo da estrutura institucional da ciência. Aliás, Merton (1968, p. 653) ressalta que “A meta institucional da ciência é a ampliação dos conhecimentos comprovados.” Assim, o autor propôs quatro imperativos:

- **Universalismo:** por seu caráter global e impessoal da ciência, nas sábias palavras de Pasteur: “a ciência não tem pátria”<sup>13</sup>. Para o autor, “Na medida em que uma sociedade é democrática, oferece oportunidade para o exercício de critérios universalistas na ciência.” (Merton, 1968, p. 657);

<sup>12</sup> A expressão ‘mecenato’ refere-se a prática do incentivo e patrocínio de artistas, literatos e cientistas de várias áreas por parte dos mecenas, pessoas ricas e poderosas tais como comerciantes, banqueiros, príncipes, condes e bispos como forma de obter reconhecimento e prestígio na sociedade. Essa prática surgiu no Império Romano e ainda existe atualmente no financiamento de produções culturais.

<sup>13</sup> Merton (1973) recorreu à frase de Pasteur em sua obra: “*Le savant a une patrie, la science n'en a pas*”.

- **Comunalismo:** também traduzido como 'comunismo', estabelece que o conhecimento é um bem comum e pertence a todos;
- **Desinteresse:** são elementos básicos da ciência que funcionam como controle da conduta dos cientistas e da manutenção da integridade científica, garantindo que este não seja movido pelos próprios interesses;
- **Ceticismo organizado:** estabelece que todas as ideias dos cientistas devem ser colocadas em suspensão, testadas e submetidas à avaliação da comunidade científica.

Como contra-normas a esses ideais, Merton (1968) apresentou ainda o **isolamento (sigilo e misantropia)**, o **particularismo**, os **interesses e o dogmatismo**. Anos mais tarde, esses princípios propostos por Merton foram retomados e discutidos por Ziman (2000), renomado físico inglês e epistemólogo da ciência, que resolveu formar o acrônimo CUDOS (Comunalismo, Universalismo, Desinteresse, Originalidade e Ceticismo organizado), ou seja, inverteu a ordem dos dois primeiros para formá-la, inspirando-se também no grego antigo: *ku'doç*, significa fama, glória e renome resultantes de um ato ou realização, além de ter acrescentado a 'originalidade' aos imperativos de Merton.

Contrastando os valores propostos por Merton, Ziman (2000) propôs a sigla PLACE (Proprietária, Local, Autoritária, Comissionada ou Encomendada e Especializada), baseada em seu próprio ideário e buscou distinguir a tecnologia da ciência. Ziman (2000) fazia parte de um contexto industrial em que imperava a mentalidade empresarial e taylorista, não obstante, o autor confere os atributos da ciência industrial como sendo:

[...] **proprietária, local, autoritária, encomendada e especializada.** [A ciência] produz conhecimento "proprietário" que não necessariamente é tornado público. Tem seu foco em problemas técnicos locais, ao invés de um entendimento geral. Pesquisadores industriais submetem-se à autoridade gerencial, em vez de agir como indivíduos. Sua pesquisa é encomendada, tendo em vista os fins práticos, em vez de empreendida em busca do conhecimento. São empregados como solucionadores de problemas especializados, em vez de colocar a sua criatividade pessoal (ZIMAN, 2000, p. 79, grifo nosso).

Dessa forma, as normas mertonianas pós-acadêmicas passaram a ser substituídas pelas da ciência industrial. (OLIVEIRA, 2011). Buscando ir além da polaridade entre Merton e Ziman, Etzkowitz e Leydesdorff (1997, 2000) comparam as relações entre a indústria, a universidade e o governo propondo o modelo da hélice tripla para explicar que a união entre essas três esferas seria capaz de conduzir à

inovação, ao desenvolvimento de novas tecnologias e à difusão de conhecimento, de modo que todos ganham e ainda podem cumprir sua responsabilidade social, dando um retorno para a sociedade e proporcionando melhorias para a vida das pessoas.

Na linha do conhecimento organizacional, por exemplo, Galbraith (1979), discutiu o desenvolvimento da ciência e tecnologia colocando o conhecimento em destaque como um fator estratégico de produção, enquanto Drucker (1968) propôs que a sociedade capitalista contemporânea passasse a ser chamada de 'sociedade do conhecimento'. Com a revolução tecnológica da informação e comunicação, cresceu extraordinariamente a quantidade e o volume de informação disponível, e, apesar disso, conforme Bresser-Pereira (2018, p. 18), “[...] o conhecimento político e organizacional permaneceu restrito à tecnoburocracia formada por altos executivos e financistas, e por intelectuais dos mais variados tipos.”

Nesse ínterim, as discussões sobre a capitalização do conhecimento ganharam força, assim, o capitalismo contemporâneo (DRUCKER, 1968), o capitalismo informacional da sociedade em rede (CASTELLS, 1999), o capitalismo digital (SCHILLER, 2011) ou capitalismo cognitivo (MOULIER BOUTANG, 2010), e o capitalismo tecnoburocrático (BRESSER-PEREIRA, 2018) – trouxeram uma maior importância para a informação e o conhecimento, reconhecendo seus papéis e as novas formas vislumbradas na cultura do *remix* e *mashup*, como a produção, criação, cocriação, *e-Science*, geração e apropriação. Segundo Albagli (2015) a ciência foi posta em nível estratégico, como símbolo de inovação e acirrou as disputas relacionadas ao acesso e uso dos conhecimentos entre os países.

Em suma, Le Coadic (2004) sintetiza os elementos que contribuíram para o crescimento da ciência na atualidade: a ampliação das indústrias de informação e do conhecimento que trouxeram novos produtos informacionais e processos, assim como novas atividades e empresas. Soma-se a isso, a profunda mudança na geografia das disciplinas científicas e em seus colégios invisíveis, a ramificação ou fusão de outras disciplinas, e, por fim, o avanço das tecnologias da informação e comunicação.

Avançando um pouco mais no tempo, em 2002, foi publicada a Declaração de Budapeste (Open Access Initiative – BOAI), considerada o marco inicial do Movimento de Acesso Aberto. Nesse mesmo ano, a Unesco também publicou a Declaração Universal sobre a diversidade cultural, estabelecendo que todas as pessoas teriam “[...] igualdade de acesso às expressões artísticas, ao conhecimento científico e tecnológico – inclusive em formato digital – e a possibilidade, para todas as culturas,



de estar presentes nos meios de expressão e de difusão [...]” (LESSA; ROSA; RODRIGUES, 2019, p. 1).

Outro marco histórico que merece ser mencionado no âmbito do *Big Data* e que possui relação com a comunicação científica e a colaboração entre os cientistas é o Projeto Genoma. Este projeto foi considerado um divisor de águas porque conseguiu reunir cientistas e indústria farmacêutica para trabalharem juntos no mapeamento do genoma humano. Outros grandes projetos colaborativos também merecem destaque, como os esforços do *Conseil Européen pour la Recherche Nucléaire* (CERN) no descobrimento da partícula subatômica bóson de Higgs; o Projeto Netuno do Observatório Oceânico de colaboração entre os Estados Unidos e Canadá; entre outros. (COSTA, 2017).

A partir desse cenário, o fluxo informacional da comunicação científica e as formas tradicionais de interação entre os cientistas entram em conflito com um novo paradigma baseado na abertura e em formação de redes colaborativas: a Ciência Aberta (*Open Science*), também denominada Ciência 2.0. Por se pautar em princípios como a abertura, colaboração, compartilhamento, troca de informações e conhecimentos de maneira livre e irrestrita, na Ciência Aberta a comunicação científica e o Movimento de Acesso Aberto estão cada vez mais inclinados a defender a abertura e o acesso às publicações científicas e aos dados de pesquisa, o que visa promover práticas científicas mais transparentes, abertas e que coadunam com a adoção de ações para a produção do conhecimento.

Nesse sentido, há um forte apelo para a mudança de comportamento dos pesquisadores a fim de adotar novas práticas, modificando o *modus operandi* de como os cientistas se relacionam e desenvolvem suas pesquisas atualmente, e, por conseguinte, afetam a construção do conhecimento, a comunicação e a divulgação científica. Para Tennant (2018), mais do que uma mudança de comportamento, trata-se de fazer a ciência do jeito certo.

Estabelecendo um paralelo entre as relações da Ciência Aberta com a *e-Science*, vale destacar o crescente volume e dilúvio de dados gerados em pesquisas e como tem sido demandado o uso de ferramentas computacionais em novas práticas de produção colaborativa do conhecimento científico (GRAY, 2009; SALES, 2014). Da mesma forma, a Ciência Aberta incentiva o desenvolvimento e uso de plataformas computacionais avançadas de apoio à pesquisa, considerando os direitos de acesso e uso dos dados assim como aqueles resultantes de pesquisa.

Também denominado como o quarto paradigma, a *e-Science* é uma forma de avançar as fronteiras do conhecimento por meio das tecnologias, influenciando diretamente em sua coleta, manipulação, análise e exibição dos dados. (GRAY, 2009). Nas palavras de Moura (2018, p. 7), a *e-Science* surgiu como uma oportunidade de se pensar os novos contextos e práticas científicas, é uma “[...] abordagem que se refere ao termo em inglês *enhance* (melhoria, aprimoramento, intensificação) e ao fortalecimento dos laboratórios e grupos de pesquisa como ambientes digitais colaborativos [...].” Para Ferreira (2018, p. 13, grifo nosso), a *e-Science* caracteriza-se pela

[...] necessidade de uma tecnologia adicional que suportasse o desenvolvimento das pesquisas e aliviasse o isolamento do pesquisador [...]; um composto de hardware, software e um preponderante cunho colaborativo. Esse foi o empurrão necessário para o desenvolvimento de **uma metodologia cujo uso intensivo de dados** fundamentava a busca para o tratamento dos insumos de um fazer científico abundante. Em seu bojo identifica-se a preocupação crescente pela **captura, curadoria e análise dos dados**, uma ajuda de peso para que o completo **ciclo da pesquisa** se concretize.

Portanto, a *e-Science* ou ciberciência pode ser definida como a infraestrutura que permite aos cientistas e pesquisadores acessar remotamente os dados científicos primários, movê-los, manipulá-los e extraí-los, possibilitando, assim, o surgimento de novas aplicações. Cabe enfatizar que a *e-Science* não é uma nova disciplina científica em si mesma, ela é um atalho para o conjunto de ferramentas e tecnologias necessárias para apoiar a ciência colaborativa e em rede. Exatamente por isso, toda a sua infraestrutura tem como objetivo capacitar os cientistas a fazer suas pesquisas de forma mais rápida, melhor e diferente. (HEY; HEY, 2006).

Por outro lado, Albagli, Appel e Maciel (2013) chamam a atenção para o fato de que **a *e-Science* nem sempre irá representar uma pesquisa colaborativa aberta, tendo em vista que nem sempre todas as ferramentas utilizadas são abertas e distribuídas**, assim, não pode ser caracterizada como *e-Science*, portanto, nesse quesito, **reside aí a diferença entre *e-Science* e *Open Science*** (Ciência Aberta).

Finalmente, merece destaque a iniciativa europeia Horizonte 2020, que pretende tornar públicas todas as publicações científicas resultantes de pesquisas financiadas pelo governo até 2020, exigindo dos cientistas que gerenciem os seus dados de pesquisa, elaborem Planos de Gestão de Dados e publiquem os resultados

das pesquisas em periódicos de acesso aberto ou em plataformas de acesso compatíveis<sup>14</sup>. Compreende-se que os dados de pesquisa sofreram influência da comunicação científica, do movimento de Acesso Aberto e da Ciência Aberta conforme o exposto nessa seção.

## 2.2 Ciência Aberta: definições

Nesta seção serão apresentadas algumas definições sobre a Ciência Aberta com a finalidade de contextualizar o tema (Gestão de Dados de Pesquisa), e assim deixar clara as relações existentes entre cada um dos assuntos. Além disso, para que se possa compreender melhor alguns conceitos, percepções e perguntas realizadas durante o processo de pesquisa, faz-se necessário situar o leitor a respeito das fronteiras entre a gestão dos dados científicos e a Ciência Aberta, e, em que ponto eles se cruzam, inclusive, em quais são divergentes (dados de pesquisa restritos e dados de pesquisa abertos). Deste modo, será preciso divergir para convergir.

Inicialmente, a Ciência Aberta “É a ideia de que os conhecimentos científicos de todos os tipos deveriam ser compartilhados abertamente tão logo quanto possível no processo de descoberta.” (NIELSEN, 2011b). Por conseguinte, em um dos cursos ofertados pelo projeto europeu Foster, a Ciência Aberta é conceituada como uma prática que permite aos outros colaborar e contribuir com os dados de investigação, notas laboratoriais e em diversos processos de pesquisa, contribuindo para a livre disponibilização e sob condições de reuso, redistribuição e reprodução da pesquisa, dos dados e métodos incluídos no estudo.(BEZJAK *et al.*, 2018, online).

Para a *Organisation for Economic Cooperation and Development* - OECD (2015, p. 9), a sua definição de Ciência Aberta incorpora a menção aos resultados de pesquisas financiadas com dinheiro público e características ligadas ao uso das TIC's:

A Ciência Aberta refere-se geralmente aos esforços para tornar os resultados da investigação financiada com recursos públicos mais amplamente acessíveis em formato digital à comunidade científica, ao setor empresarial ou à sociedade em geral. A Ciência Aberta é o encontro entre a antiga tradição de abertura na ciência e as ferramentas das tecnologias da informação e da comunicação (TIC) que remodelaram a empresa científica e exigem um olhar crítico por parte dos gestores políticos que procuram promover a investigação a longo prazo e a inovação.

---

<sup>14</sup> Matéria da Folha de São Paulo (Ciência europeia terá de ser publicada em revistas de acesso livre em 2020), disponível em: <https://www1.folha.uol.com.br/ciencia/2018/10/ciencia-europeia-tera-de-ser-publicada-em-revistas-de-acesso-livre-em-2020.shtml>.

No glossário do acesso aberto da Fiocruz (2019, online), a Ciência Aberta define-se como um “Conceito amplo que abarca diversas práticas do fazer científico, desde o acesso aberto à literatura científica e aos dados abertos de pesquisa, passando pela ciência cidadã, recursos educacionais abertos, software e hardware abertos, entre outras.” A entidade esclarece ainda que o termo se refere a um modelo metodológico em consonância com a filosofia da cultura digital, da colaboração e por meio do compartilhamento das práticas e dos conteúdos científicos.

Assim como o projeto Foster, outra iniciativa financiada pela comissão europeia, *A Open Responsible Research and Innovation to further Outstanding Knowledge – Orion* (2018, online) esclarece que a Ciência Aberta é um termo genérico que engloba uma multiplicidade de aspectos: acesso aberto a publicações, dados de pesquisa aberta, reprodutibilidade e integridade científica, avaliação de pesquisa e engajamento público, entre outros.

A Ciência Aberta é o movimento para tornar a pesquisa científica, os dados e sua disseminação disponíveis para qualquer membro de uma determinada sociedade, de profissionais a cidadãos. Este movimento tem repercussões nos princípios do desenvolvimento científico e no acesso público, incluindo práticas como a publicação de investigação aberta e a promoção do acesso aberto, com o objetivo principal de facilitar a publicação e a comunicação dos conhecimentos científicos. Desde o desenvolvimento até à difusão dos conhecimentos, vários conceitos pertencem ao conceito genérico de "Ciência Aberta".

Conforme o exposto baseado nos conceitos apresentados, a Ciência Aberta é desenvolvida e disseminada de forma que todos possam colaborar, desde pesquisadores ao público leigo, permitindo-lhes contribuir e acrescentar conhecimento à pesquisa, tais como: dados, metodologias, resultados e protocolos, disponibilizando-os gratuitamente em diferentes estágios do processo de investigação. Além disso, antes de compartilhar os dados, é preciso estabelecer as condições em que esses dados serão distribuídos a fim de divulgar e ampliar o impacto da pesquisa, mas também garantir as questões referentes aos direitos autorais e de propriedade intelectual.

Do mesmo modo, Murray-Rust *et al.* (2010, online) reforçam, em seus *Panton Principles*<sup>15</sup>, que, “para a ciência funcionar de maneira efetiva e para que a sociedade se beneficie dos esforços científicos, é crucial que os dados científicos sejam abertos.” Além disso, a Ciência Aberta incentiva essa abertura e acesso aos dados científicos

---

<sup>15</sup> Panton Principles: <https://www.pantonprinciples.org/>.

para a verificação e reprodutibilidade da pesquisa em prol do progresso da ciência, da maior velocidade na circulação das informações a fim de produzir uma ciência de qualidade. Complementando essa visão, Michael Nielsen (2011a, online), em palestra intitulada “Open Science Now!”, proferida durante a TEDx, defende que:

Nossa sociedade encara uma questão fundamental: que tipo de conhecimento esperamos e incentivamos que seja compartilhado? Continuaremos com as práticas do passado? Ou iremos abraçar as novas formas de compartilhar, que levarão a novos métodos para a solução de problemas e a uma aceleração no processo científico de maneira global? Minha esperança é que iremos abraçar a Ciência Aberta e realmente aproveitar essa oportunidade de reinventar a descoberta em si.

Em um segmento que abrange a comunicação científica e a *Open Scholarship*, Kraker *et al.* (2016), idealizadores dos princípios de Viena, explicam que a Ciência Aberta tem potencial para lidar com os déficits do sistema de comunicação científica e melhorar a produção e a disseminação do conhecimento acadêmico. Sobre o conceito ‘*Open Scholarship*’<sup>16</sup>, termo com tradução ainda indefinida no português, define-se como uma forma de “[...] tornar a pesquisa visível e compartilhar conhecimento. Ela abrange acesso aberto, dados abertos, recursos educacionais abertos e todas as outras formas de abertura no ambiente acadêmico e de pesquisa.” (GRIFFITH UNIVERSITY, 2019, online, tradução nossa).

*Open Scholarship*, por muitas vezes, está associado à Ciência Aberta, inclusive chegando a se confundir ou não deixar claro os limites entre um e outro, conforme explicitado por Tennant *et al.* (2019, p. 1, tradução nossa):

A pesquisa acadêmica pode ser idealizada como um processo inspirador para avançar nosso conhecimento coletivo em benefício de toda a humanidade. Entretanto, as práticas de pesquisa atuais muitas vezes lutam com uma série de tensões, em parte devido ao fato de que esse ideal coletivo (ou “*commons*”) entra em conflito com o sistema competitivo no qual a maioria dos estudiosos trabalha, e por outro lado porque grande parte da infraestrutura do mundo acadêmico está se tornando amplamente digital. O que é amplamente denominado como *Open Scholarship* é uma tentativa de realinhar práticas modernas de pesquisa com esse ideal. Nós não propomos uma definição de *Open Scholarship*, mas, reconhecemos que é um termo holístico que abrange muitas disciplinas, práticas e princípios, às vezes também referidos como *Open Science* [Ciência Aberta] ou *Open Research* [Pesquisa Aberta]. Nós escolhemos o termo [*Open Scholarship*] por ser o mais inclusivo.

---

<sup>16</sup> Griffith University (Open Scholarship): <https://www.griffith.edu.au/library/research-publishing/open-scholarship>.

A Ciência Aberta tem uma miríade de conceitos e escolas de pensamento (conhecimento como um bem público; pragmática; infraestrutura tecnológica; engajamento público; sistemas de valores), perpassando por várias vertentes e práticas. Comumente, é denominada como um termo guarda-chuva por abrigar uma série de outros assuntos que estão imbricados e intimamente relacionados a ela. (ALBAGLI, 2015; FECHER; FRIESIKE, 2014).

Retomando os conceitos de Ciência Aberta, Fecher e Friesike (2014) investigaram o cenário da *Open Science* a fim de criar uma tipologia que pudesse abarcar todas as atividades e conceitos sobre o assunto. Então, os autores identificaram cinco escolas de pensamento:

- a) Conhecimento como um bem público: Parte do pressuposto de que o acesso ao conhecimento é desigual, assim, garantir o direito humano de ter acesso ao conhecimento, de forma que esteja livremente disponível para todos;
- b) Pragmática: Pressupõe que a colaboração aberta é mais eficiente para a criação de conhecimento e acelera o processo de descoberta, portanto, abrir esses processos é essencial para que se tenha acesso aos conhecimentos tão rápido quanto possível;
- c) Infraestrutura eletrônica: A oferta de infraestrutura e ferramentas tecnológicas são fundamentais para a colaboração, assim, criar plataformas, ferramentas, redes e serviços abertos para os cientistas é um fator determinante;
- d) Engajamento público: Por ser considerada uma atividade pública, também deve estar disponível para todos, e, nesse aspecto, o engajamento dos cidadãos (público leigo) em pesquisas científicas também é primordial a fim de agregar valor às contribuições dos pesquisadores, além de garantir o envolvimento e representatividade dessas pessoas na Ciência Aberta;
- e) Sistemas de valores: Parte do princípio de que é necessário criar novas métricas, incentivos e outras formas mais justas de medir o impacto das pesquisas, portanto, o desenvolvimento de métricas alternativas também incentiva o estabelecimento de sistemas de avaliação mais abertos, transparentes, justos e inclusivos.

Para Albagli (2015), a Ciência Aberta pode ser considerada como um 'movimento de movimentos', pois vários outros movimentos - como o de Acesso

Aberto - estão inseridos nela. Inclusive, ainda de acordo com a autora, a Ciência Aberta pode ser discutida como um movimento social; além disso, esclarece que não há um consenso sobre o conceito, portanto, encontra-se em franca expansão. Assim, a Ciência Aberta sustenta-se em quatro pilares básicos, a saber: Dados, Códigos, Publicações e Revisões. (Figura 5).

Figura 5 - Pilares da Ciência Aberta



Fonte: Foster, Curso What is Open Science? (2019).

Baseada nesses pilares, a Ciência Aberta desdobra-se em um conjunto composto por elementos heterogêneos, variáveis e associados ao seu escopo principal, conforme o ilustrado na figura 6:

Figura 6 - Princípios da Ciência Aberta



Fonte: Elaborado pela autora (2019).

Tal como um caleidoscópio, engloba uma série de práticas e processos, tais como (WIKIVERSIDADE, 2017):

- a) Dados de Pesquisa Abertos: Dados oriundos de informações coletadas e criadas em pesquisas científicas, tais como planilhas, transcrições e áudios de entrevistas, entre muitos outros tipos de materiais que são disponibilizados livremente para outros usarem e reutilizarem;
- b) Reprodutibilidade da Pesquisa e Análise de dados Aberta: Experiência científica que permite reproduzir experimentos e procedimentos científicos a fim de analisar e verificar os resultados de pesquisas como forma de combater fraudes, garantir a integridade e a evolução do conhecimento científico, a exemplo do projeto *Brazilian Reproducibility Initiative*<sup>17</sup> e o a iniciativa

<sup>17</sup> Disponível em: <https://www.reprodutibilidade.bio.br/home>.



- coordenada pelo pesquisador Brian Nosek denominada *Estimating the Reproducibility of Psychological Science*<sup>18</sup>;
- c) Métodos, protocolos e Instrumentos científicos abertos, Códigos, *Softwares* de Pesquisa Abertos e *Open Source*: Incentivo ao uso de softwares de código-fonte abertos que permitem o livre uso, adaptação e compartilhamento em pesquisas científicas;
  - d) *Hardware* abertos: Artefatos, máquinas e outros objetos físicos e dispositivos tangíveis que são disponibilizados ao público para que possam ser usados para construir, modificar, adaptar e distribuir estes artefatos e máquinas, a exemplo do uso de Arduino, Impressoras 3D e outros equipamentos disponíveis em *Makerspaces / Hackerspaces*. Outro exemplo a ser mencionado nesse caso é o pesquisador André Maia Chagas<sup>19</sup>, desenvolvedor de equipamentos de laboratório científico a baixo custo;
  - e) Acesso Aberto às Publicações Científicas: Acesso livre e irrestrito à todas as publicações científicas na internet;
  - f) Avaliação Aberta: Modelo de avaliação por pares abertas com a finalidade de aumentar a transparência, a eficiência e a responsabilidade do processo de revisão de artigos e outras publicações científicas;
  - g) Cadernos de Laboratório Abertos: Santos (2016) afirma que cadernos de laboratório eletrônicos disponíveis na internet que divulgam na íntegra os registros de pesquisas individuais ou de um grupo por meio de licenças livres que permitam o acesso, uso e reuso para todos;
  - h) Recursos Educacionais Abertos: para a United Nations Educational, Scientific and Cultural Organization - Unesco (2018), são materiais de ensino, aprendizagem e pesquisa em quaisquer tipos de suporte ou mídia disponibilizados em domínio público ou licenciados de forma aberta que permite a utilização e adaptação por terceiros. Os materiais variam desde cursos inteiros ou em parte, módulos, livros didáticos, artigos, vídeos, teste, softwares e muitos outros elementos;

---

<sup>18</sup> Disponível em: <https://osf.io/ezum7/>.

<sup>19</sup> Palestra de André Maia Chagas sobre equipamentos científicos de baixo custo: [https://www.youtube.com/watch?v=yLed0gx\\_WKw](https://www.youtube.com/watch?v=yLed0gx_WKw). Ver também o artigo do pesquisador que descreve a experiência de produção de um microscópio com impressora 3D: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2002702>.

- i) Colaboração e Inovação aberta: Processo em que as empresas avaliam ideias internas e externas como forma de avançar no desenvolvimento de tecnologias, produtos e serviços;
- j) *Advocacy*: Incentivo e mobilização social em torno da Ciência Aberta e suas práticas;
- k) Ciência Cidadã e Engajamento Público com a Ciência: Incentivo à participação do público leigo permitindo que seja dada a sua contribuição em estudos científicos colaborativos;
- l) Políticas governamentais e mandatos institucionais de promoção à Ciência Aberta: Políticas instituídas pelos países a fim de promover a Ciência Aberta, a exemplo do lançamento do maior programa de pesquisa e inovação a União Europeia, a *Horizon 2020*<sup>20</sup>; a *Alliance of Science Organisations in Germany*<sup>21</sup> na Alemanha; na França a *Le Plan national pour la science ouverte*<sup>22</sup>; a Política Nacional de Ciência Aberta de Portugal<sup>23</sup>; e por fim o *National Open Science*<sup>24</sup> na Holanda. Outros países apresentam políticas de compartilhamento de dados, tais como a *National Data Sharing and Accessibility Policy*<sup>25</sup> da Índia; Austrália com a *Australian Government Public Data Policy Statement* e *Australian Research Council*<sup>26</sup>; os Estados Unidos da América com a *Open Government Initiative*<sup>27</sup>, *Open Data Policy*<sup>28</sup> e *NIH Data Sharing Policy and Implementation Guidance*<sup>29</sup>; e o *Common principles on data policy*<sup>30</sup> no Reino Unido.

Também podemos considerar os mapas de conhecimento aberto um elemento adicional além dos que foram mencionados acima, afinal, a ideia principal desses mapas é a construção e representação do conhecimento em várias instâncias. Apresenta uma estrutura visual do conhecimento que pode ser encontrada dentro de

<sup>20</sup> Disponível em: <https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020>.

<sup>21</sup> Disponível em: <https://www.mpg.de/alliance>.

<sup>22</sup> Disponível em: <http://www.enseignementsup-recherche.gouv.fr/cid132529/le-plan-national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-a-tous-sans-entree-sans-delai-sans-paiement.html>.

<sup>23</sup> Disponível em: <https://www.ciencia-aberta.pt/>.

<sup>24</sup> Disponível em: <https://www.openscience.nl/en/national-platform-open-science/national-plan-open-science>.

<sup>25</sup> Disponível em: <http://www.dst.gov.in/national-data-sharing-and-accessibility-policy-0>.

<sup>26</sup> Disponível em: <https://www.pmc.gov.au/resource-centre/public-data/australian-government-public-data-policy-statement>.

<sup>27</sup> Disponível em: <https://open.usa.gov/>.

<sup>28</sup> Disponível em: <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf>.

<sup>29</sup> Disponível em: [https://grants.nih.gov/grants/policy/data\\_sharing/data\\_sharing\\_guidance.htm](https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm).

<sup>30</sup> Disponível em: <https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>.

uma organização, grupo, fontes, pessoas ou domínio de conhecimento. (AMERICAN PRODUCTIVITY & QUALITY CENTER, 2019; SHAREWEB LEARNING & NETWORK, 2019; WEXLER, 2001).

Este tipo de mapa é comumente aplicado em organizações nos processos de inventariar o conhecimento existente a fim de elaborar estratégias para desenvolver e melhorar processos de compartilhamento de conhecimento entre os membros. Já os mapas de conhecimento abertos estão mais voltados à análise de domínio de conhecimento, a exemplo da iniciativa *Open Knowledge Maps*<sup>31</sup> que visa fornecer uma visualização em larga escala de mapas de conhecimento abertos, interativos e interligados com todas as áreas do conhecimento e campos de pesquisa. (KRAKER *et al.*, 2017).

Outra forma de ilustrar as ramificações de assuntos relacionados à Ciência Aberta é a Taxonomia da Ciência Aberta (Open Science) elaborada pelo projeto Foster, conforme a Figura 7. A referida figura expressa visualmente as relações da Ciência Aberta com outros temas relevantes que estão presentes no cotidiano acadêmico e da comunicação científica.

Albagli (2015, p. 21-22) afirma que “Os esforços da Ciência Aberta envolvem instâncias de ação e decisão diferenciadas, internas e externas à ciência, [...] Trata-se de [...] governança informacional – envolvendo formas de gestão e resolução de conflitos e de poder, que mantêm suas especificidades.” Essas instâncias mencionadas pela autora são assim apresentadas:

- Micro: Pesquisadores, Grupos de pesquisa, Gestores públicos e privados;
- Meso: Instituições, Organizações, Empresas, Comunidades;
- Macro: Marcos regulatórios e Políticas Públicas.

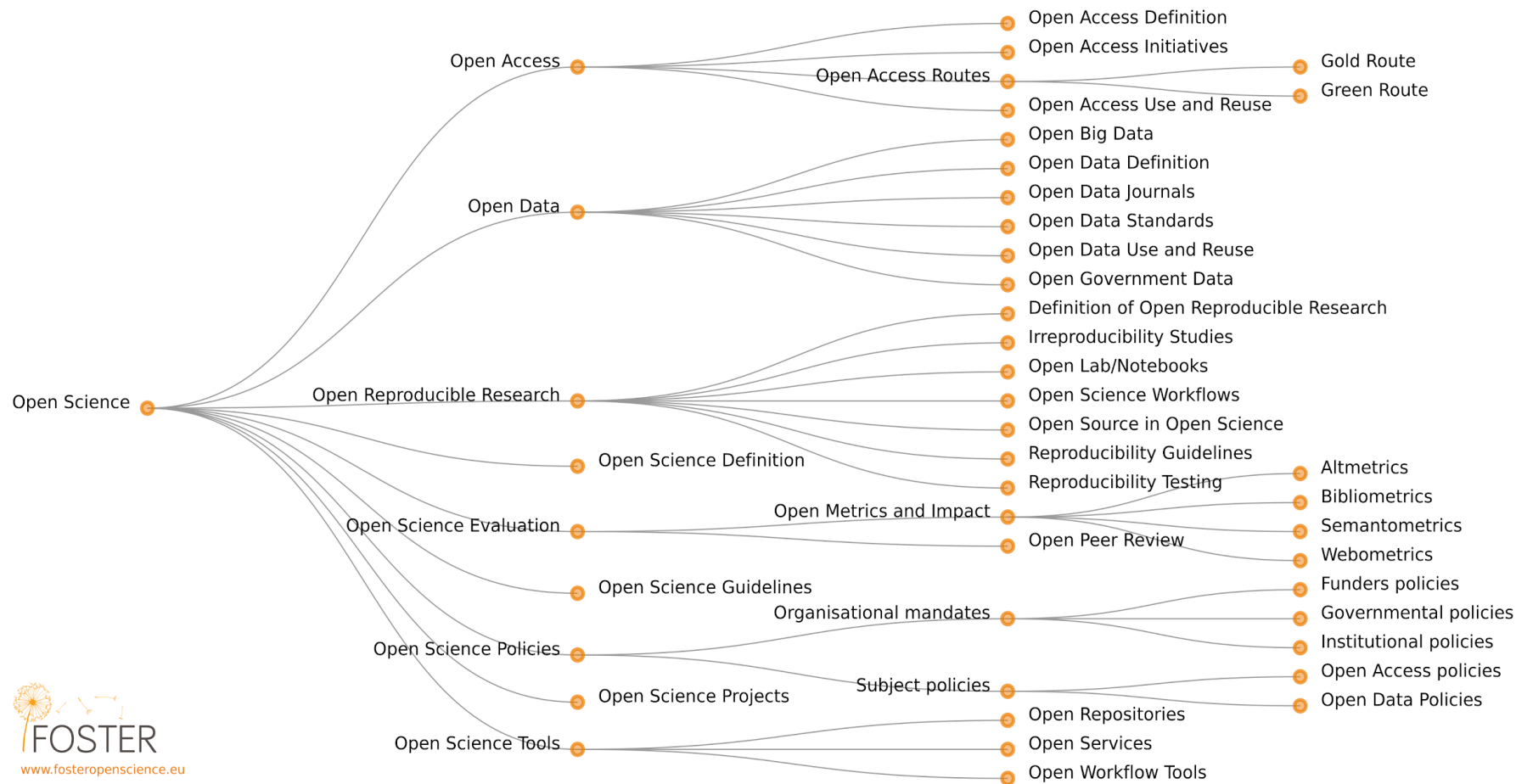
Albagli (2015) alerta que essas instâncias se influenciam de forma direta e indireta, apesar de serem desconectadas. Nesse sentido, a Ciência Aberta traz em seu bojo a proposta de mobilização desses diversos atores sociais heterogêneos que produzem conhecimento e buscam incentivá-los a compartilhar, contribuir e adotar essas práticas informacionais presentes nos regimes de informação.

---

<sup>31</sup> Open Knowledge Maps: <https://openknowledgemaps.org/>.

Figura 7 - Taxonomia da Ciência Aberta

# Open Science Taxonomy



Fonte: Foster (2018).

Além disso, a Ciência Aberta foi profundamente influenciada pela cultura livre digital, tem raízes no movimento *hacker* e no software livre (ALBAGLI, 2015; SODERBERG, 2008). Prova disso são alguns conceitos ligados a ela, como é o caso de ‘aberto’ ou ‘abertura’, definido pela *Open Knowledge Foundation* (2019)<sup>32</sup> como “O conhecimento [...] livre para acessá-lo, utilizá-lo, modificá-lo, e compartilhá-lo - restrito, no máximo, a medidas que preservam a proveniência e abertura.” Em seu site, a *Open Knowledge Foundation* explica que essa definição traz influências da *Open Source Definition* (Definição de Código Aberto), da *Definition of Free Cultural Works* (Definição de Obras Culturais Livres) e da *Debian Free Software Guidelines*.

Esta definição já tem uma segunda versão no site da entidade e esclarece que “A Open Definition torna preciso o significado de “aberto” no que diz respeito ao conhecimento, promovendo um bem comum robusto no qual qualquer um pode participar, e a interoperabilidade é maximizada.” (OPEN KNOWLEDGE FOUNDATION, 2019).

Internacionalmente, advogar pela abertura tem se estendido para além da Ciência Aberta, se expandindo para todas as esferas da sociedade, no campo das artes, educação, pesquisa, ciência e tecnologia, negócios, avaliações, governo e patrimônio cultural.

A figura 8 representa a abertura ampla e em diversos segmentos, tais como nas artes e patrimônio cultural, a exemplo da iniciativa OpenGLAM<sup>33</sup> mantida pela Open Knowledge que promove o acesso livre e aberto ao patrimônio cultural digital mantido por galerias, bibliotecas, arquivos e museus.

---

<sup>32</sup> Ver: <http://opendefinition.org/od/2.0/pt-br/>.

<sup>33</sup> OpenGlam: <https://openglam.org/>.

Figura 8 - Concepções sobre o conceito Aberto



Fonte: Baseado no Gráfico de Katja Mayer (2018), CC BY 4.0.

Apesar de existir uma definição para ‘abertura’, este também é um termo ainda em construção, assim como a palavra ‘apropriação’, ambas têm distintas formas de entendimento e conceituação (ALBAGLI, 2015). Apropriação pode ser entendida como a socialização da informação e do conhecimento, talvez, seja tomada como a ação de se apossar de um bem patrimonial, ou ainda o ato de tomar para si a responsabilidade por alguma coisa.

Por fim, após terem sido apresentados alguns conceitos, práticas e processos da Ciência Aberta, a exemplo da ciência como um bem público, a ciência orientada aos dados e que precisa dispor de uma ciberinfraestrutura, a Ciência Aberta como movimento, um ideário e um conjunto de práticas intrinsecamente ligadas ao comportamento e valores culturais dos cientistas, iremos a seguir adentrar no universos dos dados de pesquisa a partir da discussão sobre as várias definições sobre o que são dados.

### 2.3 Dados: explorando definições dos elementos em forma bruta

Há na literatura científica diversas definições para dados. Não pretendemos esgotar todas as definições existentes, mas sim traçar uma dinâmica de apresentação para alguns desses conceitos presentes na literatura científica, tendo em vista que o conhecimento sobre o assunto é vasto e inesgotável, aliás, trata-se de tema bem debatido e enseja profundas discussões na área de Biblioteconomia e Ciência da Informação, apesar de não ser este o foco desta pesquisa.

Inicialmente, Zins (2007) empreendeu um estudo crítico Delphi sobre mapas de conhecimento voltado para a área de Ciência da Informação, e, em um de seus artigos apresenta uma série de concepções a respeito das definições para 'dados', 'informação', 'conhecimento' e 'mensagem', resultando em cerca de 130 definições desses conceitos, demonstrando quão difícil é delinear qualquer descrição ou elucidação sobre estes termos, especialmente sobre o que são 'dados'. Para o autor, "Dados são frequentemente concebidos como matéria-prima para a informação, que é frequentemente concebida como matéria-prima para o conhecimento. Conhecimento é a construção de ordem superior." (ZINS, 2011, p. 155). Em semelhante perspectiva, os "[...] dados são sinais que não foram processados, correlacionados, integrados, avaliados ou interpretados de qualquer forma, e, por sua vez, representam a matéria-prima a ser utilizada na produção de informações." (RUSSO, 2010, p. 15).

Para Nielsen e Hjørland (2014), a palavra 'Dado' é bastante usada na Ciência da Informação, a exemplo de 'banco de dados', 'bases de dados', 'mineração de dados' etc. Além disso, pode ainda estar ligada à ciência dos dados e processamento de dados, ou ainda, pode aparecer apresentada na hierarquia 'dado-informação-conhecimento', nesse contexto, inserida como matéria-prima a ser processada, gerando informação para a aquisição do conhecimento.

Miranda (1999) considera que os dados são um conjunto de registros qualitativos ou quantitativos e se for organizado, agrupado, categorizado e padronizado adequadamente transforma-se em informação. Na ótica de Setzer (1999, não paginado) dado é

[...] uma sequência de símbolos quantificados ou quantificáveis. [...] um texto é um dado. [...] as letras são símbolos quantificados, já que o alfabeto por si só constitui uma base numérica. Também são dados imagens, sons e animação, pois todos podem ser quantificados [...].

O autor alerta que mesmo que os dados sejam ininteligíveis para o leitor, qualquer texto ainda se constitui em um dado ou uma sequência de dados. Dando seguimento a esta linha de pensamento, Padoveze (2000, p. 43) explicita que a “informação é o dado que foi processado e armazenado de forma compreensível para seu receptor e que apresenta valor real percebido para suas decisões correntes ou prospectivas”. Sob esse prisma, consideramos que é através da aplicação do conhecimento humano, do processamento dos dados até à sua contextualização dentro de um determinado cenário é que possibilitará a transformação desses dados em informação, e por sua vez, agregará valor e poderá ser utilizada por empresas e pessoas em processos decisórios/tomadas de decisão.

Na perspectiva de Davenport (1998), os dados são elementos brutos, sem significado, desvinculados da realidade e que correspondem a observações sobre o mundo. Para ele, há uma relação intrínseca entre o dado e a informação: os dados são observações simples sobre o mundo, podem ser facilmente estruturados, quantificados e transferíveis. Portanto, para o autor, o conjunto de dados dotados de propósito e relevância que necessitam de análise e mediação para obter significado é considerado como informação. Não há produção de informação sem dados, logo, isso reflete em sua quantidade e a necessidade de ser tratada. Na esteira desse pensamento, esclarecemos que os dados são elementos em forma bruta, e, significa que foram coletados, mas ainda precisam passar por análise e processamento.

Corroborando essa visão, Oliveira (2002, p. 51) afirma que “[...] dado é qualquer elemento identificado em sua forma bruta que, por si só, não conduz a uma compreensão de determinado fato ou situação”. Nesse sentido, o autor esclarece que os dados sozinhos não levam a compreensão porque estes estão fora de um contexto, sem a devida explicação ou um resumo sobre o que se trata, algo que os torne compreensíveis e interpretáveis, portanto, de fato, não fará sentido para aqueles que tentaram interpretá-los. Por exemplo: uma planilha com muitos valores diferentes, em que não constam as informações complementares da origem desses números e sem qualquer explanação para entender o que eles significam. Portanto, além de os dados serem contextualizados, Brookes (1980) reforça que o seu uso e o processamento se transformam em informação, elemento principal para promover as devidas transformações nas estruturas subjetivas e objetivas de um indivíduo.



Nessa linha de pensamento, comparamos esta situação à ideia da transformação da “informação” em coisa conforme Buckland (1991), assim representada em um determinado suporte informacional, e ainda, da teoria da informação como um processo, referente à ação de informar ou comunicar algo. O autor também apresenta em seus estudos a informação como algo capaz de propiciar a construção de conhecimento, reduzindo as incertezas, e, nesse aspecto, estas teorias da informação de Buckland se enquadram bem ao panorama dos dados.

Afinal, é a partir desse processamento que os dados se transformam em informação e instrumentos para tomadas de decisão em diferentes níveis sociais, tornando-se inclusive em objetos de estudo nas pesquisas científicas; ou a depender do contexto e tipologia, podem se converter em produtos e serviços para os *prosumers*, termo cunhado originalmente por Toffler (1995) e que define um novo perfil de interação da sociedade de consumo, um neologismo para descrever o usuário que é produtor e também consumidor. O termo ‘*prosumers*’ (prossumidores) foi traduzido por Armando Malheiro da Silva (2009, p. 98), é envidado na perspectiva do paradigma pós-custodial e baseado na mediação informacional cumulativa, o qual propala que:

À medida que se inovam e expandem mais as possibilidades tecnológicas (novas soluções e produtos) o papel do “prossumidor” (**produtor e usuário**) cresce enormemente, desenvolvendo um tipo de mediação cumulativa que pode abranger a de designer e de programador, e que produz efeitos e é condicionada através da activa participação em comunidades que agregam interagentes idênticos ou parecidos. (SILVA, 2009, p. 98, grifo nosso).

Sendo assim, o paradigma pós-custodial, informacional e científico de Silva (2009, 2011) apregoa a necessidade de uma mudança de postura epistemológica com foco nas abordagens instrumentais, práticas, normativas e descritivas dos documentos-artefatos, assim como na *Open Science* se exige uma mudança de postura tanto comportamental, cultural, como no uso de tecnologias que facilitem a colaboração e o compartilhamento de dados entre os pares e aos demais interessados da sociedade. Para o autor, trata-se de um dos desafios da era da informação em que a internet assume a função de ferramenta de construção de projetos desenvolvidos a partir de diferentes dimensões. Silva (2009, 2011) reforça que a Ciência da Informação brota deste novo paradigma emergente, repleto de fragilidades e potencialidades, em sintonia com as Ciências Sociais.

Retomando a discussão sobre dados, de acordo com a Organização das Nações Unidas (ONU), “Dados são a alma da tomada de decisões e a matéria-prima

para a prestação de contas. É quase impossível a concepção, o acompanhamento e a avaliação de políticas eficazes sem dados de alta qualidade que fornecem as informações corretas sobre as coisas certas no momento certo”. (UNITED NATIONS ORGANIZATION, 2014, p. 2, tradução nossa).

Dados abertos, de acordo com a *Open Knowledge International Foundation* (2017), podem ser usados livremente, reutilizados e redistribuídos por qualquer pessoa (física ou jurídica), exigindo, no máximo, a atribuição da fonte e o compartilhamento pelas mesmas regras. Em suma, os pontos mais importantes que devem ser levados em conta são a disponibilidade e acesso; reutilização e redistribuição; participação e a interoperabilidade dos dados.

Dados abertos conectados (*Linked Open Data*) emergem como um novo conceito, permitindo que os dados fiquem disponíveis em formatos legíveis por máquina, possibilitando a conexão com fontes diferentes, além de seguir padrões de representação e assim como dotá-los de elementos semânticos e requisitos básicos de qualidade, como o estabelecimento de vocabulários controlados, ontologias, entre outros recursos de descrição e representação, isto é, a partir de um dado é possível acessar outros dados e informações relacionadas a um dado específico. (ÁVILA, 2015; BAUER; KALTENBÖCH, 2012; BERNERS-LEE, 2006; ISOTANI; BITTENCOURT, 2015; WORLD WIDE WEB CONSORTIUM, 2017).

Auer, Bryl e Tramp (2014, p. 6), apresentam algumas características de justaposição dos conceitos atribuídos a Dados Conectados (*Linked Data*), *Linked Open Data* (Dados Abertos Conectados) e *Open Data* (Dados Abertos) baseados em sua representação e grau de abertura, conforme o Quadro 2 abaixo:

Quadro 2 - Justaposição dos conceitos de Linked Data (Dados conectados), Linked Open Data (Dados abertos conectados) e Open Data (Dados abertos)

<b>Representação e Grau de Abertura</b>	<b>Possivelmente fechado</b>	<b>Aberto</b>
Modelo de dados estruturados (Exemplo: XML, CSV, SQL etc.)	Dados	Dados Abertos
Modelo de dados RDF (Publicado como Dados conectados)	Dados Conectados	Dados Abertos Conectados

Fonte: Adaptado de Auer; Bryl; Tramp, 2014, p. 6.

Nesse sentido, para Marcondes e Sayão (2001, p. 24) “A convergência e o uso integrado das tecnologias de comunicação, de computação e de conteúdos em

formato digital, [...] tem contribuído nos anos recentes para criar um novo ambiente de acesso, disseminação, cooperação e promoção do conhecimento em uma escala global [...].”

Sobre os **dados de pesquisa** (*research data*), uma das clássicas definições foi cunhada pela *Organisation for Economic Cooperation and Development* - OECD (2007, p. 13, tradução nossa):

Dados de pesquisa são como registros factuais (escores numéricos, registros textuais, imagens e sons) utilizados como fontes primárias para pesquisa científica e que são comumente aceitos na comunidade científica como necessário para validar resultados da pesquisa. Um conjunto de dados de pesquisa constitui uma representação sistemática e parcial do assunto que está sendo investigado.

Entretanto, esse é um conceito que já vem sendo bastante discutido na literatura acadêmica, e alguns autores não concordam totalmente com a definição supracitada. Há uma variedade de conceitos para definir o que são dados, e notadamente são distintos em diversas disciplinas e contextos.

Dados de pesquisa, no contexto da gestão da informação, podem se apresentar como objetos digitais ou em outras formas materiais constituintes de um documento, e como tal, carrega uma estrutura e aplicação dentro de um contexto. Prova desse aspecto é a corrente de pensamento que remonta o raciocínio de Otlet (1934), Briet (1951), Day (2001) e Frohmann (2009), colocando em xeque a noção de documento e sua representação, inclusive para a ideia de que o documento em si não existe, conforme a corrente neodocumentalista. Em linhas gerais, nas palavras de Saldanha (2012, p. 9), “o documento se traveste de novos significantes – textos, discursos, registro e, principalmente, informação.”

São muitas as formas em que os dados se apresentam: números, textos, imagens, sons, vídeos, amostras biológicas, coleções físicas e digitais, simulações, códigos computacionais, sequências genéticas etc. Podem ainda se diferenciar entre dados de pesquisa e outros materiais adicionais associados, os quais são necessários para a compreensão dos dados em si, como um arquivo ‘*Read me*’ (Leia-me), que acompanha muitos dos softwares que costumamos instalar em nossos computadores. Este é um bom exemplo que faz parte do cotidiano de qualquer pessoa.

O *National Institutes of Health* (NIH)<sup>34</sup> afirma que os dados finais de pesquisa são materiais factuais registrados comumente e aceitos pela comunidade científica, sendo necessários para documentar, apoiar e validar os resultados de um estudo. O NIH esclarece ainda que, para muitos estudos, os dados de pesquisa serão conjuntos de dados computadorizados, e, para outras áreas, os conjuntos de dados podem ser compostos por dados brutos e variáveis derivadas que seriam descritos posteriormente em um plano de gestão de dados. A partir desse viés, percebemos que aqui dados científicos são chamados ‘dados finais de pesquisa’, isto é, deixando ainda mais claro que se tratam de dados coletados, processados e analisados, e acrescentam que são registrados e acessíveis para que outros possam examinar e validar os resultados de um estudo.

O NIH admite que, devido à amplitude e variedade da ciência e dos estudos realizados com o seu apoio, não são especificados precisamente os formatos, os tipos de documentação dos dados, a apresentação ou a forma de transporte dos dados, afinal, o que pode ser perfeitamente normal e aplicável para uma área talvez não sirva para outras. Assim, o NIH corrobora que os membros de diferentes disciplinas e sociedades profissionais discutam as formas de compartilhamento, os padrões e melhores práticas.

A definição da *National Science Foundation* (NSF)<sup>35</sup> congrega do mesmo pensamento do NIH ao afirmar que dados são materiais factuais registrados comumente aceitos pela comunidade científica necessários para a validação de resultados de pesquisa, todavia, a NSF esclarece quais os tipos não são considerados dados de pesquisa, tais como: análises preliminares; rascunhos de artigos científicos; planos para pesquisas futuras; revisões por pares; comunicações com colegas de profissão. Segundo a NSF, o termo ‘material registrado’ exclui objetos físicos, por exemplo, amostras de laboratório. Assim, a definição da NSF inclui dados originais, metadados, protocolos experimentais, códigos de software escritos para análises estatísticas ou experimentais etc.

Para a NSF, os dados são essenciais para a comunidade de pesquisa e oferecem exemplos variados para facilitar o entendimento sobre estes, entretanto, é

---

<sup>34</sup> NIH Sharing Policies and Related Guidance on NIH-Funded Research Resources: <https://grants.nih.gov/policy/sharing.htm>.

<sup>35</sup> National Science Foundation (NSF), Data Management Guidance for CISE Proposals and Awards: [https://www.nsf.gov/cise/cise\\_dmp.jsp](https://www.nsf.gov/cise/cise_dmp.jsp).

interessante notar que há outros tipos de conjuntos de dados que vão além dos dados quantitativos.

Nesse sentido, a *National Endowment for the Humanities* (NEH)<sup>36</sup> destaca-se por expandir a sua definição e ofertar uma variedade de exemplos que incluem citações, códigos computacionais, algoritmos, ferramentas digitais, documentos, bases de dados, coordenadas geoespaciais, escavações arqueológicas, relatórios e artigos. Assim como a NSF, a NEH exclui análises preliminares, rascunhos de artigos, planos para pesquisas futuras, avaliações por pares, comunicações entre colegas, e acrescenta o alerta de que certos materiais devem permanecer em status confidencial até que sejam publicados e divulgados os resultados ou qualquer informação pessoal que resultaria em invasão de privacidade e quebra de sigilo.

A NEH sinaliza que há muitas variáveis para o que se constitui como dados de pesquisa, e, reforça que o gerenciamento dos dados depende de cada disciplina/área, pois, cada uma possui sua própria cultura em relação aos dados. Por esse prisma, é interessante notar que nas definições do NIH, NSF e NEH estabelecem que os dados devem ser validados e presumem o compartilhamento entre a comunidade científica. Conseqüentemente, são os dados produtos de pesquisa, possuem como características a heterogeneidade e devem ser contextualizados dentro das disciplinas/áreas a qual pertencem.

Do mesmo modo, Sayão e Sales (2015, p. 7) coadunam com a visão de que o conceito de dados de pesquisa pode variar entre as áreas e os pesquisadores, assim esclarecem:

A constatação de que os dados são gerados para diferentes propósitos, por diferentes comunidades acadêmicas e científicas e por meio de diferentes processos intensifica ainda mais essa percepção de diversidade. Tipos de dados podem incluir, por exemplo, números, imagens, textos, vídeos, áudio, software, algoritmos, equações, animações, modelos, simulações. Alguns tipos de dados têm valor imediato e duradouro, enquanto outros adquirem valor ao longo do tempo; alguns dados são capturados num momento específico e irrecuperável, enquanto outros são passíveis de se reproduzir.

Para os autores, em linhas gerais, qualquer dado obtido ou produzido a partir de um estudo científico utilizado para evidenciar, validar e apresentar resultados são dados de pesquisa. Dessa forma, Para Green, Macdonald e Rice (2009); Sayão e Sales (2015) e Corti, Van den Eynden, Bishop e Morgan-Brett (2015) em publicação

---

<sup>36</sup> Data Management Plan for NEH Office of Digital Humanities: [https://www.neh.gov/sites/default/files/2018-06/data\\_management\\_plans\\_2018.pdf](https://www.neh.gov/sites/default/files/2018-06/data_management_plans_2018.pdf).

editada pela *United Kingdom Data Archive*, os dados de pesquisa classificam-se como:

- a) Dados segundo a sua origem: observacionais; computacionais e experimentais;
- b) Dados segundo a sua natureza: números; imagens; vídeos; áudio; software; algoritmos; equações; animações ou modelos e simulações;
- c) Dados segundo a sua fase de pesquisa: brutos, crus ou preliminares (*Raw data*); derivados; canônicos ou referenciais;
- d) Dados governamentais: produzidos pelo governo, são dados oriundos das atividades dos diferentes órgãos da administração pública, embora não sejam dados coletados em investigações científicas, podem se tornar dados de pesquisa.

Sobre este último (dados governamentais), Aventurier (2016, p. 9) esclarece que é fundamental entender essa distinção porque “O movimento de dados abertos é radicalmente diferente do processo de dados de pesquisas abertos, mas poderá funcionar como alavanca para disponibilizar os dados de pesquisas e também levar dados governamentais disponíveis para a pesquisa científica.”

De fato, o autor toca em um ponto central e que muitas vezes gera bastante confusão, pois, apesar de serem temas que podem se relacionar, verdadeiramente há nuances distintas: o movimento de dados abertos “[...] defende a criação de um ecossistema fundamentado no acesso à dados que possibilite a participação da sociedade no desenvolvimento de um estado mais eficiente, com oferta de melhores serviços, e fazendo amplo uso de tecnologias.” (BRASIL, 2019, p. 1). Em relação aos dados de pesquisa abertos, Murray-Rust *et al.* (2010, online, tradução nossa) assim definem em seus “*Principles for Open Data in Science (Panton Principles)*”:

Por dados abertos em ciência, entendemos que eles estão disponíveis gratuitamente na internet, permitindo que qualquer usuário possa fazer o download, copiar, analisar, reprocessar, passar para um software ou usá-los para qualquer outro propósito sem barreiras financeiras, legais ou técnicas que não sejam as inseparáveis do acesso à própria internet. Para este fim, os dados relacionados com [os resultados] da ciência publicada devem ser explicitamente colocados em domínio público.

No tocante ao acesso aberto a dados de pesquisa, este tem sido incentivado por muitas entidades internacionais, inclusive nacionalmente com o lançamento em 2016 do “Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para a Ciência

Cidadã” pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). De acordo com o Ibict (2016), o manifesto tem o objetivo de estimular o acesso livre, abrangente e ilimitado aos dados de pesquisa (fontes primárias), tornando possível o uso, reuso, compartilhamento e reprodutibilidade da pesquisa.

Dando prosseguimento à classificação dos dados de pesquisa, conforme o exposto, os autores mostraram uma tipologia mais ampla dos tipos de dados de pesquisa do que a NIH, NEH e NSF, contemplando inclusive distintas áreas do conhecimento e demonstrando como os dados de pesquisa são heterogêneos e podem ser caracterizados por sua *natureza*, *origem* e *status* no fluxo de trabalho da pesquisa.

Contrastando com todas as argumentações anteriores, dados de pesquisa, de acordo com a Universidade de Leeds (2019, não paginado, tradução nossa), são “[...] todas as informações que foram coletadas, observadas, geradas ou criadas para validar resultados da pesquisa original.”

Sobre a tipologia dos dados, a Universidade de Leeds (2019) elenca as seguintes **formas**:

- Documentos, planilhas;
- Cadernos de laboratório, diários de campo, agendas;
- Questionários, transcrições, livros de códigos (Codebooks);
- Fitras de áudio e vídeo;
- Fotografias e filmes;
- Respostas de testes;
- Slides, artefatos, espécimes, amostras;
- Coleções de saída digitais (Digital outputs);
- Arquivos de dados;
- Conteúdos de bases de dados (Vídeo, áudio, texto, imagens);
- Modelos, algoritmos, scripts;
- Conteúdo de uma aplicação específica (registros de entrada, saída, arquivos de *log* para análise de software, simulações, esquemas);
- Metodologias e fluxos de trabalho;
- Procedimentos operacionais e protocolos padrão.

Nesse sentido, a Universidade de Leeds esboça uma infinidade de tipos de dados em maior amplitude e completa sobre dados de pesquisa, especialmente ao sinalizar os materiais físicos em papel, ou seja, dados não digitais (cadernos de laboratório, amostras, cadernos de esboço (*sketchbooks*), diários de campo etc., que costumam ser materiais únicos, contemplando a realidade e os passos de uma pesquisa e não mantém o foco apenas nos dados digitais.

Exemplo clássico desse tipo de material físico em papel e os dados valiosos contidos nele que citamos aqui é o caderno de Leonardo da Vinci que foi digitalizado pela *British Library*<sup>37</sup>: há nele centenas de esboços, rascunhos, anotações importantes que nos ajudam atualmente a compreender o pensamento do grande gênio criador e inventivo que foi Leonardo da Vinci. Outro exemplo memorável são os cadernos do matemático indiano Srinivasa Ramanujan<sup>38</sup>, os quais contêm diversas fórmulas e cálculos, alguns deles, usados até hoje no estudo do comportamento de buracos negros. Sua história inspiradora foi recontada no filme “O homem que viu o infinito” em 2015 e estrelado pelo ator Dev Patel<sup>39</sup>.

Perante a variedade de dados e características supracitadas, Green, Macdonald e Rice (2009), Sayão e Sales (2015), Corti, Van den Eynden, Bishop, Morgan-Brett (2015) e a *University of Leeds* (2019), detalhamos os tipos de dados de acordo com a sua **origem**, isto é, conforme os diferentes propósitos e através de diferentes processos, quais sejam:

- a) Dados Observacionais: São capturados em tempo real, por meio de observações diretas, geralmente insubstituíveis porque não podem ser coletados outra vez, pois capta um determinado comportamento ou fenômeno particular daquele momento específico e que não se repetirá exatamente do mesmo jeito. Exemplos: Dados de sensores, erupção de um vulcão, fotografia de um fenômeno natural (aurora boreal) ou de uma foto ou filmagem com um registro histórico (queda do muro de Berlim), levantamento do comportamento de uma comunidade, registros de imagens neurais, entre outros;

---

<sup>37</sup> Para visualizar e folhear “The Leonardo Notebook”, acesse: <http://www.bl.uk/turning-the-pages/?id=cb4c06b9-02f4-49af-80ce-540836464a46&type=book>.

<sup>38</sup> Imagens dos cadernos de Ramanujan: [http://ramanujan.sirinudi.org/html/unpublished\\_notebooks.html](http://ramanujan.sirinudi.org/html/unpublished_notebooks.html). Ver também os manuscritos disponíveis no catálogo da Biblioteca da Universidade de Trinity: <https://www.trin.cam.ac.uk/library/wren-digital-library/modern-manuscripts/ramanujan/>.

<sup>39</sup> Trailer do filme: <https://www.youtube.com/watch?v=mfyAyfGrA4A>.



- b) **Dados Experimentais:** São captados de equipamentos de laboratório e respeitando determinadas situações controladas para a reprodução de um experimento, como o registro de reações químicas, sequências genéticas, cromatogramas, dados do campo magnético toroidal etc. Nem sempre esses tipos de dados podem ser reproduzidos, por conta de variáveis, custos e por serem proibidas a reprodução em casos específicos ou que causem danos ao ambiente, animais e seres humanos;
- c) **Dados de simulação ou computacionais:** São gerados a partir de modelos de teste, simulações em ambientes virtuais, como modelos climáticos e econômicos. Para este tipo, são mais importantes os modelos, metadados, os dados de entrada do que os dados de saída. Há ainda a dependência tecnológica (hardware e software) e outras dependências técnicas para a sua replicação ao longo do tempo.
- d) **Dados Brutos (*Raw Data*):** São dados preliminares advindos diretamente dos instrumentos científicos da pesquisa;
- e) **Dados Derivados ou compilados:** São dados que foram transformados, processados ou combinados a partir de dados brutos ou de outros dados. Uma de suas características principais é que pode ser reproduzível mesmo que sejam perdidos, entretanto, isso pode representar aumento nos custos do projeto. Exemplos: Mineração de dados, bancos de dados compilados, modelos 3D etc.;
- f) **Dados canônicos ou referenciais:** São conglomerações de dados estáticos ou orgânicos consolidados, ou podem ser coleções menores de conjuntos de dados, geralmente revisadas por pares, provavelmente publicados e que passaram pelo processo de curadoria. Exemplos: Bancos de dados de sequências genéticas (GenBank)<sup>40</sup>, estruturas químicas ou portais de dados espaciais.

Não obstante, é por isso que todos os materiais de pesquisa e dados devem ser avaliados e selecionados para serem preservados a longo prazo, sejam estes digitais ou não, afinal, nem todos os tipos de dados precisam ser mantidos, tendo em vista que o pesquisador deve determinar quais dados deseja manter e preservar, ou

---

<sup>40</sup> Para saber mais sobre o GenBank, acesse: <https://www.ncbi.nlm.nih.gov/genbank/>.

em casos especiais, a instituição do pesquisador determina o que será preservado, pois há custos envolvidos, entre outros fatores, e, tudo isso também deve ser informado em um documento formal, o Plano de Gestão de Dados, do qual trataremos posteriormente neste trabalho.

Logo, entender o gerenciamento de dados permite compreendê-los em todas as suas construções e interfaces, desde a fase inicial do planejamento, passando pela coleta até chegar ao arquivamento dos dados. Para Sayão e Sales (2015) a distinção feita por essa categorização auxilia na escolha das estratégias de arquivamento e preservação dos dados. Segundo os autores, os dados fazem parte de um projeto e que muitas vezes recebe financiamento, então, é preciso pensar em sua continuidade, caso o apoio financeiro seja encerrado, pessoas saiam do projeto durante sua execução, entre outros fatores, afinal, os dados dependem de etapas e fases que se retroalimentam no ciclo de vida da pesquisa e no ciclo de vida dos dados.

## **2.4 Gestão de dados de pesquisa**

Meneghini (1998), em um estudo sobre a produção científica indexada em bases de dados internacionais, comparou a ciência a um iceberg: apenas uma parte acima da água estava visível. Analogamente, também é possível compará-la aos dados de pesquisa, tendo em vista que se encontram na mesma situação: estão disponíveis e acessíveis apenas uma pequena parcela, por meios das publicações formais (artigos e livros), enquanto todos os dados que ajudaram a embasar e sustentar os resultados das pesquisas são invisíveis, tampouco estão acessíveis e à disposição para uso e reuso, conforme ilustra a Figura 9.

Figura 9 - Iceberg dos dados científicos



Fonte: Elaborada pela autora (2019).

Assim como as publicações científicas são consideradas importantes, igualmente são os dados, porque tratam-se de elementos comprobatórios e auxiliam na construção da narrativa, no delineamento da pesquisa e subsidiam as interpretações do que é escrito em um trabalho. Mais do que isso: são registros únicos que uma vez perdidos não podem ser substituídos (SAYÃO; SALES, 2012).

Retomando a discussão acerca dos dados que estão invisíveis na pesquisa científica, Sayão e Sales (2019, p. 1) assim explicam o fenômeno:

[...] apesar dos avanços na comunicação científica e no acesso e compartilhamento dos materiais de pesquisa, metodologias, códigos instrumentos e muito mais, redefinidos pelos pressupostos da Ciência Aberta, há uma parcela considerável do trabalho científico que não está visível nem para a sociedade em termos de benefícios e qualidade de vida, nem para os pares para o reuso em novos experimentos ou para impulsionar a interdisciplinaridade da ciência. Uma grande parte das descobertas científicas não aparecem na literatura publicada, ao invés disso, reside nas gavetas e nos computadores pessoais dos pesquisadores. Este fenômeno tem os contornos mais nítidos no segmento da ciência conhecido como “cauda longa da ciência”, onde um grande número de pequenas equipes de pesquisadores e laboratórios independentes gera no seu dia a dia de pesquisa uma ampla variedade de coleções de dados de pesquisa.

Por esse motivo, é vital que os pesquisadores gerenciem os seus dados, não apenas com o propósito de organizá-los, acessá-los e usá-los, mas também como uma atitude sustentável e responsável de se conduzir a pesquisa científica. **Inclusive,**

**é o pesquisador quem irá decidir 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.**

Destarte, os dados de pesquisa são considerados valiosos porque “[...] o conhecimento é o motor do avanço científico, os dados são o seu combustível.” (SILVA, 2019, p. 3), nesse sentido, permite aos pesquisadores descobrir novos campos de pesquisa por meio de uma gestão adequada dos seus dados. Na literatura vigente, a gestão de dados científicos é conceituada como um conjunto de atividades que visa à coleta, armazenamento, gerenciamento, preservação e compartilhamento dos dados provenientes de uma pesquisa. Nesse sentido, cobre todos os aspectos relacionados à manipulação, organização, documentação, curadoria e agrega valor aos dados, garantindo-lhes sustentabilidade e acessibilidade a longo prazo. Tais ações permitem que os dados sejam reusados em outros projetos. A gestão de dados científicos tem como funções a racionalização de recursos, promoção da preservação, além de auxiliar os pesquisadores no planejamento, organização e documentação de seus projetos de pesquisa. (BALL, 2012; BORGMAN, 2007, 2010; CONSORTIUM OF EUROPEAN SOCIAL SCIENCE DATA ARCHIVES, 2017; CORTI; VAN DEN EYNDEN; BISHOP; MORGAN-BRETT, 2011; EUROPEAN COMMISSION, 2016; FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER, 2019; GREEN; MACDONALD; RICE, 2009; INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH, 2012; LEAGUE OF EUROPEAN RESEARCH UNIVERSITIES, 2018; SALES, 2014; SAYÃO; SALES, 2015; UNIVERSIDADE DE SÃO PAULO, 2019).

Sobre a gestão de dados, Bécard *et al.* (2016, p. 2, tradução nossa) afirma que “[...] consagra o compartilhamento e reuso de produções científicas como um princípio básico, para uma participação da ciência em um contexto.” Por sua vez, Mesguich *et al.* (2012) define a curadoria de dados como a identificação de vários conteúdos digitais de acordo com um determinado tema e consiste em selecionar e filtrar os temas mais relevantes, organizá-los e estruturá-los por meio de dispositivos digitais visando a sua promoção, disseminação e uso.

A iniciativa *Facilitate Open Science Training for European Research* (FOSTER), elaborou uma taxonomia (Figura 10) para a gestão de dados de pesquisa (Research Data Management). Nela, estão contidas as relações com os Planos de Gestão de Dados de Pesquisa (Research Data Management Plans), as Políticas de

Gestão de Dados de Pesquisa (Research Data Management Policies), os Serviços de Gestão de Dados de Pesquisa (Research Data Management Services), os Padrões em Gestão de Dados de Pesquisa (Research Data Management Standards) e as Ferramentas em Gestão de Dados de Pesquisa (Research Data Management Tools).

Nesse quesito, é preciso enfatizar que apesar de estar sendo exigido por diversas agências de fomento em todo o mundo, também as revistas e editoras científicas tem recomendado o depósito dos dados científicos em plataformas de gestão de dados e/ou repositórios generalistas e temáticos, “Os dados devem ser tão abertos quanto possível e fechados quando necessário.” (DIGITAL SCIENCE; FIGSHARE, 2017, p. 5, tradução nossa), tendo em vista que nem todo tipo de dado deve ser disponibilizado, evidentemente, por motivos que envolvem aspectos éticos e legais, e, portanto, serão respeitados. (SALES; SAYÃO, 2019).

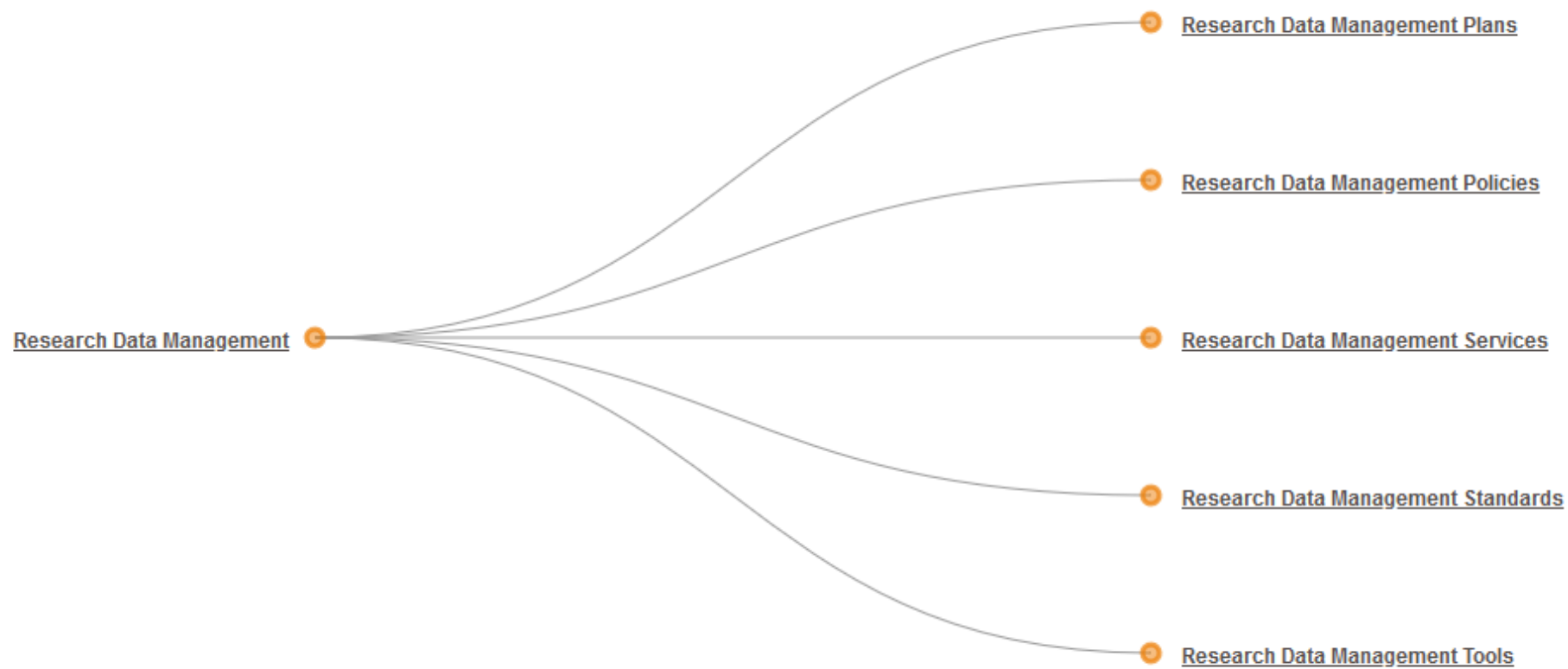
Dados de pesquisa se mostraram importantes no cenário nacional, especialmente a partir do projeto nacional que inclui o portal ‘Dados de Pesquisa Abertos’<sup>41</sup>, instituído pela Rede Nacional de Ensino e Pesquisa (RNP) e pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), em parceria com a Universidade Federal do Rio Grande do Sul (UFRGS) e a Universidade do Rio Grande (FURG) e visam ao desenvolvimento de um projeto em conjunto sobre o Acesso Aberto a Dados de Pesquisa (AADP).

O objetivo dessa Rede de Dados de Pesquisa Brasileira (RDP Brasil) “[...] é executar atividades que contribuam para a identificação de práticas, mapeamento de requisitos e prototipação de sistemas que facilitem a disseminação de informações científicas, [...] com o objetivo de promover e incentivar o compartilhamento de dados entre pesquisadores, conferindo maior eficiência na produção de conhecimento científico no Brasil.” (REDE DE DADOS DE PESQUISA BRASILEIRA, online, 2018). A referida rede já executou uma série de ações e pesquisas que buscaram mapear a situação nacional a respeito do compartilhamento de dados de pesquisa.

---

<sup>41</sup> Portal Dados de Pesquisa Abertos, disponível em: <https://dadosdepesquisa.rnp.br/>.

Figura 10 - Taxonomia da Gestão de Dados de Pesquisa



Fonte: Foster (2018).

Passando ao cenário global, hodiernamente, as principais agências de fomento do mundo passaram a exigir que os pesquisadores gerenciem seus dados e, conseqüentemente, submetam um ‘Plano de Gestão de Dados’ (PGD), a exemplo da: *Wellcome Trust*<sup>42</sup>; *National Science Foundation (NSF)*<sup>43</sup>; *National Endowment for the Humanities (NEH)*<sup>44</sup>; *National Institutes of Health (NIH)*<sup>45</sup>; Comissão Europeia (Projeto *Horizon 2020*)<sup>46</sup>; *Economic and Social Research Council*<sup>47</sup>; *Bill & Melinda Gates Foundation*<sup>48</sup>; *Office of Science and Technology Policy (OSTP)*; *Cancer Research UK*; *Research Councils UK*; *Netherlands Organisation for Scientific Research*; *Australian Research Council*; *National Health and Medical Research Council*; *Genome Canada*, entre outras. E no Brasil: a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp)<sup>49</sup> e a Universidade de São Paulo (USP)<sup>50</sup>. Cada vez mais editores estão exigindo o depósito dos dados em um repositório, como o *Dataverse*<sup>51</sup>; *Dryad*<sup>52</sup>; *Figshare*<sup>53</sup>, *Mendeley Data*<sup>54</sup>; *Zenodo*<sup>55</sup>, ou em repositórios de dados temáticos, por exemplo, o *Knowledge Network for Biocomplexity (KNB)*<sup>56</sup>, o repositório *protocols.io*<sup>57</sup>, exclusivamente para protocolos, o popular *GitHub*<sup>58</sup>, usado para projetos que envolvem programação, softwares e código-fonte aberto, entre outras aplicações, e até de revistas que possuem repositório próprio, como é o caso das publicações da *Institute of Electrical and Electronics Engineers (IEEE)*<sup>59</sup>, seguidos da abertura ou acesso aos dados correspondentes às publicações científicas para verificação e, em

<sup>42</sup> Disponível em: <https://wellcome.ac.uk/funding/guidance/developing-outputs-management-plan>.

<sup>43</sup> Disponível em: <https://www.nsf.gov/eng/general/dmp.jsp>.

<sup>44</sup> Disponível em: [https://www.neh.gov/sites/default/files/2018-06/data\\_management\\_plans\\_2018.pdf](https://www.neh.gov/sites/default/files/2018-06/data_management_plans_2018.pdf).

<sup>45</sup> Disponível em: [https://grants.nih.gov/grants/policy/data\\_sharing/data\\_sharing\\_guidance.htm](https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm).

<sup>46</sup> Disponível em: [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm).

<sup>47</sup> Disponível em: <https://esrc.ukri.org/funding/guidance-for-grant-holders/research-data-policy/>.

<sup>48</sup> Disponível em: <https://docs.gatesfoundation.org/documents/faq.pdf>.

<sup>49</sup> Informações sobre o Plano de Gestão de Dados no site da Fapesp: <http://www.fapesp.br/gestaodedados/>.

<sup>50</sup> Medidas adotadas pela Pró-Reitoria de Pesquisa (PRP) da USP para tornar obrigatória a apresentação do PGD: <https://prp.usp.br/gestao-de-dados-cientificos/>.

<sup>51</sup> Harvard Dataverse: <https://dataverse.harvard.edu/>. No Brasil, o Instituto Brasileiro de Ciência e Tecnologia (Ibict), também adotou o Dataverse como opção para repositório de dados. Ver também: <https://repositoriopesquisas.ibict.br/>.

<sup>52</sup> Dryad: <http://datadryad.org/>.

<sup>53</sup> Figshare: <https://figshare.com/>.

<sup>54</sup> Mendeley Data: <https://data.mendeley.com/>.

<sup>55</sup> Zenodo: <https://zenodo.org/>. O Zenodo também oferece um ambiente de teste para API's e outras aplicações, o Zenodo Sandbox, disponível em: <https://sandbox.zenodo.org/>.

<sup>56</sup> KNB: <https://knb.ecoinformatics.org/about>.

<sup>57</sup> Protocols.io: <https://www.protocols.io/>.

<sup>58</sup> GitHub: <https://github.com/>.

<sup>59</sup> IEEE Dataport: <https://ieee-dataport.org/>.

alguns casos, para reproduzir experimentos e métodos descritos no manuscrito submetido.

Sobre os repositórios, estes têm sido empregados como uma das formas de democratizar o acesso ao conhecimento científico, não é por acaso que diversos países incentivaram a criação de repositórios institucionais há alguns anos, os quais são um “Sistema de informação para armazenar, preservar e difundir a produção intelectual de dada instituição, geralmente uma comunidade universitária.” (FUNDAÇÃO OSWALDO CRUZ, 2019, não paginado). A iniciativa foi coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e apesar dos esforços, ainda existe um reflexo das desigualdades em acesso ao conhecimento e a falta de concretização de políticas e ações que contribuem para a gestão de dados de pesquisa e para a Ciência Aberta, como a implantação de repositórios de dados, estes definidos pela Fiocruz (2019, online) como uma “Estrutura tecnológica e gerencial que permite que pesquisadores depositem seus dados de pesquisa para armazenamento e amplo acesso.”

De acordo com o RE3DATA (Figura 11) muitas regiões possuem repositórios implantados, por outro lado, outros lugares ainda não possuem sequer um repositório.

Figura 11 - Mapa de repositórios por países



Legenda: Em azul, países sem repositório. Os demais, em verde, possuem repositórios.

Fonte: RE3DATA (2018).

A América Latina é composta por quase todos os países da América do Sul e América Central, com exceção das Guianas (Guiana e Guiana Francesa), Suriname, Belize e Jamaica. Ao todo são vinte países: Argentina, Bolívia, Brasil, Chile, Colômbia,



Costa Rica, Cuba, Equador, El Salvador, Guatemala, Haiti, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, República Dominicana, Uruguai e Venezuela. Destes vinte países, doze ainda não possuem nenhum tipo de repositório (institucional, de dados ou disciplinar) e apenas oito os têm.

Tabela 1 - Total de repositórios por países da América Latina

PAÍS	QUANTIDADE	PAÍS	QUANTIDADE
Argentina	4	Haiti	0
Bolívia	0	Honduras	0
Brasil	8	México	11
Chile	1	Nicarágua	0
Colômbia	2	Panamá	2
Costa Rica	1	Paraguai	0
Cuba	0	Peru	2
Equador	0	República Dominicana	0
El Salvador	0	Uruguai	0
Guatemala	0	Venezuela	0

Fonte: Elaborada pelas autoras, baseado nos dados do RE3DATA (2019).

Um dado relevante a ser considerado sobre o Brasil em relação aos dados da tabela acima é que apesar de ter o total de oito repositórios ao país, apenas metade<sup>60</sup> são repositórios genuinamente brasileiros, enquanto os demais são parcerias internacionais com outros países e nas quais o Brasil coopera e trabalha em estudos científicos. Coincidência ou não, os países que ainda não possuem repositórios e

<sup>60</sup> Disponível em: [https://www.re3data.org/search?query=&countries\[\]=BRA](https://www.re3data.org/search?query=&countries[]=BRA).

políticas voltadas para as práticas abertas e colaborativas da ciência são países que possuem conflitos internos no campo político e social.

Para Costa (2006, p. 49), “não parece restar dúvidas, portanto, de que a filosofia aberta tende a se constituir no fundamento inexorável da comunicação científica daqui por diante”. Contudo, sob a ótica da autora, faz-se necessário alertar que nem todas as informações de uma instituição estão disponibilizadas em seus repositórios porque envolvem diversos fatores, tais como: grandes volumes documentais de materiais diversos que precisam ser previamente analisados; algumas instituições ainda não adotaram o autoarquivamento; a falta de profissionais disponíveis e capacitados; os custos com digitalização e espaço de armazenamento em servidores; restrições em políticas rígidas de divulgação de informações de algumas instituições; restrições que envolvem direitos autorais e patentes; segurança dos dados e informações sensíveis que não podem ser divulgadas; entre muitos outros casos.

Para além dos repositórios como ferramenta de democratização de conhecimento e como plataforma de gestão de dados, uma novidade que merece destaque no que diz respeito aos dados de pesquisa é o surgimento de uma nova modalidade de publicação de artigos: os *Data Papers*. Basicamente, eles descrevem como foram criados os *datasets*, exploram vários outros aspectos e até extraem novos *insights* desses conjuntos de dados, a exemplo das revistas (*Data Journals*): *Atomic Data and Nuclear Data Tables*<sup>61</sup>; *Biodiversity Data Journal*<sup>62</sup>; *Biomedical Data Journal*<sup>63</sup>; *BMC Research Notes*<sup>64</sup>; *Chemical Data Collections*<sup>65</sup>; *Data*<sup>66</sup>; *Data in Brief*<sup>67</sup>; *Data Observer*<sup>68</sup>; *Data Papers in Science*<sup>69</sup>; *Ecology*<sup>70</sup>; *Earth System Science Data*<sup>71</sup>;

---

<sup>61</sup> Disponível em: <https://www.journals.elsevier.com/atomic-data-and-nuclear-data-tables/>.

<sup>62</sup> Disponível em: <https://bdj.pensoft.net/>.

<sup>63</sup> Disponível em: <http://biomed-data.eu/>.

<sup>64</sup> Disponível em: <https://bmcrsnotes.biomedcentral.com/>.

<sup>65</sup> Disponível em: <https://www.journals.elsevier.com/chemical-data-collections>.

<sup>66</sup> Disponível em: <https://www.mdpi.com/journal/data>.

<sup>67</sup> Disponível em: <https://www.journals.elsevier.com/data-in-brief>.

<sup>68</sup> Título anterior desta publicação era “European Data Watch”. Disponível em: <https://www.ratswd.de/publikationen/data-observer>.

<sup>69</sup> Disponível em: <https://www.hindawi.com/journals/dpis/>.

<sup>70</sup> Disponível em: [https://esajournals.onlinelibrary.wiley.com/hub/journal/19399170/resources/data\\_paper\\_inst\\_ecy](https://esajournals.onlinelibrary.wiley.com/hub/journal/19399170/resources/data_paper_inst_ecy).

<sup>71</sup> Disponível em: <https://www.earth-system-science-data.net/>.

*F1000Research*<sup>72</sup>; *Genomics Data*<sup>73</sup>; *Geoscience Data Journal*<sup>74</sup>; *GigaScience*<sup>75</sup>; *IEEE Journals*<sup>76</sup>; *Internet Archaeology*<sup>77</sup>; *Journal of Chemical & Engineering Data*<sup>78</sup>; *Journal of Open Archaeology Data*<sup>79</sup>; *Journal of Open Health Data*<sup>80</sup>; *Journal of Open Psychology Data*<sup>81</sup>; *Journal of Physical and Chemical Reference Data*<sup>82</sup>; *Nuclear Data Sheets*<sup>83</sup>; *Open Data Journal for Agricultural Research*<sup>84</sup>; *Open Journal of Bioresources*<sup>85</sup>; *Research Data Journal for the Humanities and Social Sciences*<sup>86</sup>; *Scientific Data*<sup>87</sup>; *The International Journal of Robotics Research*<sup>88</sup>; entre outras.

Além da diversidade de periódicos científicos citados, que publicam os *data papers*, existem as revistas que solicitam o envio do Plano de Gestão de Dados relacionado ao artigo ou ao *data paper* como pré-requisito para a sua publicação, a exemplo do periódico *Research Ideas and Outcomes - RIO Journal*<sup>89</sup>.

Diante desse cenário de tantas novidades no ambiente acadêmico, editorial e na comunicação científica, conforme vimos acima, há muitas possibilidades por meio do uso dos dados, vários recursos, revistas e repositórios à disposição dos pesquisadores, ademais, a gestão de dados científicos torna-se cada vez mais premente, uma vez que "Disponibilizar os dados de maneira oportuna e responsável garante que outras pessoas possam verificá-los, desenvolvê-los e usá-los para avançar o conhecimento e melhorar a saúde". (DIGITAL SCIENCE; FIGSHARE, 2017, p. 4, tradução nossa). Logo, os dados científicos se tornaram importantes recursos informacionais que representam a garantia da qualidade e integridade da pesquisa.

<sup>72</sup> Disponível em: <https://f1000research.com/>.

<sup>73</sup> Disponível em: <https://www.journals.elsevier.com/genomics-data>.

<sup>74</sup> Disponível em: <https://rmets.onlinelibrary.wiley.com/journal/20496060>.

<sup>75</sup> Disponível em: <https://academic.oup.com/gigascience>.

<sup>76</sup> As revistas da IEEE possuem opções para publicação dos dados em seu próprio repositório (IEEE Dataport): <https://iee-dataport.org/>. Disponível em: <https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-article/authoring-tools-and-templates/about-sharing-your-data-and-code/>.

<sup>77</sup> Disponível em: <http://intarch.ac.uk/authors/data-papers.html>.

<sup>78</sup> Disponível em: <https://pubs.acs.org/journal/jceaax>.

<sup>79</sup> Disponível em: <https://openarchaeologydata.metajnl.com/>.

<sup>80</sup> Disponível em: <https://openhealthdata.metajnl.com/>.

<sup>81</sup> Disponível em: <https://openpsychologydata.metajnl.com/>.

<sup>82</sup> Disponível em: <https://aip.scitation.org/toc/jpr/current?expanded=47>.

<sup>83</sup> Disponível em: <https://www.journals.elsevier.com/nuclear-data-sheets/>.

<sup>84</sup> Disponível em: <https://library.wur.nl/ojs/index.php/odjar/>.

<sup>85</sup> Disponível em: <https://openbioresources.metajnl.com/>.

<sup>86</sup> Iniciou como um projeto piloto. Ver também: <http://dansdatajournal.nl/>. Disponível em: <https://brill.com/abstract/journals/rdj/aop/issue.xml>.

<sup>87</sup> Disponível em: <https://www.nature.com/sdata/>.

<sup>88</sup> Disponível em: <https://journals.sagepub.com/home/ijr>.

<sup>89</sup> RIO Journal: <https://riojournal.com/>.

Por sua importância, a gestão de dados científicos configura-se um desafio aos pesquisadores e instituições. Os dados brutos, durante e após a pesquisa, precisam ser armazenados, organizados e preservados para sua recuperação futura, o que possibilitará sua reutilização e a retroalimentação da ciência. Para que isso seja uma realidade, é preciso formar equipes especializadas que ofereçam suporte aos pesquisadores na realização de uma gestão satisfatória dos seus dados. (LEHMKUHL; GAVRON; ALVES; MORAES DE BEM, p. 147).

Inclusive, pode auxiliar a dirimir questões sobre os direitos autorais, de propriedade intelectual e licenciamento. Com efeito, a autoria e a autoralidade colaborativa tem sido objeto de estudo ligado também aos dados científicos e à Ciência Aberta, a exemplo do trabalho de Oliveira (2016). Então, os pesquisadores podem adotar modelos, diretrizes ou políticas sugeridas por sua área, instituição ou pelos periódicos científicos, a exemplo das diretrizes do *Committee on Publication Ethics* (COPE)<sup>90</sup>, do *Council Science Editors* (CSE)<sup>91</sup>, do *International Committee of Medical Journal Editors* (ICMJE)<sup>92</sup> e da *Contributor Roles Taxonomy* (CRediT)<sup>93</sup>.

Outrossim, a gestão de dados e a elaboração do PGD são fundamentais porque promovem a integridade científica, a ética, a reprodutibilidade da pesquisa, conforme veremos adiante.

#### **2.4.1 Ciclos de vida da pesquisa e dos dados**

Coincidência ou não, **contexto** é a palavra-chave para sistematizar, processar, tratar, descrever e documentar dados de pesquisa, afinal, para que possam ser preservados, compartilhados e reutilizados, é preciso que façam sentido por meio de modelos e processos detalhados, além de uma documentação que tenha por base uma descrição minuciosa e precisa dos dados a exemplo de um plano de gestão de dados (PGD), bem como na criação de metadados, dando-lhes um contexto.

<sup>90</sup> O COPE possui vários fluxogramas e diretrizes norteadoras para pesquisadores e editores de periódicos a fim de incentivar as melhores práticas relacionadas à produção de conhecimento em publicações científicas (Authorship and Contributorship): <https://publicationethics.org/authorship>.

<sup>91</sup> Na publicação do CSE intitulada "White Paper on Publication Ethics", na seção 2.4.1.5 Access to and Provision of Data, trata a respeito do acesso, provisão e compartilhamento dos dados de pesquisa, e, na seção 3.1.2 Falsification and Fabrication of Data, há orientações para coibir a falsificação e fabricação de dados: <https://www.councilscienceeditors.org/resource-library/editorial-policies/white-paper-on-publication-ethics/>.

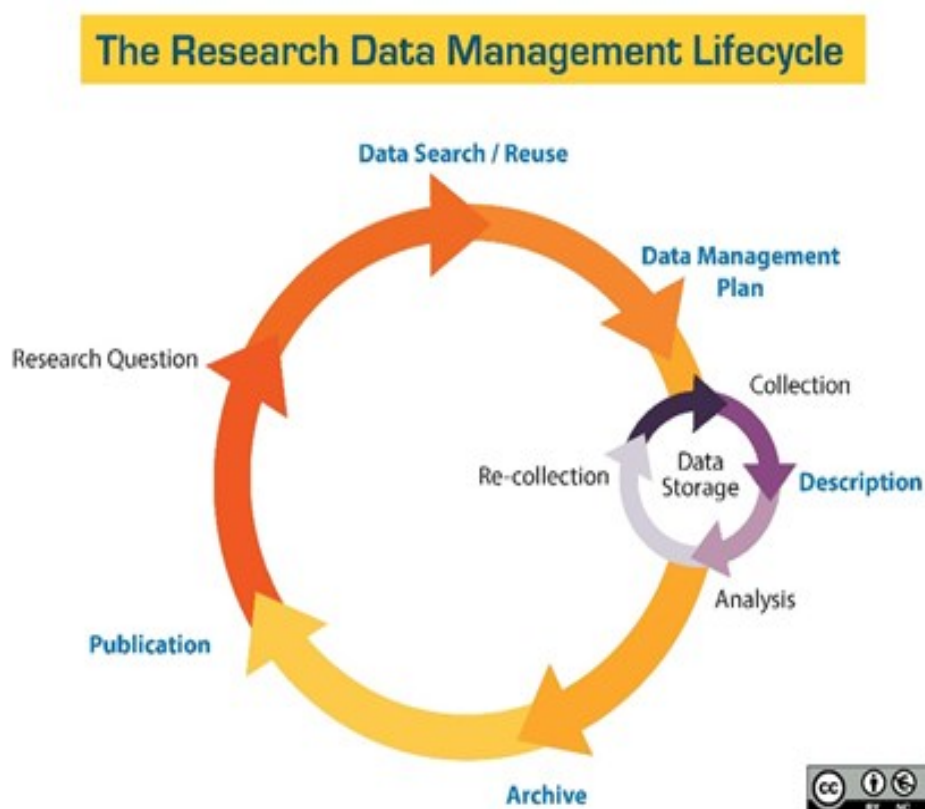
<sup>92</sup> Recomendações do ICJME: <http://www.icmje.org/recommendations/>.

<sup>93</sup> O CRediT é uma taxonomia e inclui 14 papéis que podem representar os autores e colaboradores em funções desempenhadas por eles na produção acadêmica científica. Foi criada pelo Consortia Advancing Standards in Research Administration Information (Casrai), uma iniciativa internacional e associação sem fins lucrativos liderada por instituições de pesquisa. Para saber mais sobre a Taxonomia CRediT, acesse: <https://casrai.org/credit/>.

Nesse sentido, diante da obsolescência dos objetos digitais e o rápido avanço das tecnologias, a curadoria digital e a Ciência da Informação cumprem um importante papel a fim de preservar, tratar, representar, disponibilizar, disseminar e preparar os dados para as fases de descoberta, uso e reuso durante o ciclo de vida de um objeto digital. Assim sendo, existem muitos modelos de ciclo de vida dos dados com a finalidade de sistematizar essas etapas e recomendar boas práticas para instituições, bibliotecários, arquivistas, pesquisadores em geral e outros profissionais afins, isto é, os ciclos de vida dos dados oferecem uma visão dos processos que devem ser realizados, tal como a curadoria digital.

Nesse sentido, vale destacar nessa seção alguns modelos de ciclo de vida dos dados, os quais possuem fases que se retroalimentam no ciclo de vida da pesquisa e no ciclo de vida dos dados, a exemplo da Universidade de Santa Cruz, na Califórnia, a qual combinou dois ciclos (da pesquisa e dos dados) para exemplificar a influência de um sobre o outro, conforme ilustra a figura 12.

Figura 12 - Ciclo de vida da Gestão de Dados

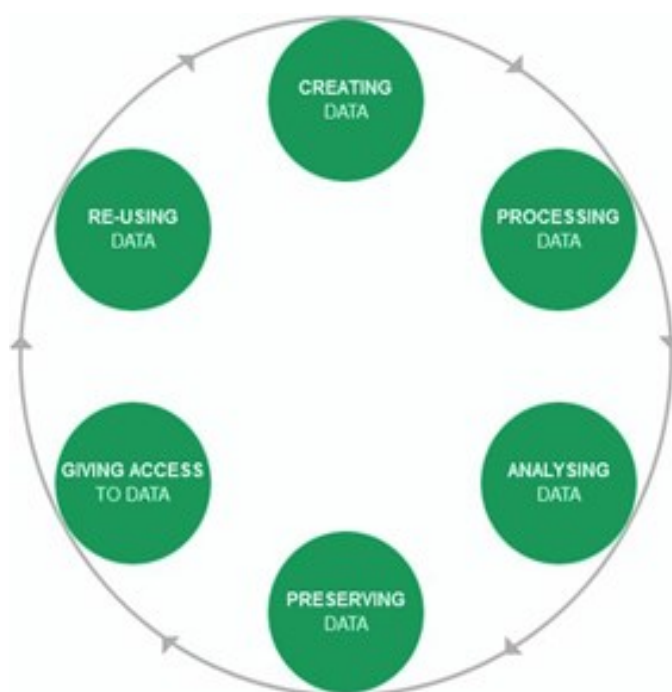


Fonte: The University of California, Santa Cruz (2015).

Ilustrado na figura acima, a pesquisa inicia com uma pergunta (*Research question*), seguida das fases de busca dos dados e reuso (*Data search/reuse*), a elaboração do PGD (*Data Management Plan*); a descrição (*Description*), subdividida em coleta (*Collection*), análise (*Analysis*) e nova coleta (*Re-collection*) e no centro deste ciclo menor está o armazenamento dos dados (*Data storage*); seguido do arquivamento (*Archive*) e da publicação (*Publication*).

Há outros exemplos de ciclo de vida dos dados de pesquisa e com representações distintas, mas, que envolvem as etapas principais deste processo. O ciclo de vida dos dados do *United Kingdom Data Archive* (2014) envolve as etapas de criação dos dados (*Creating data*), Processamento (*Processing data*), análise (*Analysing data*), preservação (*Preserving data*), acesso (*Giving access to data*) e reutilização (*Re-using data*).

Figura 13 - Research Data Life Cycle



Fonte: UK Data Archive (2014).

No ciclo do *United Kingdom Data Archive* (figura 13) a etapa da criação dos dados abarca a concepção da pesquisa e seu delineamento; a elaboração do PGD; a coleta dos dados e sua captura. A etapa do processamento corresponde ao tratamento dos dados coletados: compreende a transcrição, tradução e codificação, entrada ou digitalização. Em seguida, é necessário realizar a verificação dos dados, a validação e limpeza de dados (por exemplo, eliminação de dados duplicados,

incompletos e inconsistências). Se for preciso, a anonimização dos dados deve ser feita para preservar detalhes da investigação relacionados aos sujeitos de pesquisa que não devem ser expostos ou divulgados. O controle de acesso aos dados deve ser discutido nesta etapa inicial, uma vez que certos dados só podem estar disponíveis para os membros que fazem parte do projeto de pesquisa. Assim, uma das formas de garantir o controle de acesso é por meio de login e senha, por exemplo. Em seguida, há a descrição dos dados coletados, gestão, *backup* e armazenamento dos dados.

Após o tratamento e organização dos dados, a etapa de análise envolve a interpretação dos dados, produção dos resultados e sua preparação para a publicação em publicações científicas e preparação para a preservação. Na fase de preservação dos dados, são consideradas as extensões dos arquivos de dados e aspectos relacionados à migração para outros formatos. Backup e armazenamento dos dados são também uma preocupação dessa etapa descrita no ciclo de vida dos dados, e, portanto, deve ser estabelecida uma frequência para se executar essa tarefa. Alguns *backups* podem ser diários, semanais, mensais etc., tudo irá depender da necessidade e melhor planejamento para evitar perdas dos dados e contribuir para a sua preservação e acesso no futuro. Considera-se ainda que os metadados e a documentação que acompanham os dados devem ser atualizados se alguma alteração ocorrer ou se for necessária antes de proceder ao arquivamento dos dados.

Para a fase de acesso aos dados, pensar em sua distribuição e compartilhamento também exige pensar, antes de tudo, no controle de acesso estabelecido ou não na fase inicial do ciclo. É preciso esclarecer se haverá embargo para acesso aos dados, e, nesses casos, deve-se justificar os motivos e estabelecer um período de embargo. Tais medidas estão relacionadas à aspectos jurídicos de reutilização dos dados, como patentes, pesquisas em fase de desenvolvimento etc.

A última fase do ciclo da *United Kingdom Data Archive* é a de reutilização dos dados. Envolve o acompanhamento da pesquisa e dos dados produzidos, pode gerar uma nova pesquisa, além de fazer uma síntese ou revisão sobre a pesquisa a fim de examinar os resultados. Lembramos que os dados podem ser reutilizados tanto pelo pesquisador que produziu os dados como por outros pesquisadores que tiveram acesso aos dados através de repositórios de dados, demonstrando o potencial de utilização dos dados.

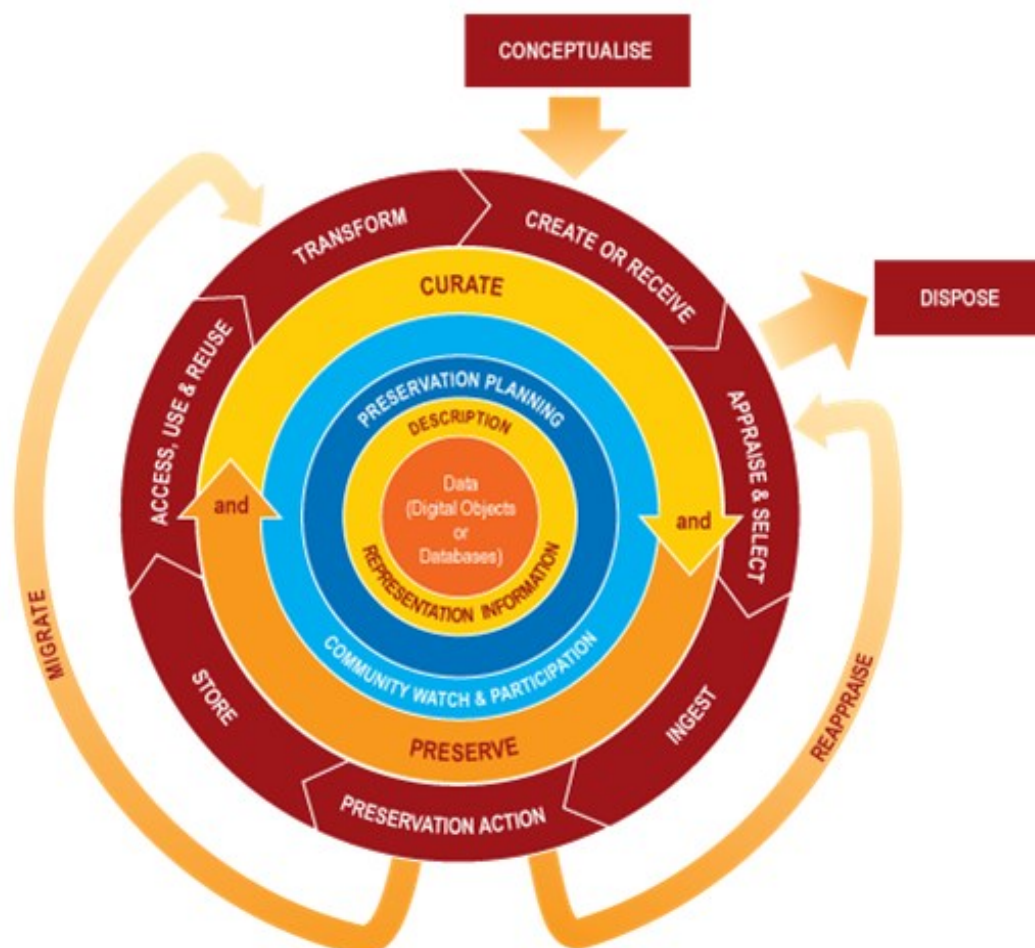
O modelo de Ciclo de vida dos dados do *Digital Curation Centre* (DCC) (figura 14) apresenta a atividade da curadoria atrelada a diferentes papéis profissionais,

como o bibliotecário e o arquivista em diferentes partes deste ciclo. Esse modelo é bastante completo porque permite aos curadores identificar pontos fortes e fracos, lacunas e limites no processo de gestão de dados de pesquisa.

O DCC esclarece que este é um modelo ideal, ajuda a fornecer uma visão geral dos estágios necessários para uma boa curadoria dos dados e bem-sucedida em todos os estágios. Este modelo pode ser utilizado para planejar as atividades de curadoria e gerenciamento de dados em uma instituição.

Outro diferencial deste modelo do DCC é que os usuários deste podem entrar em qualquer estágio do ciclo de vida, dependendo de sua área, função e necessidade. Assim, conforme o próprio DCC exemplifica em seu site, um gestor de repositório digital pode se engajar a partir da etapa de processamento (*Ingest*) e se inserir ou retroceder a outras etapas a fim de que possa melhorar o gerenciamento dos dados e a curadoria a longo prazo.

Figura 14 - Modelo do ciclo de vida dos dados e curadoria (DCC)



Fonte: DCC (2019).



No ciclo de vida do DCC os dados (*Data*) estão no núcleo, no centro da curadoria e constituem-se como qualquer informação em formato digital binário, incluindo os objetos digitais simples (arquivos textuais, imagéticos, sonoros etc.) e bancos de dados (coleções estruturadas de registros). A camada acima do núcleo presente no ciclo do DCC corresponde à descrição (*Description*) e representação da informação (*Representation Information*): nela deve-se atribuir os metadados, realizar a descrição, informar aspectos técnicos, estruturais e de preservação utilizando padrões apropriados a fim de assegurar o controle adequado a longo prazo. Essa etapa precisa ter informações suficientes para entender e processar os dados.

As etapas seguintes representadas no ciclo do DCC são o planejamento de preservação (*Preservation planning*), a vigilância e a participação da comunidade (*Community watch & participation*), curadoria e preservação (*Curate & preserve*). Após todas essas camadas, há ações sequenciais para conceber e planejar a criação dos dados, incluindo as opções de método de captura e armazenamento. São elas: Conceitual (*Conceptualise*)<sup>94</sup>, Criar ou receber (*Create or receive*)<sup>95</sup>, Avaliar e selecionar (*Appraise & select*)<sup>96</sup>, Processar/transferir (*Ingest*)<sup>97</sup>, Ação de preservação (*Preservation action*)<sup>98</sup>, Armazenar (*Store*)<sup>99</sup>, Acesso, uso e reuso (*Access, use & re-use*)<sup>100</sup>, Transformar (*Transform*). Ações ocasionais podem ser executadas nesse modelo de ciclo, como o descarte (*Dispose*) de alguns dados que não foram selecionados, a reavaliação (*Reappraise*) de dados que apresentam alguma falha em procedimentos de verificação e validação, e ainda a migração (*Migrate*) de alguns dados em um formato diferente de acordo com a obsolescência de hardware e software, para melhor armazenamento, preservação, garantia de integridade, entre outros.

Outro modelo de ciclo de vida dos dados do DataOne (Figura 15) é bastante popular e conhecido entre a comunidade acadêmica, pois apresenta vários processos envolvidos na gestão de dados de pesquisa. O ciclo de vida dos dados é uma parte de cada projeto de pesquisa científica, com decisões que giram em torno de cada

<sup>94</sup> Conceptualise checklist: [http://www.dcc.ac.uk/sites/default/files/Conceptualisation Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Conceptualisation%20Checklist.pdf).

<sup>95</sup> Create or receive checklist: [http://www.dcc.ac.uk/sites/default/files/Create Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Create%20Checklist.pdf).

<sup>96</sup> Appraise & select checklist: [http://www.dcc.ac.uk/sites/default/files/Select and Appraise Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Select%20and%20Appraise%20Checklist.pdf).

<sup>97</sup> Ingest checklist: [http://www.dcc.ac.uk/sites/default/files/Ingest and Store Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Ingest%20and%20Store%20Checklist.pdf).

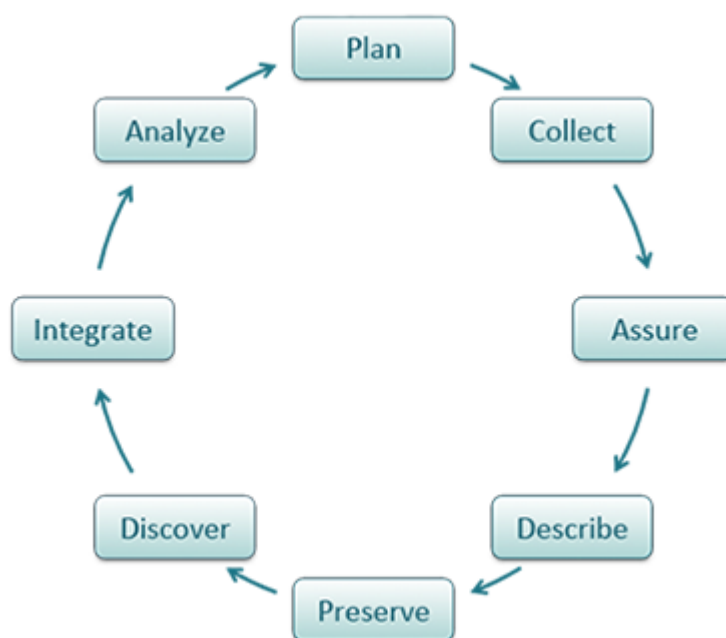
<sup>98</sup> Preservation action checklist: [http://www.dcc.ac.uk/sites/default/files/Preservation Action Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Preservation%20Action%20Checklist.pdf).

<sup>99</sup> Store checklist: [http://www.dcc.ac.uk/sites/default/files/Ingest and Store Checklist\\_0.pdf](http://www.dcc.ac.uk/sites/default/files/Ingest%20and%20Store%20Checklist_0.pdf).

<sup>100</sup> Access, use & re-use checklist: [http://www.dcc.ac.uk/sites/default/files/Access and Reuse Checklist.pdf](http://www.dcc.ac.uk/sites/default/files/Access%20and%20Reuse%20Checklist.pdf).

aspecto que são feitos por padrão ou com intento estratégico. Além disso, o ciclo é representado de forma lógica, suas etapas podem ocorrer em uma ou várias vezes, em sequências diferentes, com algumas etapas ocorrendo em conjunto e algumas repetidas com maior frequência do que outras. (DATAONE, 2013).

Figura 15 - Ciclo de vida dos dados (DataOne)



Fonte: DataOne (2019).

Possui oito etapas, a saber:

- Planejar (Plan): Etapa em que se realiza a descrição dos dados, informa como serão gerenciados e disponibilizados;
- Coletar (Collect): Corresponde à fase de descrição da coleta dos dados, quais os instrumentos foram utilizados, em que formas os dados se apresentam etc.;
- Assegurar (Assure): Verificação e controle dos dados são realizados neste passo a fim de garantir a qualidade dos dados;
- Descrever (Describe): Descrição minuciosa dos dados, utilizando-se inclusive de padrões de metadados, vocabulários controlados, ontologias, taxonomias etc.;

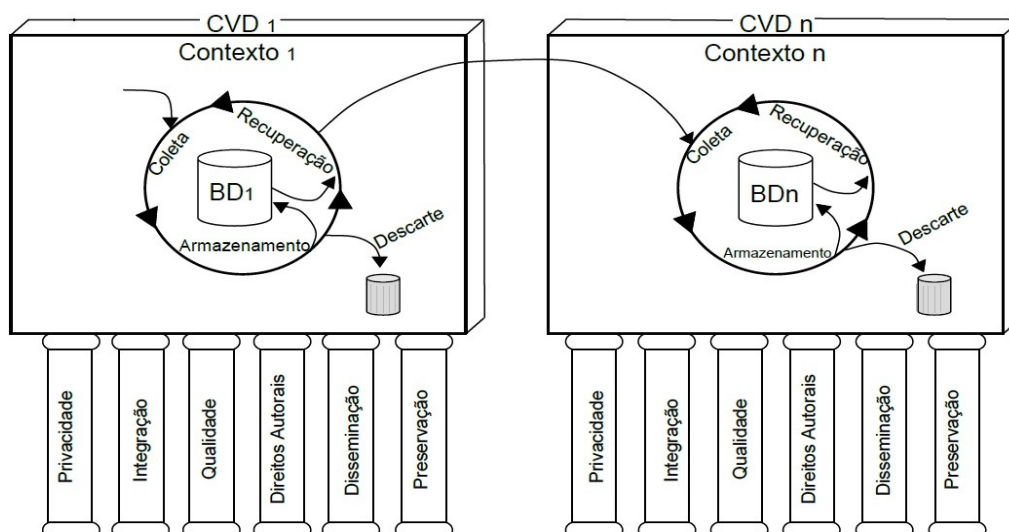
- Preservar (Preserve): De acordo com este modelo de ciclo, essa é a fase em que os datasets são transferidos para centros de dados apropriados a fim de serem preservados a longo prazo;
- Descobrir (Discover): Corresponde à etapa em que os dados descritos e disponibilizados são descobertos em processos de busca e acessados por meio da detalhada descrição dos metadados;
- Integrar (Integrate): Nesta fase, dados de fontes diferentes podem ser combinados e dar origem a novos datasets que podem ser analisados;
- Analisar (Analyse): Etapa correspondente à análise dos dados.

Um aspecto interessante que deve ser evidenciado é o fato de que nem todas as etapas dos ciclos de vida dos dados podem ser cumpridas e seguir um caminho linear de acordo com cada modelo, pois, isso dependerá do tipo de pesquisa a ser realizada. Alguns pesquisadores, talvez, cumpram todas as fases, entretanto, outros podem apenas seguir as etapas de descoberta, integração, análise e então sintetizar os dados existentes em seus resultados, a exemplo das metanálises, isto é, os ciclos de vida dos dados e seus diferentes modelos existem para auxiliar, dar um norte para a tarefa de gerenciamento dos dados científicos, contudo, maior do que o desenho desses ciclos é o delineamento da pesquisa e seu próprio ciclo de vida. (FOSTER, 2019; SAYÃO; SALES, 2015).

Na área de Ciência da Informação, Sant'Ana (2016) criou um modelo de ciclo de vida dos dados (CDV-CI), conforme ilustrado na figura 16, e baseia-se em quatro fases: coleta, armazenamento, recuperação e descarte. Além disso, são acrescentadas ao modelo os objetivos relacionados a cada uma das fases mencionadas. Os objetivos são: privacidade, integração, qualidade, direitos autorais, disseminação e preservação.

O modelo idealizado ilustra o papel complementar da Ciência da Informação nesse quesito, “[...] na proposta de novos caminhos para que os usuários, de forma mais democrática e aberta possam acessar e utilizar dados, reduzindo a assimetria informacional que pode surgir entre os que os detêm e os que dele precisam.” (SANT’ANA, 2016, p. 138).

Figura 16 - Ciclo de vida dos dados em Ciência da Informação de Sant'Ana (CDV-CI)



Fonte: Sant'Ana (2016).

Em suma, as quatro fases do modelo de Sant'Ana (2016) sintetizam as atividades previstas no ciclo de vida da curadoria de dados e no uso destes. O autor destaca que a Ciência da Informação tem um papel importante nesta temática porque oferece um arcabouço teórico capaz de contribuir na construção de novos conhecimentos e técnicas, especialmente nas fases de coleta e recuperação as quais têm tido maior concentração de estudos.

Adiante, cumprindo uma das etapas presentes nos diversos ciclo de vida dos dados apresentados nesta seção, trataremos sobre o documento formal que estabelece a minuciosa descrição dos dados oriundos de pesquisa: o Plano de Gestão de Dados (PGD).

#### **2.4.2 Plano de Gestão de Dados (PGD)**

Um PGD se constitui em um documento que define claramente como serão criados e documentados os dados; estabelece quais serão as pessoas a ter acesso aos dados tanto no decorrer da pesquisa quanto no final, ao disponibilizar o acesso para os usuários, descreve a qual público-alvo se destinam os dados; informa onde serão armazenados os dados e como isso será feito; discrimina quem são os responsáveis pela preservação dos dados; e define se estes poderão ser compartilhados e reutilizados, dependendo, evidentemente, de questões autorais, de patentes ou de dados sensíveis.

Os Planos de Gestão de Dados são considerados como elementos chave para uma boa gestão no decorrer de uma investigação. (FOSTER, 2019). O PGD descreve todo o ciclo de vida da pesquisa e dos dados, desde a fase da coleta, processamento e tratamento até a disponibilização.

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

Em suma, o PGD é um documento que descreve como os dados produzidos por uma pesquisa serão gerenciados, armazenados, preservados e divulgados em repositórios confiáveis de dados. Nas palavras de Sayão e Sales (2015, p. 15, grifo dos autores), o Plano de Gestão de Dados:

[...] descreve o ciclo de vida de gestão para todos os dados que serão coletados, processados ou gerados por um projeto de pesquisa. De uma forma abreviada, ele se constitui em um documento formal que estabelece um compromisso de como esses dados serão tratados durante todo o desenvolvimento do projeto, e também após a sua conclusão. [...] Posto que o PGD espelha uma situação dinâmica, é necessário observar que ele **não é um documento fixo no tempo**, [...] o PGD não é um documento burocrático e sua elaboração não deve ser pensada como uma mera tarefa administrativa [como o preenchimento de relatórios da Capes ou CNPq] na qual um texto padronizado possa ser utilizado para todos os projetos. **Ele deve ser tratado como uma carta de intenções que considere o que realmente é necessário para a preservação, compartilhamento e reuso dos dados.**

O PGD é um documento orgânico, ou seja, é um documento vivo, e, apesar de ser elaborado no início de um projeto de pesquisa, ainda poderá ser atualizado ao longo do processo de desenvolvimento da investigação, ganhando diferentes versões do início ao fim, sempre que necessário. O versionamento, também denominado controle de versões, é uma característica herdada da área de Ciência da Computação e Sistemas de Informação, as quais possuem a prática de nomear e salvar cada nova atualização de softwares e documentos técnicos. Portanto, o PGD pode ser entregue no começo de cada projeto; durante o seu andamento, se houver alterações

substanciais e significativas; no final, após a conclusão do projeto de pesquisa. (FOSTER, 2019).

Outrossim, documentar bem os dados de pesquisa contribui para a sua organização, preservação, descoberta e compartilhamento, além de reforçar as boas práticas na ciência. Além disso, o PGD publicado pode funcionar como comprovação, prova e registro para um pesquisador ou grupo e assim corroborar sua produção e primazia sobre determinados temas/temáticas.

Dessa forma, estabelecemos que os processos envolvidos com os referidos conhecimentos e serviços na gestão de dados podem ser representados em seis fases e aplicadas ao PGD (Figura 17). A figura a seguir representa todo o ciclo de vida e de ações a serem executadas na gestão de dados de pesquisa, inclusive este processo deve contar com o suporte do bibliotecário.

Nesse sentido, Sayão e Sales (2015) esclarecem que a gestão de dados científicos é uma responsabilidade compartilhada entre os pesquisadores e os bibliotecários, pois, ambos devem se ajudar mutuamente para realizar todas as etapas do processo e assegurar que todos os dados serão curados, tratados, organizados, disponibilizados e preservados.

Com a finalidade de descrever cada uma das seis fases dispostas na figura 17, na fase de desenvolvimento da proposta de pesquisa é realizado o primeiro contato com os responsáveis pelo arquivamento a fim de obter aconselhamento e consultoria sobre a elaboração do PGD e traçar as estratégias de preservação a longo prazo, assim como deve verificar a viabilidade dos recursos. Após a etapa anterior, dá-se início a segunda fase, onde começa a execução do projeto de pesquisa e as tarefas pertinentes à condução de testes, elaboração de termos, construção da metodologia etc.

A terceira fase corresponde à coleta dos dados e a organização destes, bem como dos arquivos gerados, estabelecendo estratégias para rotulação e nomeação dos arquivos. Nessa etapa é importante verificar as questões de segurança e integridade dos dados, verificação de falhas; utilização de padrões consolidados para os dados e metadados, tais como a criação de glossários, ontologias, taxonomias etc. A quarta fase diz respeito à análise dos dados; o gerenciamento dos arquivos criados; realização de *backups* e de toda a documentação da pesquisa.

Figura 17 - Fases da Gestão de Dados de Pesquisa aplicadas ao PGD



Fonte: Elaborado pela autora (2019), baseado em ICPSR (2012).

A quinta fase é a preparação dos dados para compartilhamento e compreende a divulgação de todos os riscos e limitações, informações sobre responsabilidades, formas de compartilhamento; atribuição de licenças; descrição dos formatos para o depósito dos dados etc. Por fim, na sexta fase é realizado o depósito dos dados em repositórios confiáveis. Cabe ressaltar que a ajuda de um bibliotecário é essencial em todas essas etapas, portanto, o pesquisador pode buscar aconselhamento em todas as fases ou quando sentir a necessidade.

Após a descrição as etapas apresentadas, vale ressaltar as vantagens em elaborar um PGD. Assim, elencamos a seguir, com base na literatura sobre o tema e diversos cursos sobre dados de pesquisa (EUROPEAN COMMISSION, 2016; FOSTER, 2019; SAYÃO; SALES, 2015), alguns motivos para elaborar um PGD:

- Tornou-se uma exigência das principais agências de fomento;
- Adequação do projeto de pesquisa às políticas mandatórias institucionais ou das agências de fomento;
- Assegura a integridade da pesquisa, permite a verificação e reprodutibilidade, viabiliza a reprodução de experimentos;
- Possibilita documentar as estratégias para o armazenamento, backup, manipulação, segurança e demais procedimentos e processos relacionados aos dados coletados, minimizando os riscos de perda ou dano;
- Visa à promoção do compartilhamento de informações e de dados a fim de que possam ser reutilizados por outros pesquisadores e assim gerarem novas pesquisas;
- Promove o compartilhamento de dados, a colaboração entre os pares e contribui para o avanço da ciência em todo o mundo;
- Representa economia de esforços, uma vez que racionaliza a aplicação de recursos e contribui para as boas práticas da Ciência Aberta;
- Evita a duplicação de trabalho;



- Facilita a busca por dados e permite compreender o seu contexto – se bem descrito e documentado – quando precisar usá-los / reusá-los oportunamente no presente ou futuramente;
- Garante a continuidade do projeto de pesquisa mesmo que uma ou mais pessoas deixem de fazer parte do grupo de trabalho, ou em casos em que novos integrantes entrem para a equipe;
- Permite a validação de resultados através dos dados de apoio ou como documentos suplementares em publicações científicas;
- Contribui para o aumento da segurança dos dados e outros registros de pesquisa a fim de garantir a sua inteireza, completude, autenticidade e confiabilidade;
- Amplia o alcance, impacto e visibilidade de um estudo e do pesquisador;
- Aumenta a eficiência do pesquisador, uma vez que tenha organizado e armazenado seus dados, além de tê-los documentado, estará mais focado na pesquisa e capacitado para localizar, compartilhar e usar seus dados com os outros colaboradores;
- O compartilhamento de dados contribui para que outras pessoas citem os dados (Os autores ganham mais citações com suas publicações e com os dados);
- Facilita a preservação digital e o arquivamento;
- Reduz custos e representa economia, uma vez que os dados foram devidamente coletados, selecionados, processados e organizados para serem depositados em um repositório de dados confiável.

Para a elaboração de um PGD, devem constar as seguintes informações básicas (EUROPEAN COMMISSION, 2016; FOSTER, 2019; SAYÃO; SALES, 2015):

- Que tipo de dados serão gerados na pesquisa;
- Quais são os formatos de arquivos utilizados;
- Como serão nomeados os arquivos;

- Como serão coletados os dados;
- Como serão processados os dados;
- Que tipo de documentação e metadados serão descritos e integrados aos dados;
- Qual a quantidade / volume de dados a pesquisa irá gerar;
- Qual o custo previsto com os dados e sua gestão;
- Como serão cobertos os custos;
- Se a pesquisa irá utilizar dados existentes e provenientes de outras fontes;
- Que tipos de dados serão mantidos e preservados;
- Quais os níveis e tipos de acesso serão implementados para a segurança dos dados;
- Qual a estratégia de gestão e manutenção dos dados a curto prazo (estratégias de manutenção, frequência de backups durante o projeto);
- Qual a estratégia de gestão e manutenção, disponibilização e preservação dos dados a longo prazo;
- Quem são os responsáveis pela gestão dos dados;
- Se há questões éticas e de privacidade associadas aos dados;
- Se há questões de propriedade intelectual, copyright etc.;
- Qual o tipo de licença será aplicado aos dados;
- Como serão arquivados e compartilhados os dados;
- Quais são as obrigações impostas ao público para compartilhar os dados;
- Quais são os usuários em potencial, público-alvo a que se destinam os dados de pesquisa produzidos e de que forma podem utilizá-los;

- Como citar os dados<sup>101</sup>.

Nesse sentido, o PGD deve conter a descrição completa dos dados do projeto, como será realizado o backup dos arquivos e dados, quais os níveis e controles de acesso, como será realizado o controle de versões, como os dados serão armazenados, de que modo será feita a curadoria desses dados, como serão compartilhados, além de informar todos os detalhes de quem faz parte da equipe de pesquisa, entre outros detalhes.

O PGD ajuda outros pesquisadores a entender como foram coletados os dados antes, durante e depois de uma pesquisa; em qual contexto estão inseridos; qual a sua ligação com as análises produzidas ao longo do estudo e que contribuirão para fazer sentido quando for publicado o trabalho final, produto destes dados coletados.

Existem ferramentas online gratuitas para fazer um PGD, como o DMPTool<sup>102</sup>, desenvolvido pela *University of California Curation Center*; e o DMPonline<sup>103</sup>, criado pelo *Digital Curation Centre (DCC)*. O DMPonline possui um checklist<sup>104</sup> de cada etapa para que o pesquisador verifique se todos os passos foram seguidos. Há outras ferramentas online para elaborar um PGD, tais como o DMP OPIDoR<sup>105</sup> (França), DMP Assistant (Canadá)<sup>106</sup> e mantido pela *Canadian Association Research Libraries (CARL)* e componente da rede de gestão compartilhada de dados de pesquisa Portage<sup>107</sup>, DMP Tuuli<sup>108</sup> (Finlândia, Projeto Tuuli), Pagoda – PGDonline<sup>109</sup> (tradução espanhola do DMPonline), e como exemplo de template de PGD institucional citamos o da Delft University of Technology (TUDelft)<sup>110</sup>, na Holanda. No Brasil, a Universidade Federal do Paraná (UFPR) disponibilizou o seu modelo de PGD em um formulário do

<sup>101</sup> Para orientações sobre a citação de dados, consulte as diretrizes elaboradas pela Force 11 Joint Declaration of Data Citation Principles: <https://www.force11.org/datacitationprinciples> e o esquema de metadados da DataCite: <https://schema.datacite.org/>.

<sup>102</sup> DMPTool: <https://dmptool.org/>.

<sup>103</sup> DMPonline: <https://dmponline.dcc.ac.uk/>.

<sup>104</sup> Checklist DMPonline (DCC): [http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP\\_Checklist\\_2013.pdf](http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf).

<sup>105</sup> OPIDoR: <https://dmp.opidor.fr/>.

<sup>106</sup> DMP Assistante: <https://assistant.portagenetwork.ca/>.

<sup>107</sup> Portage: <https://portagenetwork.ca/>.

<sup>108</sup> DMP Tuuli: <https://www.dmptuuli.fi/>.

<sup>109</sup> Pagoda – PGDonline: <http://dmp.consorciomadrono.es/>.

<sup>110</sup> Exemplo de template do PGD da TUDelft: [https://docs.google.com/viewer?url=https%3A%2F%2Fd1rkab7tqy5f1.cloudfront.net%2FLibrary%2FThemaportalen%2FRD%2F4TU\\_Data\\_Management\\_Plan.docx](https://docs.google.com/viewer?url=https%3A%2F%2Fd1rkab7tqy5f1.cloudfront.net%2FLibrary%2FThemaportalen%2FRD%2F4TU_Data_Management_Plan.docx). Ver também: [https://dmponline.tudelft.nl/?perform\\_check=false](https://dmponline.tudelft.nl/?perform_check=false).

Google<sup>111</sup>. A Universidade de Edimburgo possui mais de 21 modelos de PGD, um para cada curso, todos disponibilizados no DMPonline e em seu site institucional. Além disso, na Internet estão disponíveis vários modelos de PGD's públicos elaborados por pesquisadores de todo o mundo nos sites do DMPtool<sup>112</sup>, DMPonline<sup>113</sup> e também é possível encontrar alguns exemplos no repositório de dados Zenodo<sup>114</sup>.

As ferramentas mais recomendadas ainda são o DMPonline e o DMPtool por possuírem características e funcionalidades que permitem o registro livre e gratuito; autenticação federada (login e senha); disponibiliza modelos diferentes que atendem aos pré-requisitos de muitas instituições, áreas e disciplinas, agências financiadoras e de fomento; facilita o processo de preenchimento, pois fornece exemplos e mostra um painel de perguntas e respostas sugeridas; suporta múltiplas fases da elaboração de um PGD (Início, meio e fim); fornece permissões de leitura, escrita e compartilhamento para outros usuários colaboradores na elaboração do PGD, função ideal para equipes e acompanhamento de um orientador da pesquisa; e permite a exportação do PGD em diferentes formatos.

Na elaboração de um PGD é preciso considerar que tipos de dados serão produzidos na pesquisa. Assim, deve-se caracterizar os dados quanto à sua natureza, origem e processamento, tais como: observacionais, experimentais, brutos ou derivados, simulações, coleções físicas, modelos, software, imagens, vídeos etc. (SAYÃO; SALES, 2015).

Uma dúvida recorrente é sobre a quantidade de dados que será gerada em uma pesquisa, afinal, isso também deve ser descrito no PGD. A dica é se basear na hipótese da pesquisa e no plano de amostragem, assim, torna-se mais fácil fazer uma estimativa do volume de dados que serão gerados. (FOSTER, 2019; SAYÃO; SALES, 2015). Conforme dito anteriormente neste manuscrito, o PGD pode ser alterado sempre que necessário, portanto, não há problema se a estimativa da quantidade de dados mudar ao longo da pesquisa e em seu final.

Os dados que serão coletados precisam ser especificados e detalhados: informar quais foram os métodos utilizados para a obtenção dos dados; quem, o que,

---

<sup>111</sup> PGD UFPR: <https://is.gd/TESYRE>.

<sup>112</sup> PGD's públicos no DMPtool: [https://dmptool.org/public\\_plans](https://dmptool.org/public_plans).

<sup>113</sup> PGD's públicos no DMPonline: [https://dmponline.dcc.ac.uk/public\\_plans](https://dmponline.dcc.ac.uk/public_plans).

<sup>114</sup> Para ilustrar o exemplo, recomendamos consultar o Zenodo e também visualizar o PGD sobre Política de Informação para Gestão de imagens em movimento e acervos sonoros em repositório audiovisual, disponível em: <https://zenodo.org/record/2538427#.XJVgFCJKjX4>.

quando e onde foram obtidos; como foram selecionadas as amostras e quais foram os critérios; como serão analisados os dados e se é necessário algum processamento por software, algoritmos, códigos ou máquina específica; quais os instrumentos que serão usados etc. Para auxiliar em todo o processo de pesquisa e na tarefa de gerenciamento dos dados, além do que será descrito no PGD, o *DataOne* indica uma lista<sup>115</sup> com diversos softwares.

Outro detalhe crucial no PGD é declarar os formatos de arquivo dos dados e justificar a escolha, afinal, a obsolescência dos formatos, softwares e equipamentos é uma constante nos dias atuais, já que novas tecnologias surgem rapidamente a cada dia. É aconselhável o uso de formatos padronizados e não proprietários para o arquivamento dos dados a longo prazo, pois facilita a interpretação desses no futuro, independente de plataforma tecnológica (hardware e software). (FOSTER, 2019).

Algo importante, porém, desconsiderado por muitos é a maneira como os arquivos serão nomeados. Portanto, é adequado adotar uma convenção de nomes para arquivos, afinal, a ordenação das pastas é fundamental: facilita a organização, ajuda na localização dos dados armazenados, evita a perda, mudança ou reorganização dos arquivos durante a execução do projeto/pesquisa. Para auxiliar nessa tarefa, recomenda-se o uso de ferramentas para renomeação de múltiplos arquivos. Então, a fim de facilitar a descrição, a recuperação e a descoberta dos conjuntos de dados [*datasets*] é necessário definir como serão nomeados os arquivos.

#### **Exemplos:**

- **Memoria\_Locus\_CE\_2019\_DM\_MA.csv**

**Memoria\_Locus:** Nome do projeto;

**CE:** Abreviação do estado do Ceará;

**2019:** Corresponde ao ano civil de início e criação do arquivo;

**DM\_MA:** Representa como os dados foram extraídos (**Data Mining** = Mineração de Dados) e a partir de que tipo de instrumento / documento (**Mapas Afetivos**);

**csv:** Representa o tipo de extensão do arquivo (.csv), planilha com valores separados por vírgula e em caracteres ASCII.

---

<sup>115</sup> Lista de recursos indicados pelo DataOne: <https://www.dataone.org/all-software-tools>.

- **GF1\_USU\_2010-02-12.rtf**: Representa a transcrição de entrevista do primeiro Grupo Focal com os **usuários** da biblioteca, realizada em 12 de fevereiro de 2010).

Identificar quais são as garantias e medidas de controle de qualidade adotadas para os dados é essencial: devem ser incluídas as informações do que será feito no início, durante e depois da coleta dos dados, inclusive se a análise dos dados ainda estiver em andamento. Devem ser descritas as estratégias para garantir a segurança dos dados, privacidade, confidencialidade, anonimização e até mesmo a sua destruição segura, principalmente em pesquisas que geraram dados sensíveis. O delineamento dos possíveis usuários dos dados também não pode ser esquecido.

Pensar em uma estratégia para a eliminação segura e destruição dos dados é importante porque nem todos os dados poderão ser conservados. Isso exige, em grande medida, saber desde o começo ao fim, quais dados devem ser escolhidos para a preservação e quais serão excluídos. (FOSTER, 2019; SAYÃO; SALES, 2015).

A proveniência dos dados é outro fator relevante que deve ser descrito no PGD, assim, em caso de utilização de dados secundários existentes em outras fontes e que podem ser usados na pesquisa, deve-se identificá-los e determinar sua origem, por exemplo, *datasets* obtidos de repositórios de dados ou de portais de dados governamentais, como o Portal Brasileiro de Dados Abertos<sup>116</sup>.

Gerenciar dados científicos de uma pesquisa requer a descrição de uma estratégia de manutenção dos dados a curto e longo prazo. Então, inicialmente é preciso planejar e descrever como será mantido o controle de versões dos dados e das análises; inclusive deve-se informar como será realizado o *backup* dos dados de pesquisa e quem será o responsável (pesquisador ou instituição) e se há equipamentos destinados para esse fim. Em relação aos backups, as práticas recomendadas enfatizam que se deve ter no mínimo três cópias de segurança em pelo menos dois locais, por exemplo, a cópia original salva no computador, um backup salvo na nuvem e um salvo em outros dispositivos. As cópias devem ser distribuídas geograficamente a fim de reduzir riscos, tais como incêndio, falta de energia, danos aos equipamentos que armazenam os dados etc. Também é preciso testar os backups regularmente para verificar se estão funcionando, se os dados estão sendo

---

<sup>116</sup> Portal Brasileiro de Dados Abertos: <http://dados.gov.br/>.

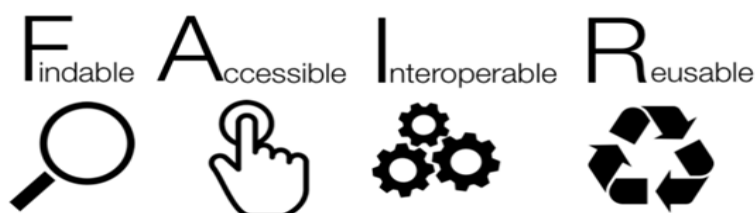
recuperados e se os arquivos estão abrindo perfeitamente. (ERWAY *et al.*, 2016; FOSTER, 2019; SAYÃO; SALES, 2015).

Após a descrição desses detalhes, também é preciso documentar e registrar a fase de depósito e disponibilização dos dados em um repositório de dados. Nesse sentido, além de cumprir todo o ciclo de gerenciamento dos dados, o pesquisador deve fornecer o máximo de informações possíveis sobre os dados de sua investigação, abordando ainda as questões de licenciamento, período de embargo, política de uso e reuso etc.

Silva (2016) justifica que uma gestão de dados bem documentada reduz esforços e a quantidade de trabalho necessária para a interpretação e compilação de informações obtidas no final da pesquisa, ademais, a investigação em andamento não precisará ser reconstituída em uma data posterior. Por isso, o autor defende que a assistência do bibliotecário prestada aos pesquisadores para avaliar e entender como organizar seus dados, como planejar sua pesquisa, além de subsidiar tomadas de decisões corretas sobre o acesso e preservação dos dados para os seus projetos.

Além de todas as etapas supracitadas, o pesquisador deve ter em mente que não basta se perguntar o que descrever e como redigir um PGD, também deve seguir certos princípios: FAIR<sup>117</sup> (Figura 18).

Figura 18 - FAIR Principles



Fonte: Sangya Pundir (2016).

FAIR, é um acrônimo para *Findable* (Fácil de encontrar), *Accessible* (Acessível), *Interoperable* (Interoperável) e *Reusable* (Reutilizável). Os princípios FAIR foram criados e propostos em um Workshop realizado na Holanda em 2014<sup>118</sup> e apoiados pela *Future of Research Communications and e-Scholarship* (FORCE 11).

<sup>117</sup> FAIR Principles: <https://www.force11.org/group/fairgroup/fairprinciples>.

<sup>118</sup> Workshop no Lorentz Center, em Leiden, Holanda: <https://www.lorentzcenter.nl/lc/web/2014/602/info.php3?wsid=602>.

Os resultados do estudo sobre os *FAIR Principles* foram publicados em 2016<sup>119</sup>. Abaixo apresentam-se os seguintes princípios expostos no quadro 3:

Quadro 3 - Princípios FAIR

(continua)

<b>Findable (Fácil de encontrar)</b>
O primeiro passo para (re)utilização de dados é conseguir encontrá-los. Os metadados e dados devem ser fáceis de encontrar para seres humanos e computadores. Os metadados legíveis por máquina são essenciais para a descoberta automática dos <i>datasets</i> e serviços relacionados.
F1. Os (Meta)dados recebem um identificador globalmente exclusivo e persistente. (Exemplo: Digital Object Identifier – DOI).
F2. Os dados são descritos com metadados ricos em informações (Este princípio está definido e interligado ao R1 mais adiante)
F3. Os metadados explicitam e incluem claramente o identificador dos dados que descrevem.
F4. Os (Meta)dados estão registrados ou indexados em um recurso pesquisável
<b>Accessible (Acessível)</b>
Quando o usuário encontra os dados necessários, precisa saber como eles podem ser acessados, possivelmente incluindo autenticação e autorização.
A1. Os (Meta)dados são recuperáveis pelo 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. Os metadados são acessíveis, mesmo quando os dados não estão mais disponíveis.
<b>Interoperable (Interoperável)</b>
Os dados geralmente precisam ser integrados a outros dados. Além disso, estes precisam interoperar com aplicativos ou fluxos de trabalho para análise, armazenamento e processamento.

<sup>119</sup> Artigo sobre os princípios FAIR publicado no periódico Scientific Data: <https://www.nature.com/articles/sdata201618>.



Quadro 3 - Princípios FAIR

(conclusão)

<b>Interoperable (Interoperável)</b>
I1. Os (Meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.
I2. Os (Meta)dados usam vocabulários que seguem os princípios do FAIR.
I3. Os (Meta)dados incluem referências qualificadas a outros (meta)dados.
<b>Reusable (Reutilizável)</b>
O objetivo final é otimizar a reutilização de dados. Para conseguir isso, os metadados e os dados devem ser bem descritos para que possam ser replicados e / ou combinados em diferentes configurações.
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 com proveniência detalhada
R1.3. Os (Meta)dados atendem aos padrões da comunidade relevantes ao domínio.

Fonte: GO FAIR (2016).

Com base nesses 15 princípios, um conjunto de 14 métricas foi definido para quantificar os níveis em que os dados estão de acordo com os princípios FAIR. Para auxiliar os pesquisadores a seguir esses princípios, foram desenvolvidas algumas ferramentas<sup>120</sup> para verificar, por meio de uma espécie de *checklist* ou questionário, os requisitos que os dados já possuem e assim sinalizar o que ainda precisa ser ajustado, atribuindo inclusive uma classificação com cinco estrelas, baseada também na classificação das cinco estrelas dos dados abertos proposta por Tim Berners-Lee (2006)<sup>121</sup>.

<sup>120</sup> Para conhecer as ferramentas, acesse: a) FAIR self-assessment tool, da Monash University em parceria com a Australian National Data Service; o National eResearch Collaboration Tools and Resources Project (Nectar) e a Australian Research Data Commons (ARDC): <https://www.and-s-nectar-rds.org.au/fair-tool/>; b) DANS FAIRdat Tool: <https://www.surveymonkey.com/r/fairdat/>; c) 5 Star Data Rating Tool: <http://oznome.csiro.au/5star/>, para obter mais informações sobre esta última ferramenta, veja também: <https://data.csiro.au/collections/#collection/CiCSIRO:27133v4/DItrue>.

<sup>121</sup> Para saber mais sobre a Classificação das 5 estrelas dos dados abertos: <https://5stardata.info/pt-BR/>.

Um ponto relevante sobre os dados é a atribuição de licenças, tais como a Creative Commons ou a Open Data License, afinal, estas são importantes para definir como os dados podem ser reutilizados.

Em suma, a gestão de dados científicos abrange desde a etapa de idealização da pesquisa, planejamento e elaboração de um plano de gestão de dados até as fases posteriores que envolvem a curadoria. A gestão de dados visa a garantia de dados organizados, tratados, curados e prontos para distribuição, uso e reuso, contemplando os requisitos dos dados FAIR (Findable - Fácil de encontrar, Accessible - Acessível, Interoperable - Interoperável, Reusable - Reutilizável).

### **2.4.3 O papel do bibliotecário na Ciência Aberta e gestão de dados**

Nos últimos anos, bibliotecários têm enxergado oportunidades na área da gestão de dados de pesquisa. De acordo com esse fato, Cunha (2010, p. 10) já alertava há anos que “[...] outra área do acervo informacional que tende a crescer no futuro é o chamado conjunto de dados científicos (ou ciência eletrônica, *e-Science*).” Conforme o referido autor, essa seria uma área composta por bases de dados numéricos, conjuntos de dados variados baseados em resultados das pesquisas realizadas em institutos, faculdades e departamentos.

Diversas pesquisas revelam que os bibliotecários têm ganhado espaço na gestão de dados de pesquisa nos últimos anos. (BURNETT, 2013; CORTI; EYNDEN; BISHOP; GREEN; MACDONALD; RICE, 2009; CUNHA, 2010; MORGAN-BRETT, 2011; PINFIELD, COX, SMITH, 2014; SAYÃO; SALES, 2014; SAYÃO; SALES, 2015; TENOPIR; BIRCH; ALLARD, 2012). Internacionalmente, há muitos relatos de experiências que tratam de questões pragmáticas e operacionais, sem se prender estritamente aos aspectos ligados às competências que o bibliotecário deve ter para atuar no gerenciamento de dados de pesquisa. Embora alguns abordem o assunto, a maior ênfase está na capacitação da equipe para lidar com esta demanda.

Vários relatos de experiência demonstram que os serviços de gestão de dados se iniciaram com orientações básicas sobre o PGD. Em outras experiências, foram realizados levantamentos e um diagnóstico institucional, seguido da fase de planejamento da instalação dos serviços e de infraestrutura necessária, por exemplo, um repositório de dados, como foi o caso da Universidade Federal do Paraná (UFPR).

Aos poucos, as bibliotecas universitárias têm se engajado e voltado seus olhares para os dados de pesquisa, a exemplo das bibliotecas da Universidade de São Paulo (USP)<sup>122</sup>, da Universidade Federal do Paraná (UFPR)<sup>123</sup> e da Universidade Federal do Rio Grande do Sul (UFRGS)<sup>124</sup>, que instituíram serviços em gestão de dados científicos. Destacam-se ainda as ações desenvolvidas pelo Instituto Brasileiro de Ciência e Tecnologia (Ibict)<sup>125</sup>, ou seja, é um nicho ativo em território nacional que vem se consolidando.

Diversos autores corroboram que existe uma preocupação dos bibliotecários e das bibliotecas universitárias em trabalhar com a *e-Science* e com a gestão de dados, inclusive, demonstrando disposição para redefinir seus papéis, produtos e serviços de informação a fim de que possam atender à demanda emergente no campo da Ciência da Informação. (COSTA; CUNHA, 2014; CUNHA, 2010; HEY; TREFETHEN, 2005; LUCE, 2010; MARCUM; GEORGE, 2010; MARTÍNEZ-URIBE; MACDONALD, 2008; SALES, 2014; SAYÃO; SALES, 2015; SILVA, 2016).

Para Cunha (2010), a biblioteca passou a tomar para si a responsabilidade da gestão do conhecimento gerado no campus, além disso, busca conhecer os conteúdos e as estruturas dos recursos informacionais disponíveis nos laboratórios e gabinetes dos professores. Xia e Wang (2014), por exemplo, apontam as responsabilidades e competências do bibliotecário de dados para atuar no gerenciamento de dados de pesquisa das Ciências Sociais e elencam como requisitos que os profissionais tenham experiência em métodos e softwares estatísticos; Serviços de Bibliotecas Universitárias; Dados espaciais, Dados numéricos e informação governamental; Desenvolvimento de pesquisas acadêmicas; Gestão, supervisão e liderança; Conhecimento em Tecnologia da Informação (Web e programação); Serviços de dados; Ensino e treinamento; Arquivamento e preservação digital; Repositórios de dados; Documentação de dados e metadados, Gerenciamento de dados; Concessões de dados; Bases de dados; Análise de dados; Criação e coleta de dados; Compras e contratos relacionados a dados.

Conforme podemos observar, os autores supracitados fornecem evidências de que as principais qualificações exigidas dos bibliotecários de Ciências Sociais são

---

<sup>122</sup> Ações desenvolvidas pela USP: <http://www.sibi.usp.br/apoio-pesquisador/dados-pesquisa/iniciativas-universidade-sao-paulo-gestao-dados/>.

<sup>123</sup> Base de Dados Científicos da UFPR: <https://bdc.c3sl.ufpr.br/>.

<sup>124</sup> CEDAP: <https://cedap.ufrgs.br/jspui/>.

<sup>125</sup> Ibict Dataverse: <https://repositoriopesquisas.ibict.br/>.

competências técnicas, embora também tenham sido identificadas outras competências comportamentais, a exemplo de comunicação, gestão e liderança.

Nos estudos de Tenopir, Birch e Allard (2012), a biblioteca é o local ideal para apoiar à gestão de dados de pesquisa acadêmica, e trabalha em parceria com os laboratórios de pesquisa da universidade. Nesse caso, conforme os autores, o desafio imposto aos bibliotecários requer um perfil profissional capaz de mediar, estabelecer contatos, fazer ligações e alianças com os potenciais pesquisadores/produtores de dados, pelo fato de ser um profissional que tem a obrigação de sair do seu lugar comum e espaço físico de trabalho para ir até os laboratórios, departamentos, escritórios e gabinetes para a captação dos dados a fim de dar-lhes um sentido e inseri-los dentro de um contexto, possibilitando o tratamento, a curadoria, a preservação digital e a recuperação desses dados, ou seja, semelhante ao perfil de um bibliotecário *liaison*.

Sobre os *liaison librarians* (bibliotecários de ligação), este se insere no contexto universitário, possui atribuições variadas e um perfil altamente especializado, mas, sua principal característica é a de estabelecer um vínculo / ligação com o público a qual atende, a ponto de ser o contato principal entre os departamentos, os cursos e a biblioteca. Nesse contexto, os *liaison librarians* aproximam-se dos docentes e discentes de uma forma personalizada, não apenas identificando suas necessidades de informação ou encaminhando pesquisas especializadas, mas também propondo e colaborando com soluções para o engrandecimento da área de atuação do pesquisador, seja ele docente ou discente, inclusive na criação de novos serviços e produtos com a finalidade de atender as demandas de seus usuários. (HARDY; CORRAL, 2007; JENKINS, 2005; MON; HARRIS, 2011; SANTOS; LIMA; ARAÚJO, 2019; TENNANT; CATALDO, 2002).

Em outras palavras, constituem-se como fatores diferenciais para demonstrar a importância do bibliotecário na gestão de dados porque:

Os bibliotecários possuem habilidades únicas para organizar informações díspares com a finalidade de desenvolver um projeto de pesquisa e para encontrar fontes de referência [...] que podem ajudar os pesquisadores a compreender melhor as ferramentas que podem ser úteis, o contexto da investigação e o formato necessário para indexação dos dados. [...] O bibliotecário também pode ajudar os pesquisadores a fazer conexões com outras pessoas da mesma comunidade disciplina, as quais podem fornecer suporte com ferramentas e todas as informações necessárias sobre o tema em questão. Este é o ponto onde as habilidades do bibliotecário e sua

capacidade de encontrar informações são úteis, porque; apesar de não saber todos os repositórios ou padrões de metadados, o bibliotecário pode encontrar as respostas, graças ao seu conhecimento universal de fontes de informação. (SILVA, 2016, p. 394).

Sales (2014) afirma que quando se trata do papel do bibliotecário frente à *e-Science*, sua contribuição vai ainda além do desenvolvimento de ferramentas e técnicas de representação da informação, portanto, são os profissionais competentes para realizar todos os procedimentos relacionados à gestão de dados.

Sobre a curadoria de dados, Nielsen e Hjørland (2013) discutem o papel do bibliotecário e as responsabilidades das bibliotecas nesta atividade, além de apresentar alguns conceitos e tipologias sobre dados apoiados na tríade dados-informação-conhecimento. Os autores abordam a curadoria e gestão de dados de pesquisa na perspectiva de que as bibliotecas e profissionais da informação são desafiados a assumir um papel de liderança na curadoria, ademais, evidenciam que as qualificações dos profissionais da informação a ser exigida mudarão totalmente o seu papel, dando lugar a uma outra faceta como especialista em documentação científica ou em dados, e ainda como especialista em informação e conhecimento.

Outro ponto observado pelos autores a respeito da curadoria de dados é o fato de ser um processo diferente da indexação de um documento, portanto, os conceitos e ideias que os bibliotecários têm sobre o que é um documento devem ser repensados, e, para isso, os autores recorrem à obra seminal de Suzanne Briet intitulada 'O que é a documentação?' para apoiar o seu argumento.

Ainda sobre a curadoria de dados, a preservação digital, o acesso e os metadados, Soehner, Steeves e Ward (2010) explicam que são áreas da *e-Science* em que as bibliotecas são afinadas, então, dominam as respostas de pesquisa e as tarefas que envolvem desde a captura, a seleção dos dados de pesquisa, a representação, a preservação e a recuperação, pois fazem parte de seu cotidiano.

Por outro lado, Martínez-Urbe e Macdonald (2008) admitem que seja comum surgir o questionamento se este é mesmo um trabalho para bibliotecários. Os autores advertem que estes profissionais podem não se sentir confortáveis com o gerenciamento de materiais tão complexos e heterogêneos. Ainda assim, de acordo com os autores, há um grupo que apresenta maior predisposição para participar do novo *status quo*: o **bibliotecário de dados**. Destarte, é preciso estimular o desenvolvimento de competências profissionais para atuar como bibliotecário de

dados, desenvolver habilidades e instituir o cargo, levando-se em consideração o alto grau de especialização e competência exigida.

Nesse sentido, Costa e Cunha (2014, p. 203-204) asseveram que os bibliotecários em universidades ou instituições científicas e tecnológicas precisam se perguntar:

[...] como organizar esses dados produzidos nas universidades e centros de pesquisa? Quais serão as habilidades requeridas para o profissional da informação que deseja trabalhar com organização de dados de e-Science no Brasil? Como organizar esses dados de forma a disponibilizá-los em rede? Quais as necessidades de informações dos pesquisadores que trabalham com esses dados? Como irei me capacitar para gerir e organizar esses dados?

Diante disso, em 2012 a *Association of College & Research Libraries* (ACRL) realizou uma pesquisa e identificou que o envolvimento das bibliotecas universitárias com a curadoria de dados é uma das dez maiores tendências para o futuro das bibliotecas.

De acordo com dois estudos publicados pelos membros da *Joint Task Force on Librarians' Competencies in Support of E-Research and Scholarly Communication* (CALARCO; SHEARER; SCHMIDT; TATE, 2016; SCHMIDT; SHEARER, 2016) sobre os perfis de competências e habilidades que destacam o papel dos bibliotecários na Ciência Aberta, em particular, em dados de pesquisa, ressaltam, primeiramente, que as atividades desenvolvidas por bibliotecários em comunicação científica e acesso aberto se enquadram em quatro categorias, a saber:

- a) Serviços de publicação acadêmica;
- b) Serviços de repositórios de acesso aberto;
- c) Aconselhamento e consultoria sobre direitos autorais e acesso aberto;
- d) Avaliação de recursos acadêmicos.

Para os bibliotecários que atuam em gestão de dados de pesquisa, é necessário um nível de conhecimento sobre a maioria das atividades descritas acima. Particularmente, os bibliotecários devem ter uma ampla perspectiva e compreensão dos modelos tradicionais (comerciais) e acesso aberto de publicação, questões de propriedade intelectual e economia da publicação acadêmica (CALARCO; SHEARER; SCHMIDT; TATE, 2016). Basicamente, traduzimos essas competências e habilidades a seguir (quadro 4):

Quadro 4 - Competências do bibliotecário de dados

(continua)

<b>COMPETÊNCIAS DO BIBLIOTECÁRIO EM GESTÃO DE DADOS DE PESQUISA</b>			
<b>1. Serviços de publicação acadêmica</b>	<b>2. Serviços de repositórios de acesso aberto</b>	<b>3. Serviços de aconselhamento e consultoria sobre direitos autorais e acesso aberto</b>	<b>4. Avaliação de recursos acadêmicos</b>
<p><b>a) Conhecimento de:</b></p> <ul style="list-style-type: none"> <li>- Plataformas de publicação de acesso comercial e aberto;</li> <li>- Publicação de fluxos de trabalho e modelos operacionais;</li> <li>- Processos editoriais;</li> <li>- Padrões: identificadores de objetos digitais (DOI), números padronizados internacionais (ISBN e ISSN), URL persistente e opções de citação, como OpenURL e CNRI Handle;</li> <li>- Padrões mandatórios e requisitos dos financiadores;</li> <li>- Padrões de metadados e ferramentas de descoberta.</li> </ul>	<p><b>b) Conhecimento de:</b></p> <ul style="list-style-type: none"> <li>- Políticas e requisitos de acesso aberto;</li> <li>- Software de repositório, padrões de metadados e ferramentas de descoberta;</li> <li>- Formatos de dados, design de banco de dados, gerenciamento de dados, ferramentas de manipulação de dados.</li> </ul>	<p><b>c) Conhecimento de:</b></p> <ul style="list-style-type: none"> <li>- Problemas de direitos autorais e licenciamento relacionados ao conteúdo acadêmico, incluindo a tradição direitos autorais, bem como creative commons e outras licenças OA;</li> <li>- Políticas e requisitos de acesso aberto; Tendências e questões atuais em acesso aberto e comunicação acadêmica.</li> </ul>	<p><b>d) Conhecimento de:</b></p> <ul style="list-style-type: none"> <li>- Critérios de avaliação para revistas e outros recursos;</li> <li>- Bibliometria e teoria e prática de métricas alternativas (Altmetrics).</li> </ul>
<p><b>a) Compreensão de:</b></p> <ul style="list-style-type: none"> <li>- Tendências e questões atuais em acesso aberto e comunicação acadêmica;</li> <li>- Práticas de conservação de dados e preservação;</li> <li>- Problemas de licenciamento relacionados ao acesso aberto.</li> </ul>	<p><b>b) Compreensão de:</b></p> <ul style="list-style-type: none"> <li>- Tendências e questões atuais em acesso aberto e comunicação acadêmica;</li> <li>- Problemas de direitos autorais e licenciamento relacionados ao conteúdo acadêmico;</li> <li>- Práticas de conservação de dados e preservação.</li> </ul>	<p><b>c) Compreensão de:</b></p> <ul style="list-style-type: none"> <li>- Sistemas de publicação acadêmico tradicionais.</li> </ul>	<p><b>d) Compreensão de:</b></p> <ul style="list-style-type: none"> <li>Políticas e procedimentos de promoção e posse de professores;</li> <li>Avaliação institucional / interesses do planejamento na produção acadêmica.</li> </ul>

Quadro 4 - Competências do bibliotecário de dados

(conclusão)

<b>COMPETÊNCIAS DO BIBLIOTECÁRIO EM GESTÃO DE DADOS DE PESQUISA</b>			
<b>1. Serviços de publicação acadêmica</b>	<b>2. Serviços de repositórios de acesso aberto</b>	<b>3. Serviços de aconselhamento e consultoria sobre direitos autorais e acesso aberto</b>	<b>4. Avaliação de recursos acadêmicos</b>
<p><b>a) Capacidade para:</b></p> <ul style="list-style-type: none"> <li>- Gerenciar os serviços de software de publicação em acesso aberto, como os sistemas produzidos pela Public Knowledge Project: Open Project Journal Systems (OJS) e Open Monograph Press (OMP);</li> <li>- Trabalhar com a equipe de TI local para desenvolver capacidade e infraestrutura.</li> </ul>	<p><b>b) Capacidade para:</b></p> <ul style="list-style-type: none"> <li>- Gerenciar a plataforma do repositório e atualizar o software ao longo do tempo;</li> <li>- Trabalhar com pesquisadores no depósito de resultados de pesquisa no repositório;</li> <li>- Ligação com editores sobre questões relacionadas a políticas de arquivamento incluindo períodos de embargo e, quando aplicável, taxas de processamento de artigos.</li> </ul>	<p><b>c) Capacidade para:</b></p> <ul style="list-style-type: none"> <li>- Sensibilizar pesquisadores sobre o OA, incluindo questões práticas, como financiamento e de aderência à política aderência;</li> <li>- Fornecer aconselhamento aos professores e estudantes de pós-graduação sobre alternativas à transferência de direitos autorais para seus trabalhos acadêmicos originais.</li> </ul>	<p><b>d) Capacidade para:</b></p> <ul style="list-style-type: none"> <li>- Oferecer apoio e consultoria na escolha de repositórios para o depósito dos dados;</li> <li>- Realizar estudos para o diagnóstico institucional sobre os dados de pesquisa;</li> <li>- Avaliar o impacto e desempenho dos serviços de dados ofertados.</li> </ul>

Fonte: Elaborado pela autora (2019), baseado em Calarco; Shearer; Schmidt; Tate (2016).

Conforme as competências supracitadas, esse conjunto de conhecimentos exige do bibliotecário uma postura proativa, além de incluir a mediação da informação como um dos fatores principais para o desenvolvimento dessas atividades, tendo em vista que, no âmbito da Biblioteconomia e Ciência da Informação, a mediação tem sido trabalhada sob diferentes aspectos, consolidando-se gradativamente como aporte teórico da área.

Outrossim, Almeida Júnior (2009) explica que a mediação pode ser empregada de maneira implícita e explícita. Na mediação implícita, as ações se desenvolvem sem a presença física e imediata dos usuários. Exemplos disso são a seleção, a indexação, o armazenamento, tratamento e o processamento da informação. Já na mediação explícita, a presença do usuário é uma exigência e



condição essencial, mesmo que esta presença não seja física, a exemplo de comunicações mediadas por computador e a distância.

Inter-relacionando a mediação da informação, a Ciência Aberta e a exigência de competências específicas para a atuação do bibliotecário nesse cenário, é patente que

A mediação da informação [...] desperta um novo comportamento dos bibliotecários que se distancia de uma mera execução de tarefas técnicas e repetitivas, porque toma a CI no sentido social e intersubjetivo, na medida em que os fenômenos de informação e tecnologia evoluem. Assim, podemos ter uma mediação intencional, efetiva e transformadora que é, ao mesmo tempo, pautada nos princípios básicos da Biblioteconomia e da CI [...] A mediação não é neutra, não pode ser imparcial, mas, sim, intencional, o bibliotecário deve assumir seu papel, isto é, se posicionar perante a sociedade mostrando a que veio, e não simplesmente esperar que os usuários busquem a informação somente ao se depararem com uma necessidade informacional. (SANTOS NETO; ALMEIDA JÚNIOR, 2015, p. 365).

Em face do exposto, a postura do bibliotecário é a chave para determinar o tipo e o grau de mediação em sua relação com os usuários. Ademais, é preciso construir um posicionamento ativo frente às necessidades informacionais de sua comunidade, e para isso, o bibliotecário deve se ver como a peça fundamental nesse processo de ensino-aprendizagem, nas quais pode desenvolver atividades que tenham o potencial de ir além de uma simples capacitação, pois também estará estimulando o pensamento crítico, a autonomia e o empoderamento ao “[...] criar as possibilidades para a sua produção ou a sua construção [...]” (FREIRE, 1996, p. 24) do conhecimento. Reforçando essa ideia, Farias (2015, p. 107) argumenta que o bibliotecário “[...] precisa ainda libertar-se de ideias conservadoras e estáticas, para poder sentir as demandas de um grupo, se conscientizando de seu papel na sociedade perante a responsabilidade social da área em que atua.”

Destarte, evidencia-se que realizar a gestão de dados científicos tem grande relevância no contexto das instituições acadêmicas, afinal, reforça a capacidade de gestão e curadoria dos dados e informações (mediação implícita), de maneira que os recursos indispensáveis para atender as necessidades de pesquisadores do mundo inteiro possam ser viabilizados. Assim, Sayão e Sales (2015) argumentam que:

**A gestão de dados de pesquisa** – pela amplitude do seu alcance na ciência contemporânea e pelo seu valor como recurso informacional – **não é responsabilidade somente dos pesquisadores que criaram ou coletaram os dados**. Muitas pessoas estão envolvidas nos processos de pesquisa e têm papéis importantes na garantia da qualidade, integridade, proveniência e preservação dos dados. **Porém, o papel crucial ainda é do pesquisador.** [...] **Os bibliotecários estão bem posicionados para trabalhar com os**

**dados pelos seus conhecimentos em gestão de informação, metadados, descoberta de recursos, preservação digital, além disso, eles sempre estabeleceram uma relação longa e produtiva com os pesquisadores.** (SAYÃO; SALES, 2015, p. 5-6, grifo nosso).

Segundo Martínez-Uribe e Macdonald (2008, p. 276, tradução nossa), "Os bibliotecários de dados trabalham lado a lado com os pesquisadores, aconselhando sobre sistemas de busca de dados, formas de acesso e auxiliando na sua gestão". Na mesma linha de pensamento, Silva (2016, p. 395) corrobora que "[...] os bibliotecários podem prestar assistência aos investigadores, com o depósito de dados ou a oferta de instruções com um passo a passo para repositórios comuns com o auxílio de guias da biblioteca."

Schmidt *et al.* (2016) afirma que pelo fato de os bibliotecários gerenciarem diferentes tipos de informações e dados, acrescentando-se ainda que fornecem uma série de detalhes que podem ser explorados baseados em métodos e ferramentas digitais, também devem buscar o domínio de novas habilidades e ampliar suas competências, aprimorando-se em temas específicos, tais como a Ciência Aberta, o gerenciamento de dados de pesquisa, a curadoria digital, a ciência dos dados, aos métodos estatísticos etc.

Dessa forma, para atuar na gestão de dados científicos e no auxílio ao preenchimento do PGD, propomos que o bibliotecário deve ter as seguintes competências, de acordo com o quadro 5: domínio de tecnologias da informação e comunicação; postura inovadora para criar e melhorar serviços de dados implantados nas bibliotecas; formando parcerias e trabalhando em colaboração com outros setores, pessoas e instituições, reforçando uma vez mais o caráter '*liaison*' e mediador exigido na composição do perfil deste bibliotecário que lida com os serviços de gestão de dados, incluindo outros elementos da Ciência Aberta.

Quadro 5 - Competências do bibliotecário na gestão de dados de pesquisa

Competência	Descrição
<b>Habilidade para lidar com as Tecnologias da Informação e Comunicação (TIC)</b>	O bibliotecário deve saber utilizar eficientemente a tecnologia da informação disponível.
<b>Postura inovadora e empreendedora</b>	O bibliotecário deve ter a capacidade de melhorar os serviços de dados de acordo com as necessidades e mudanças em diversas áreas. Deve se manter atualizado, de olho nas tendências que surgem na profissão, além disso, deve saber buscar onde o conhecimento está disponível (redes), a fim de disseminar as informações relevantes para as pessoas chave na organização. Frequentemente consulta seus usuários e avalia sua área sobre as necessidades percebidas no campo, por exemplo, realizando pesquisas e levantamentos, usando instrumentos como questionários, entrevistas ou grupos focais. Contribui ativamente com a comunidade, realiza visitas e promove sessões de treinamento, webinars, conferências etc.
<b>Visão Holística</b>	O bibliotecário deve ter uma visão global da instituição em que atua, deve saber reconhecer que os dados produzidos são apenas parte do ciclo de vida da pesquisa científica, além de estar ciente da importância que os dados da pesquisa têm para a realização de novas pesquisas e para o progresso da ciência. Enxerga oportunidades de serviços de dados, informações e processos como parte de um todo maior no qual as decisões são tomadas.
<b>Habilidade para prestar consultoria</b>	O bibliotecário deve estar sempre preparado e disposto a prestar serviços de consultoria em sua instituição quando se fizer necessário. Tem a capacidade de lidar com perguntas habilmente. Sabe quando e como aconselhar, encaminhar perguntas específicas, auxiliar em dúvidas sobre o gerenciamento de dados (perguntas sobre formatos, documentação, armazenamento, citação de dados, identificadores persistentes, preservação digital, elaboração de planos de gestão de dados (PGD), questões sobre propriedade intelectual, licenças, privacidade, confidencialidade e requisitos de financiamento). Pode captar outras percepções de seus usuários e solicitar feedback sobre suas habilidades de consultoria com a finalidade de ajustar-se conforme suas necessidades informacionais.
<b>Habilidades cooperativas</b>	O bibliotecário deve ter capacidade para trabalhar em cooperação e colaboração com outras pessoas (funcionários, pesquisadores, alunos, instituições) a fim de melhorar a prestação, avaliação e criação de serviços. Pode sugerir e instituir um fórum ou comitê de usuários e de apoiadores sobre dados a fim de se comunicar, levantar e abordar temas importantes, tais como políticas de dados, direitos autorais e infraestrutura de informações, entre outros aspectos relevantes. Nesse sentido, o bibliotecário assume a responsabilidade por sua contribuição e pelas parcerias firmadas.

Fonte: Elaborado pela autora, baseado em *Research Data Netherlands*, Programa de curso do *Essentials 4 Data Support* (2019).

Diante do exposto, acreditamos que essas são as competências básicas que os bibliotecários devem ter para atuar na Ciência Aberta e no gerenciamento de dados, uma vez que estes profissionais necessitam mobilizar os conhecimentos já adquiridos durante a prática laboral e combiná-las com os novos saberes exigidos para obter êxito nesse novo cenário que se apresenta.

Prosseguindo nessa linha, o estudo realizado pela *United Nations Organization* (2014) comprova o quanto é primordial que todas as pessoas - inclusive o bibliotecário - desenvolvam um conjunto das seguintes competências: a *Information literacy* (competência em informação), *Statistical literacy* (competência em estatística) e as *technical skills* (habilidades técnicas/competências técnicas) para alcançar à *data literacy* (competência em dados), conforme o diagrama abaixo (Figura 19), afinal, como entender, trabalhar e desenvolver pesquisas com dados sem compreender o seu significado?

Figura 19 - Diagrama de Venn sobre Competência em dados



Fonte: Baseado em Data Revolution Group (2014).

Portanto, tanto o bibliotecário - com a finalidade de atender aos usuários pesquisadores em estudos que envolvam dados científicos -, assim como os seus usuários devem buscar desenvolver a competência em estatística e a competência em dados (*data literacy*), exigindo esforço, engajamento e disposição para incentivar o uso e reuso de dados, além de promover a capacitação e o aprendizado do seu

público-alvo. Destarte, a utilização de estatísticas e de dados governamentais abertos em pesquisas científicas representam confiança ao usar fontes fidedignas para a construção de um embasamento teórico. Aliás, se o bibliotecário interage com os pesquisadores da instituição em que atua e possibilita o desenvolvimento de atividades das mais tradicionais até as mais inovadoras para que acessem, usem, reutilizem e se apropriem da informação, também ele (bibliotecário) anseia por apropriar-se dela e dar-lhe sentido.

O pesquisador Milo Schield (2004, p. 6) defende que “A avaliação da informação é o elemento chave na competência em informação, na competência estatística e na competência em dados.” Por isso, é importante trabalhar essas três competências juntas porque estão inter-relacionadas. De fato, o autor tem fortes argumentos que comprovam isso na realidade, uma vez que, segundo ele, é difícil promover a competência em informação ou a competência em dados sem promover a competência estatística.

Baseado nisso, no quadro 6, propomos as competências que o bibliotecário deve ter para atuar na Ciência Aberta e na gestão de dados de pesquisa. A partir do quadro, elencamos as competências em dados, em estatística e em informação, além de outras características altamente exigidas para a sua atuação e no desenvolvimento de habilidades para análise, interpretação e visualização de dados. Essas habilidades também facilitarão o manuseio de ferramentas e softwares estatísticos para o processamento dos dados e a correta aplicação dos métodos mais adequados para a interpretação dos dados. Essas condições são fundamentais para aqueles que pretendem reusar dados oriundos de outras fontes que não sejam de pesquisa própria, por exemplo, utilizar *datasets* disponíveis em portais de dados governamentais.

Quadro 6 - Competências do Bibliotecário na Ciência Aberta

Competências do Bibliotecário na Ciência Aberta e Gestão de Dados Científicos		
<ul style="list-style-type: none"> <li>- Competência em Dados;</li> <li>- Competência em Estatística;</li> <li>- Competência em Informação;</li> <li>- Criação, produção e processamento de dados;</li> <li>- Gestão de Dados;</li> <li>- Curadoria de Dados;</li> <li>- Humanidades Digitais;</li> <li>- Compartilhamento de Dados (Uso e reuso);</li> <li>- Visualização de Dados;</li> <li>- Experiência e domínio de métodos e softwares estatísticos;</li> <li>- Análise, síntese e interpretação de dados;</li> <li>- Conhecimento sobre a natureza e tipologia dos dados;</li> <li>- Dados Abertos;</li> <li>- Dados de Pesquisa;</li> <li>- Dados Governamentais;</li> <li>- Linked Data;</li> <li>- Ciclo de Vida dos Dados;</li> <li>- Princípios FAIR;</li> <li>- Verificação e reprodutibilidade da pesquisa;</li> <li>- Verificação da disponibilidade dos dados;</li> <li>- Ferramentas de descoberta dos dados;</li> <li>- Domínio das Tecnologias e Informação e Comunicação;</li> <li>- Conhecimento em programação (linguagens e ferramentas específicas para o tratamento, refinamento dos dados e extração de insights);</li> <li>- Big Data;</li> <li>- e-Science;</li> <li>- Business Intelligence (BI);</li> </ul>	<ul style="list-style-type: none"> <li>- Self-Service BI;</li> <li>- Interoperabilidade;</li> <li>- Didática e capacidade para ministrar treinamentos, capacitações e outras formas de promover o compartilhamento de conhecimento para os usuários e sua equipe de trabalho;</li> <li>- Capacidade para prestar aconselhamento e consultoria;</li> <li>- Capacidade para trabalhar em equipe;</li> <li>- Capacidade em colaboração;</li> <li>- Preservação Digital;</li> <li>- Padrões de Metadados gerais e específicos;</li> <li>- Ontologias;</li> <li>- Taxonomias;</li> <li>- Plano de Gestão de Dados (PGD);</li> <li>- Conhecimento sobre requisitos de financiamentos;</li> <li>- Serviços de Bibliotecas Universitárias e Especializadas;</li> <li>- Análise de domínio;</li> <li>- Gestão do conhecimento;</li> <li>- Segurança da Informação e em dados;</li> <li>- Conhecimento sobre confidencialidade, privacidade e anonimização de dados;</li> <li>- Conhecimento sobre questões éticas em pesquisa;</li> <li>- Mecanismos de controle de acesso;</li> <li>- Repositórios de Dados;</li> <li>- Infraestrutura de dados;</li> <li>- Comunicação científica;</li> <li>- Divulgação científica;</li> </ul>	<ul style="list-style-type: none"> <li>- Redes Sociais (acadêmicas ou não);</li> <li>- Publicações científicas;</li> <li>- Enhanced Publications (Publicações ampliadas);</li> <li>- Legislação (Lei de Acesso à Informação - LAI; Lei de Proteção de Dados Pessoais; Marco Civil da Internet etc.);</li> <li>- Políticas e diretrizes de Informação;</li> <li>- Regimes de Informação;</li> <li>- Economia da Informação;</li> <li>- Identificadores digitais persistentes (Exemplo: DOI);</li> <li>- Identificadores persistentes para autores (Exemplo: Orcid);</li> <li>- Citação dos dados;</li> <li>- Almetrics;</li> <li>- Direito Autoral, Propriedade Intelectual e Patentes;</li> <li>- Licenças livres e flexíveis (Creative Commons - CC);</li> <li>- Licenças para dados (Open Data Commons - ODC);</li> <li>- Licenças para softwares livres;</li> <li>- Workflow;</li> <li>- Blockchain;</li> <li>- Conhecimento sobre o Movimento de Acesso Aberto;</li> <li>- Boas práticas em Ciência Aberta;</li> <li>- Princípios da Ciência Aberta;</li> <li>- Cartas, declarações, postulados, planos e outros documentos de apoio ao Acesso Aberto e Ciência Aberta;</li> <li>- Cultura Livre Digital.</li> </ul>

Fonte: Elaborado pela autora (2019).

Nesse panorama, o bibliotecário deve buscar conhecer tecnologias como a *Blockchain*<sup>126</sup> aplicada à ciência, sendo esta mais uma tendência no campo da Ciência Aberta. Trata-se de um protocolo de confiança que tem como principal objetivo a auditoria e descentralização como medida de segurança das informações e dos dados. Trata-se da mesma tecnologia usada em criptomoedas como o Bitcoin.

*Blockchain* é uma tecnologia para dados descentralizados e auto-regulados. Através de dados blockchain [a ciência] pode ser gerenciada e organizada de uma maneira nova e revolucionária: aberta, permanente, verificada e compartilhada, sem a necessidade de uma autoridade central. (DIGITAL SCIENCE; FIGSHARE, 2017, p. 2, tradução nossa).

Outrossim, a *Blockchain* permite avaliar aspectos críticos concernentes à comunicação científica tais como a transparência, reprodutibilidade, confiança e os devidos créditos autorais e de propriedade intelectual. Não é à toa que surgiram diversos projetos voltados para essa tecnologia nos últimos anos a fim de mitigar alguns problemas enfrentados na atualidade, por exemplo, a abertura da pesquisa científica, ética, segurança, transparência, acesso às informações científicas, auditoria, os custos envolvidos nas publicações científicas e revisão por pares.

Outro exemplo a ser citado é a preservação digital, atividade que visa a criação de políticas e estratégias de preservação, tais como a digitalização, migração, estudo e avaliação das tecnologias voltadas para essa aplicação, estabelecimento de padrões de interoperabilidade, emulação, implantação e manutenção de repositório apropriado, entre outras tarefas.

Além da preservação digital, o bibliotecário precisa se reinventar no que concerne à descrição de metadados e catalogação dos dados, afinal, realizar essa tarefa é mais complexa e não se compara a catalogação de um livro ou artigo de periódico. Descrever dados, oferecer-lhes um contexto, torná-los fáceis de encontrar em um sistema de busca tem sido um dos maiores desafios para os bibliotecários em todo o mundo, inclusive nos países em que estas atividades já estão bastante avançadas, como é o caso relatado no artigo de Nielsen e Hjørland (2013) e também por Cox, Kennan, Lyon e Pinfield (2017), os quais mostram que diante da diversidade de instituições produtoras de conhecimento em gestão de dados científicos na

---

<sup>126</sup> Transações financeiras com Bitcoins são conduzidas de maneira autorizada e criptografada, usando assinaturas digitais e transmitidas para a rede ponto a ponto. Essas transações são verificadas em pacotes por membros do público em geral que ajudam a validar e registrar a hora das transações. A validação gera "blocos" e ao final de uma cadeia de blocos semelhantes em intervalos regulares (aproximadamente a cada 10 minutos) são compartilhadas na rede. O uso da criptografia garante que todas as transações anteriores não foram alteradas, isto é, há um registro permanente de transações criado e mantido em cada nó gerado pelo participante, assegurando que não haja um único ponto de falha e nem uma única entidade controlando os dados.

Austrália, Canadá, Alemanha, Irlanda, Holanda, Nova Zelândia e Reino Unido: os autores relatam um estudo sobre as atividades, serviços e recursos de gerenciamento de dados de pesquisa em bibliotecas do ensino superior.

De acordo com os autores supracitados, nestes países, há a predominância de serviços em consultoria e suporte a planejamento de gerenciamento de dados e na realização de treinamentos, com forte apelo para a conscientização (*awareness*) e o *advocacy* para o compartilhamento e gestão de dados de pesquisa ao invés do foco em serviços técnicos como a curadoria de dados, a descrição e catalogação dos conjuntos de dados, a disponibilização de um catálogo de dados, entre outros, devido à complexidade e natureza deste trabalho.

Assim, como a habilidade para o aconselhamento e consultoria tem sido escolhida como a primeira estratégia para estabelecer serviços de dados a fim de auxiliar os pesquisadores, não seria diferente no que diz respeito a ajuda prestada pelos bibliotecários sobre a elaboração de um plano de gestão de dados.

De acordo com os pontos levantados expostos nesta seção, é possível perceber os desafios que os bibliotecários têm que enfrentar para atender a demanda de gestão de dados científicos, tendo em vista que é preciso capacitar-se, buscar a união dos conhecimentos em Biblioteconomia, Ciência da Informação e do cotidiano já adquiridos e incorporá-los às práticas e processos da gestão de dados, pois, somente dessa maneira será possível advogar pela criação e suporte a tais serviços. Evidentemente, há as questões estruturais e de infraestrutura que devem ser pensadas e planejadas antes da efetiva implantação do rol de serviços, mas antes se faz necessário que os profissionais se preparem e se capacitem para treinar a comunidade acadêmica e replicar estes conhecimentos.



### 3 METODOLOGIA

Serão apresentados nesta seção o percurso metodológico e todas as escolhas que sustentaram a abordagem da pesquisa, desde a seleção do método, dos sujeitos, o ambiente de pesquisa, até a exposição dos procedimentos de coleta e análise dos dados. Desta maneira, adotamos dimensões que venham a priorizar as experiências e a percepção dos indivíduos a respeito do fenômeno investigado.

Com a finalidade de cumprir os ritos exigidos na academia e para atender aos requisitos básicos da ética e integridade da pesquisa científica, além de ser um estudo que envolve a participação de seres humanos, esta pesquisa seguiu os ditames da Resolução n.º 466 de 12 de dezembro de 2012<sup>127</sup> e da Resolução n.º 510 de 7 de abril de 2016<sup>128</sup>, ambas do Conselho Nacional de Saúde, tendo sido submetida ao Comitê de Ética em Pesquisa da Universidade Federal do Ceará via Plataforma Brasil sob o número 18708719.0.0000.5054, em seguida foi apreciado e aprovado conforme o parecer n.º 3.512.152. (Vide Anexo A). Solicitou-se ainda a cada um dos entrevistados que assinasse o Termo de Consentimento Livre e Esclarecido (TCLE), documento que formaliza a participação do sujeito de pesquisa, assim como informa sobre os riscos e benefícios da aplicação deste estudo (Apêndice E).

Outra medida adotada para documentar esta pesquisa diz respeito a elaboração de um Plano de Gestão de Dados (PGD)<sup>129</sup>, documento que descreve como este estudo foi planejado, conduzido, como os dados coletados serão usados, preservados e disponibilizados no futuro. (Vide Apêndice F).

#### 3.1 Caracterização da Pesquisa e Processo de construção metodológica

Quanto a sua **natureza**, esta é uma **pesquisa aplicada**, pois gera novos conhecimentos visando a aplicação prática. Quanto à forma de **abordagem**, desde já esclarecemos que será dada maior ênfase no aspecto qualitativo, portanto, trata-se de uma **pesquisa predominantemente qualitativa**, entretanto, alguns dados numéricos serão apresentados de forma sucinta. A **abordagem quantitativa** é objetiva, lida com a **quantificação dos dados**, traduz opiniões e números em

---

<sup>127</sup> Resolução n.º 466/2012 do Conselho Nacional de Saúde: <http://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf>

<sup>128</sup> Resolução n.º 510/2016 do Conselho Nacional de Saúde: [http://bvsmis.saude.gov.br/bvs/saudelegis/cns/2016/res0510\\_07\\_04\\_2016.html](http://bvsmis.saude.gov.br/bvs/saudelegis/cns/2016/res0510_07_04_2016.html).

<sup>129</sup> A primeira versão do PGD dessa dissertação está disponível em: <https://dmponline-test.dcc.ac.uk/plans/41640/export.pdf>. A segunda versão está disponível em: <https://dmponline.dcc.ac.uk/plans/41640/export.pdf>.

informações que poderão ser analisadas e classificadas, dando suporte aos achados e abordagem qualitativa.

Na **abordagem qualitativa**, são estabelecidas relações que não podem ser quantificadas, ou seja, não se traduzem em números. Portanto, a pesquisadora **aprofunda-se em questões subjetivas**, coletando os dados por meio de entrevistas, observações, narrativas e documentos. As opiniões também são analisadas em pesquisas qualitativas. As amostras não utilizam métodos estatísticos, não são representativas nesta abordagem, possuem caráter exploratório e são feitas em pequenas amostras.

Quanto aos seus **objetivos**, esta pesquisa apresenta um caráter **descritivo-exploratório**, pois visa a descrição sobre o fenômeno a partir do ponto de vista dos participantes (KOZINETS, 2006), mas também é um estudo organizado de acordo com as categorias que emergem dos dados e servem como base para a construção de teorias e hipóteses. (GLASER; STRAUSS, 1967). Para Prodanov e Freitas (2013) são os tipos de estudo habitualmente realizados por pesquisadores sociais que têm uma preocupação com a atuação prática. Na visão de Gil (2010), o propósito das pesquisas exploratórias é proporcionar maior familiaridade com o problema pesquisado a fim de torná-lo mais explícito ou para construir hipóteses. Já as pesquisas descritivas, para o referido autor, são indicadas para levantar opiniões, crenças e atitudes de uma determinada população.

Quanto aos **procedimentos técnicos**, esta é uma **pesquisa participante** em que se pretende realizar uma **intervenção (ação de informação**, também denominada como **ação educacional** ou **ação de educação** em outras áreas do conhecimento). Para Haguette (2010), considera-se a pesquisa participante como uma pesquisa alternativa se comparada a outros métodos. Pode ser entendida ainda como uma técnica de observação participante que se consolidou na década de 80. Trata-se de um processo que envolve a investigação, educação, ação e organização. Para a autora, “[...] a ação é a fonte de conhecimento” (HAGUETTE, 2010, p. 158), logo, a autora reforça que a marca da pesquisa participante é a interação entre pesquisador e os sujeitos investigados. Ademais, acrescenta que é um tipo de pesquisa que se diferencia dos demais métodos pela sua capacidade de resolver problemas de grupos sociais.

Para compor a fase de planejamento da pesquisa e na escolha das estratégias de entrada no campo e quais instrumentos de coleta e técnicas devam

ser utilizadas, pensou-se na realização de um levantamento e de instrumentos que possam captar opiniões, percepções, entender comportamentos e necessidades informacionais, assim, para o prosseguimento da segunda fase do estudo, optou-se pela adoção de um diário de campo eletrônico para o registro de observações no campo de pesquisa e um caderno de laboratório eletrônico de uso técnico para elaborar uma síntese de conteúdos e links a serem disponibilizados *a posteriori*; aplicação de questionário e entrevista semiestruturada; transcrição das entrevistas e separação do material para a terceira fase (análise e comparação constante); e finalmente a preparação do material para a realização de um minicurso online como uma ação de informação / intervenção e a coleta por meio de anotações de observações e aplicação do questionário de avaliação do minicurso. Em linhas gerais, este estudo empregou a combinação de diferentes técnicas que se contrapõem, mas que se complementam, com a função de dar credibilidade e validade aos resultados da pesquisa.

Como um dos procedimentos de pesquisa, utilizou-se como técnica a pesquisa documental, parte de um dos processos da Teoria Fundamentada, empreendida para a análise da documentação desenvolvida e que ainda não recebeu tratamento analítico, tais como os registros observacionais, notas de campo, memorandos, entre outros materiais. Executada a partir de documentos considerados cientificamente autênticos e de natureza diversificada, a pesquisa documental se constitui em uma importante técnica usada na pesquisa qualitativa e auxilia na complementação das informações colhidas por outras técnicas e revela aspectos essenciais dos fenômenos estudados (LÜDKE; ANDRÉ, 1986).

Nesse seguimento, a pesquisa adota uma abordagem construtivista, à medida que se identifica com a realidade múltipla e construída, ressaltando que o sujeito e o objeto de observação são inseparáveis dos fatos sociais (BECKER, 1994; DUARTE, 2009), de forma que, entender o contexto social requer um estudo qualitativo.

Thiollent (2009) afirma que a pesquisa empírica se ocupa da descrição de situações concretas, para intervenção ou ações orientadas em função de problemas detectados na coletividade considerada. Meksenas (2007) enfatiza que as pesquisas empíricas são realizadas no espaço social, afinal, lidam com processos de interações face a face e demandam que o pesquisador se insira no campo e se insira no lugar em que se dão as relações de pesquisa dos sujeitos-pesquisados.

[...] Trata-se de compreender precisamente como **a pesquisa pode ser criadora**, como pode fazer aparecer não apenas novos resultados mas novos métodos, como nela a própria **idéia da cientificidade** pode assumir pouco a pouco **contornos** cada vez **mais precisos**, como pode **controlar** cada vez melhor **a si mesma** de modo a se tornar **progressivamente mais eficaz**, a se **adaptar** de modo cada vez **mais estreito a seu objeto**. (LADRIÈRE, 1977, p. 17, grifo nosso).

Por ser uma pesquisa empírica, envolve a coleta direta de dados de um grupo de informantes em dada realidade social concreta; a utilização de métodos qualitativos e quantitativos, mantendo-se a coerência do estudo e de seus objetivos, além do cuidado em estar ancorado em um fundamentado quadro referencial teórico no sentido de que “Não nos parece haver incompatibilidade no fato de progredir na teorização a partir da observação e da descrição de situações concretas [...]” (THIOLLENT, 2009, p. 9).

Sobre a observação supracitada, Bruyne, Herman e Schoutheete (1991, p. 16) explicam perfeitamente que “A pesquisa social é sempre tateante, mas ao progredir elabora critérios que lhe permitem orientar-se de modo cada vez mais preciso e que, aliás, ela não pára [sic] de aperfeiçoar confrontando de modo crítico os métodos utilizados e os resultados.”

Em relação à análise e interpretação dos dados, esta pesquisa segue na linha da abordagem qualitativa e interpretativa, na qual segundo Flick (2009), utiliza como material empírico textos ao invés de números, tendo em vista que faz parte da realidade em estudo e tem interesse na perspectiva dos participantes, por seus conhecimentos e práticas rotineiras relacionadas ao objeto de estudo.

Nessa perspectiva, na abordagem qualitativa os instrumentos de trabalho utilizados no campo permitem uma mediação entre o marco teórico-metodológico e a realidade empírica investigada (MINAYO, 2008). Assim, foram adotados métodos (*Grounded Theory* e Netnografia) e técnicas de pesquisa (Pesquisa documental), além de instrumentos para a coleta dos dados (entrevistas, questionários e anotações de observações em diário de campo eletrônico) e posterior análise visando a obtenção de resultados que permitam complementar a proposta de criação de um programa de apoio e suporte aos pesquisadores (professores e alunos de pós-graduação) da Universidade Federal do Ceará, assim como a elaboração de diretrizes e a sugestão de criação dos serviços de informação (Programa de apoio e suporte ao pesquisador) voltados para a Ciência Aberta e gestão de dados de pesquisa.

Exatamente por isso, partiu-se do princípio de que a realização inicial de um levantamento sobre a percepção dos pesquisadores é necessária para investigar o grau e estágio de compreensão sobre o tema que estes pesquisadores possuem, assim como será útil para a identificação das lacunas e necessidades informacionais latentes e ainda desconhecidas.

Para tanto, conforme Morse (1994), assim como diferentes "lentes" ou perspectivas resultam do uso de métodos distintos, muitas vezes, mais de um método pode ser usado dentro de um projeto para que o pesquisador possa ter uma visão holística do cenário. Então, recorreremos à triangulação metodológica, de acordo com Triviños (1987, p. 138) “[...] tem por objetivo básico abranger a máxima amplitude na descrição, explicação e compreensão do foco em estudo.” Na triangulação, simultaneamente ocorrem as etapas de definição dos procedimentos metodológicos, pré-testes, coleta e análise dos dados (HERNÁNDEZ SAMPIERI; FERNÁNDEZ COLLADO; BAPTISTA LUCIO, 2013).

A triangulação não é nova, teve início na Psicologia e tinha como proposta complementar ou testar empiricamente os resultados obtidos em pesquisa utilizando diferentes técnicas quantitativas. Mais tarde, sua aplicação foi ampliada com a justificativa de que a obtenção dos dados oriundos de diferentes fontes e análise devem recorrer a estratégias variadas porque melhoraria a qualidade dos resultados (CAMPBELL; FISKE, 1959; DUARTE, 2009). Há quatro tipos de triangulação segundo Duarte (2009), Jensen e Jankowski (1993):

- a) Triangulação de dados: trata de diferentes dimensões de tempo, espaço e analítica que serão utilizadas pelo pesquisador na busca por informações para a sua investigação;
- b) Triangulação de investigador: trata-se da formação de equipes de pesquisadores de várias áreas do saber;
- c) Triangulação de teoria: propõe abordagem empírica do objeto por meio de perspectivas conceituais e teóricas distintas;
- d) Triangulação de métodos: quando são utilizados métodos diferentes de investigação para coletar e analisar os dados do objeto de estudo em questão.

Segundo Günther (2006), a triangulação utiliza distintas abordagens metodológicas do objeto empírico a fim de prevenir possíveis distorções relacionadas

à aplicação de um único método, uma única teoria ou um pesquisador. Assim sendo, “o uso de múltiplos métodos, ou da triangulação, reflete uma tentativa de assegurar uma compreensão em profundidade do fenômeno em questão.” (DENZIN; LINCOLN, 2006, p. 19).

As vantagens da triangulação, expostas por Denzin e Lincoln (2006) são a exposição simultânea de realidades múltiplas; possibilita aos leitores explorar visões semelhantes ou contrastadas do contexto, facilitando a compreensão da realidade apresentada. Diante do exposto, Minayo e Minayo-Goméz (2003, p. 136) asseveram que

[...] nenhum método pode se arrogar a pretensão de responder sozinho as questões que a realidade social coloca. [...] [A triangulação] pode iluminar a realidade a partir de vários ângulos, o que permite confluências, discordâncias, perguntas, dúvidas, falseamentos, numa discussão interativa e intersubjetiva na construção e análise dos dados.

Para a triangulação neste estudo aplicar-se-á a Teoria Fundamentada em Dados e a Netnografia como métodos de pesquisa. De acordo com registros da literatura, a Teoria Fundamentada costuma utilizar como métodos e instrumentos de coletas de dados aliados a pesquisa bibliográfica, análise documental, análise textual, entrevistas, observação participante e diários de campo. (CHARMAZ, 2009; GLASER; STRAUSS, 1967; STRAUSS; CORBIN, 2008). Utilizaremos a observação participante como técnica de coleta dos dados na netnografia, de forma que o pesquisador busque tornar-se membro do grupo ou organização estudada. Procedimentos aliados são a pesquisa documental, por meio da análise de documentos, notas de campo, memorandos e outros tipos na descrição e análise dos dados coletados, os quais já fazem parte da Teoria Fundamentada, como ilustra a figura 20 e conforme será explicado em detalhes nas seções posteriores.

Figura 20 - Triangulação metodológica aplicada à pesquisa



Fonte: Elaborado pela autora (2019), baseado em Flick (2009).

Com a finalidade de analisar e interpretar os resultados gerados no ambiente virtual de aprendizagem (*Google Classroom*) e comparar com os resultados obtidos em entrevistas e questionários a fim de estabelecer correlações e na extração de categorias, optamos por combinar a Teoria Fundamentada em Dados e a Netnografia. Estas são metodologias complementares, ambas têm raízes no interacionismo simbólico e representam uma combinação perfeita para descrever e explorar o fenômeno em questão. De acordo com Haguette (2010), por serem métodos de perspectiva interacionista, emergem da interpretação, da percepção ou do significado que determinada situação ou objeto tem para um ator social.

A Teoria Fundamentada também é conhecida como uma abordagem ou método geral de análise comparativa e estabelece um conjunto de procedimentos

sistemáticos capazes de gerar uma teoria embasada nos dados. (GLASER; STRAUSS, 1967; TAROZZI, 2011). Em outras palavras, “[...] trata-se do modo de construir indutivamente uma teoria assentada nos dados, através da análise qualitativa destes e que, agregada ou relacionada a outras teorias, poderá acrescentar ou trazer novos conhecimentos à área do fenômeno.” (CASSIANI; CALIRI; PELÁ, 1996, p. 78).

Nesse sentido, Kathy Charmaz (2006) considera a *Grounded Theory* como uma ‘constelação de métodos’, além de enfatizar que a qualidade e credibilidade de um estudo começa com os dados. Esta é considerada a mais completa metodologia de pesquisa qualitativa disponível porque valoriza a compreensão das ações através do ponto de vista do próprio agente (HAIG, 1995).

Outro traço marcante na Teoria Fundamentada é o tipo de amostragem e a seleção dos sujeitos de pesquisa: o procedimento típico utilizado é a **amostragem teórica**. Esse tipo de amostragem na pesquisa qualitativa é completamente diferente da amostragem probabilística estatística porque não se refere ao número de participantes em relação ao universo da população em estudo, afinal, a amostra não se forma *a priori*, mas sim ao longo da pesquisa, como parte do processo analítico e vai aumentando a sua extensão progressivamente conforme as exigências do trabalho de conceituação teórica no decorrer da análise, do número de participantes e suas características até a saturação dos dados.

Ademais, pode ser aplicada em qualquer área, em inúmeros fenômenos e tem abertura o bastante para o pesquisador interpretar o contexto de onde emergem os dados, isto é, até mesmo a ação desenvolvida pelo pesquisador é englobada como parte do processo de formação da teoria (PIDGEON, 1996). Pesquisas em Ciência da Informação também já adotaram a Teoria Fundamentada em suas investigações, a exemplo dos trabalhos de Gama (2013) sobre competência informacional, Berrío-Zapata (2015) sobre a análise de domínio e Costa (2017) sobre diretrizes para o desenvolvimento de uma política de gestão de dados científicos no Brasil.

Esta metodologia foi escolhida em virtude da complexidade do tema investigado, tornando difícil o enfoque em apenas um aspecto. Destarte, o método permite que as teorias surjam dos dados coletados em campo e passem por processos de comparação, ordenamento, classificação e codificação. Aliás, Tarozzi (2011) destaca que o método da comparação constante é o coração da Teoria Fundamentada. Por sua natureza cíclica, favorece o alcance dos objetivos almejados



porque permite ao pesquisador iniciar nova coleta de dados e fazer análises comparativas até a saturação teórica dos dados ou até que não apareçam mais novos elementos, conforme descrito na figura 21 abaixo.

Figura 21 - Ciclo da Teoria Fundamentada



Fonte: Baseado em Gama (2013).

Como apresentado na figura acima, a Teoria Fundamentada em dados é um complexo processo de análise. Nessa perspectiva, entendemos que a complexidade é um desafio que nos motiva a realizar esta pesquisa. Por esse prisma, Morin (2001, p. 38) afirma que:

*Complexus* significa o que foi tecido junto; de fato, há complexidade quando elementos diferentes são inseparáveis constitutivos do todo (como o econômico, o político, o sociológico, o psicológico, o afetivo e o mitológico), e há um tecido interdependente, interativo e retroativo entre o objeto de conhecimento e seu contexto, as partes e o todo, o todo e as partes, as partes entre si. Por isso a complexidade é a união entre a unidade e a multiplicidade.

Além disso, é um método que exige intenso envolvimento do pesquisador desde o início ao fim da pesquisa. Assim, Morin (1998, p. 335) enfatiza que “[...] a teoria não é o fim do conhecimento, mas, um meio-fim em permanente recorrência.”

A Teoria Fundamentada foi proposta como método em 1967, inicialmente pelos sociólogos americanos Glaser e Strauss, no livro intitulado “*The Discovery of Grounded Theory: Strategies for qualitative Research*”. A *Grounded Theory* inspira-se nas vertentes teórico-metodológicas do Pragmatismo e do Interacionismo-Simbólico, busca sistematizar as formas de perceber a teoria a partir do campo empírico de maneira exclusiva e como contraponto aos métodos tradicionais. (GOULDING, 1999).

O interacionismo-simbólico prioriza a visão da participação das pessoas como agentes ativos em oposição a ideia de que os sujeitos são partes de um grande organismo que sofrem passivamente a ação de forças externas. Assim, os interacionistas-simbólicos consideram a sociedade como um caleidoscópio em constante mutação de indivíduos interagindo uns com os outros, na medida em que a natureza das interações muda, também a sociedade estará em constante mudança, portanto, essa é uma abordagem mais dinâmica do estudo da vida social (ANGROSINO, 2009).

Então, a ideia central da *Grounded Theory* funda-se no campo empírico, na teorização emergente a partir de observações, constantes comparações, classificações e análises sistemáticas dos dados coletados, processo este denominado como codificação. A partir dessas análises sistemáticas o pesquisador deve elaborar memorandos teóricos, uma espécie de anotação, desenho e até mapas mentais que o ajudarão futuramente na teorização. Portanto, “A teoria, assim, emerge dos dados e não é construída pela reflexão teórica, construção de hipóteses e posterior verificação no campo empírico.” (FRAGOSO; RECUERO; AMARAL, 2011, p. 85). Complementando a afirmação anterior, Pidgeon (1996) reforça que na Teoria Fundamentada o pesquisador tem muita importância porque a partir de sua sensibilidade e percepção subjetiva é que os dados emergirão.

A Teoria Fundamentada “[...] é um método com perspectiva bastante interessante para quem trabalha com muitos dados empíricos, pois proporciona uma certa liberdade para lidar com esses dados.” (FRAGOSO; RECUERO; AMARAL, 2011, p. 83). Para as autoras citadas, o mote da *Grounded Theory* é que a teoria deve emergir dos dados, por meio de observação sistemática, comparação, classificação e análise de aspectos similares e diferentes entre si. Por seu turno, prevê a inversão

radical do método de pesquisa tradicional em que o pesquisador já tem um problema e este é confrontado com um referencial teórico, servindo como base para confrontar conceitos, elaborar hipóteses etc.

Justamente por valorizar o contato do pesquisador com o objeto e estimular a criação de uma sensibilidade para os dados, a Teoria Fundamentada traz uma contribuição diferenciada. Ela permite ao pesquisador que foca um fenômeno bastante novo que tenha a chance de experimentar o campo empírico, observando os novos elementos e construindo suas percepções através da análise e reflexão sistemáticas dos dados encontrados em campo. Essa valorização da experiência do pesquisador em campo é fundamental e é uma das principais vantagens do pesquisador que utiliza essa abordagem. (FRAGOSO; RECUERO; AMARAL, 2011, p. 87).

Na *Grounded Theory*, teorização e observação empírica caminham juntas, acontecem simultaneamente, portanto, o pesquisador deve ir a campo livre de pré-conceitos, noções e hipóteses, e, somente a partir da experiência empírica com base nos dados coletados e da execução do método é que se inicia a elaboração de hipóteses e conceitos teóricos, ou seja, sua principal característica reside no campo de pesquisa e na observação, as quais irão fornecer as hipóteses, a delimitação do problema e a construção da teorização. Nesse sentido, os dados serão fundamentais nesse processo. No entanto, é preciso esclarecer que isso nada tem a ver com o fato de o pesquisador adentrar no espaço de pesquisa sem nenhum tipo de conhecimento sobre o tema de sua investigação, mas sim de que ele deve permitir realizar esse processo por meio da análise dos dados obtidos e pela sua experiência vivenciada no campo. (CHARMAZ, 2006, 2008a, 2008b; FRAGOSO, RECUERO, AMARAL, 2011; STRAUSS; CORBIN, 2008).

Corbin e Strauss (1990) argumentam que na Teoria Fundamentada os procedimentos são traçados para desenvolver um conjunto de conceitos provenientes de uma explicação teórica aprofundada sobre um fenômeno social estudado. Destarte, a teoria deve explicar assim como descrever. A Teoria Fundamentada não possui uma única perspectiva. De acordo com o quadro 7, existem três versões principais:

Quadro 7 - Características das principais correntes da Teoria Fundamentada em Dados

Abordagens	Clássica	Descrição conceitual	Construtivista
	Glaser e Strauss	Strauss e Corbin	Charmaz
<b>Pergunta de Pesquisa</b>	Não é uma afirmação que identifica o problema a ser investigado. É impossível defini-lo antes de ir a campo.	É uma afirmação que identifica claramente o problema de pesquisa. Permite restringir e gerenciar a área de estudo.	Não existe. O que inicia a pesquisa são os conceitos sensibilizantes, os interesses pessoais e disciplinares.
<b>Tipos de dados</b>	Tudo é dado.	Indiferente. Sobretudo observações.	Entrevistas semiestruturadas e análise textual. Construção do conjunto de dados.
<b>Categoria Central</b>	Emerge de forma mágica, se sobressai das demais categorias, é intuída de improviso no início ou no fim da pesquisa.	Para que emerjam os dados é necessária a manipulação dos dados, e, não existe categoria central única.	Existe uma categoria central predominante.
<b>Tipos de codificação</b>	Substantiva teórica.	Aberta, axial e seletiva.	Inicial, focalizada, axial, teórica.

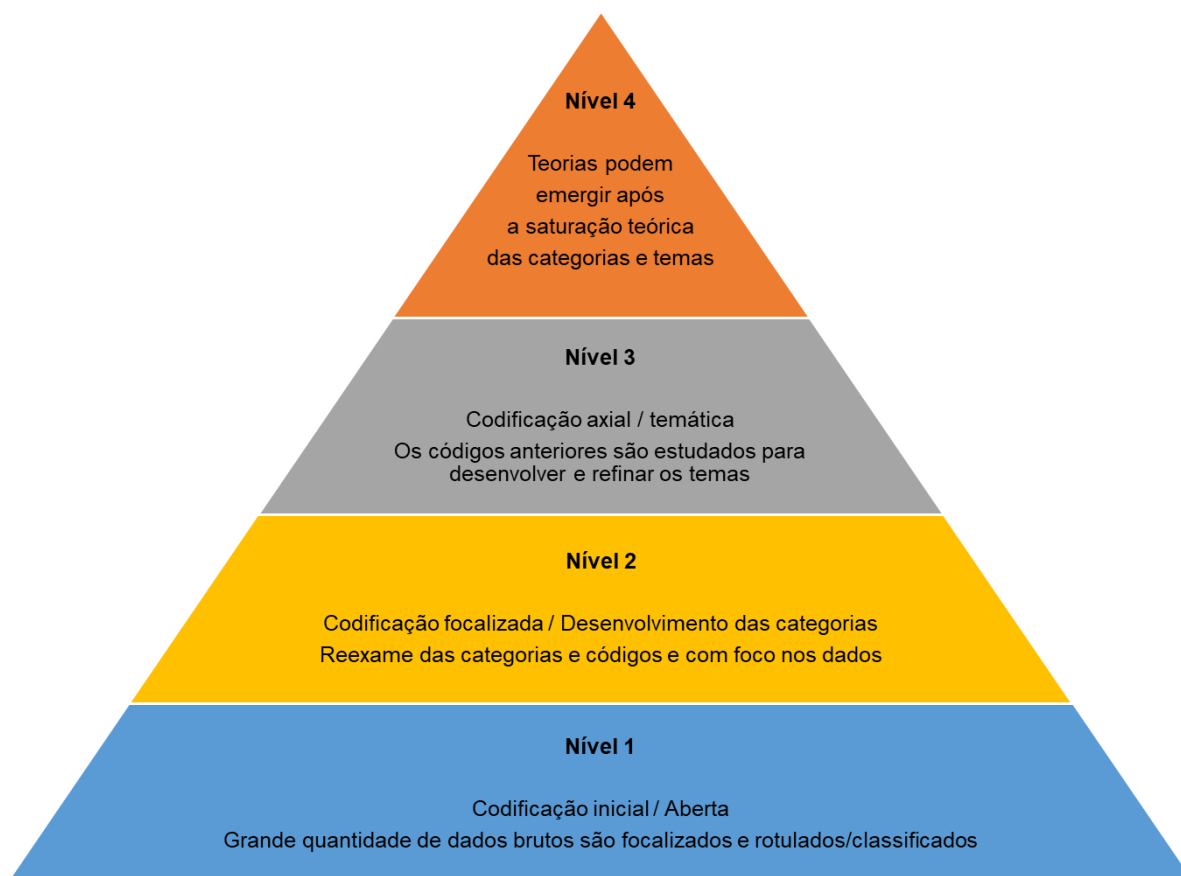
Fonte: Adaptada de Tarozzi (2011, p. 56) e Bittencourt (2017, p. 153).

Em suma, o quadro acima sintetiza cada uma das escolas relacionadas ao método: Na Teoria Fundamentada Clássica, elaborada por Glaser e Strauss, posteriormente, os dois autores se separaram por conta de discordâncias de abordagem do método e divergências ideológicas. Na Teoria Fundamentada Straussiana, segunda vertente elaborada por Strauss após juntar-se a Juliet Corbin e publicarem sua versão da Teoria Fundamentada intitulada “*Basics of Qualitative Research: Techniques and Procedures for developing Grounded Theory*”. A última abordagem é a Teoria Fundamentada Construtivista elaborada por Kathy Charmaz, a qual introduziu uma evolução ao método, cujos elementos constitutivos passam a ser a atenção ao contexto; o posicionamento dos atores envolvidos, as situações e as ações desenvolvidas; a apresentação de múltiplas realidades e a subjetividade do pesquisador como um recurso mobilizador.

Sobre os procedimentos executados na Teoria Fundamentada, Glaser e Strauss (1967), Strauss e Corbin (1990), Cassiani, Caliri e Pelá (1996) e Charmaz (2006, 2008a, 2008b, 2009) descrevem as seguintes etapas a serem seguidas neste

método: a) Coleta dos dados empíricos ou os procedimentos de codificação e análise dos dados; b) Codificação aberta, codificação axial para a formação e desenvolvimento de conceitos; c) Codificação seletiva ou modificação e integração de conceitos e delimitação da teoria, conforme ilustra a figura 22 a seguir.

Figura 22 - Processo de codificação na Teoria Fundamentada



Fonte: Adaptado de Glaser e Strauss (1967), Strauss e Corbin (1990) e Charmaz (2009).

A **codificação aberta** é o processo que consiste na comparação dos incidentes aplicáveis a cada categoria. O pesquisador atribui códigos e classifica os dados em categorias e subcategorias. Essas categorias podem ser construídas pelo próprio pesquisador ou identificadas e extraídas da própria pesquisa. Durante todo esse processo de codificação, os memorandos ou 'memos' são essenciais porque registram formulações de teorias e podem se apresentar em forma de anotações, notas teóricas, notas metodológicas, notas codificadas, esquemas, desenhos, mapas mentais e uma infinidade de notas em outros formatos. A elaboração dos memorandos é fundamental porque também serão analisados e incorporados à análise textual ou ao texto de relatórios. (GLASER; STRAUSS, 1967).

No estágio da **codificação axial** se executa o procedimento da redução, onde as categorias formadas são agrupadas, analisadas e comparadas à luz dos novos dados coletados em campo e, em seguida, realiza-se a identificação das que são mais significativas a fim de reduzir o número de categorias, afinal, esta é a etapa em que as categorias se tornam melhor definidas e organizadas. Essa etapa auxilia na integração das categorias, reunindo os dados e estabelecendo conexões entre categorias principais e subcategorias. (CASSIANI; CALIRI; PELÁ, 1996).

A **codificação seletiva** é a fase em que emerge a variável central e a integração das categorias. “A categoria central emerge no final da análise e forma o pivô ou o principal tema ao redor do qual todas as categorias giram.” (CASSIANI; CALIRI; PELÁ, 1996, p. 82). A codificação seletiva demanda que o pesquisador desenvolva uma estrutura teórica denominada como ‘paradigma de análise’. Em suma, assim se constitui:

Neste paradigma as **condições causais** são definidas como o conjunto de eventos, incidentes e acontecimentos que levam à ocorrência ou desenvolvimento do fenômeno. O **fenômeno**, por sua vez, é a idéia central, o evento, acontecimento e incidente sobre o qual um grupo de ações ou interações são dirigidas ou estão relacionadas. O **contexto** é tratado como um grupo específico de propriedades que pertencem ao fenômeno, representando um grupo particular de condições dentro do qual as estratégias de ação/interação são tomadas. As **condições intervenientes** são aquelas condições estruturais que se apoiam nas **estratégias de ação/interação** e que pertencem ao fenômeno. Elas facilitam ou bloqueiam as estratégias tomadas dentro de um contexto específico. As estratégias para lidar, para serem tomadas ou responder ao fenômeno são denominadas de estratégias de ação/interação. E finalmente as **conseqüências** [sic] são identificadas como os resultados ou expectativas da ação/interação. (CASSIANI; CALIRI; PELÁ, 1996, p. 82, grifo do autor).

Depois de todos os dados estarem codificados, a próxima etapa é a redução das categorias e a delimitação da teoria. De acordo com Strauss e Corbin (1990), a saturação teórica das categorias ocorre quando nenhum dado novo ou relevante emerge. Nesse sentido, deve-se relacionar as subcategorias em torno da categoria central por meio do modelo de paradigma, validar as relações e complementar com dados adicionais aquelas categorias que necessitem de refinamento e desenvolvimento.

Charmaz (2009), autora da mais recente abordagem da *Grounded Theory*, recomenda três tipos de codificação: a) Codificação inicial, a qual implica em transcrever todo o material coletado, fragmentar todos os dados em palavras, linhas ou segmentos, e depois analisar as frases selecionadas, extraíndo palavras-chave. A

construção dos memorandos é realizada em todas as etapas; b) Codificação focalizada, etapa em que faz uso de códigos e/ou categorias iniciais que aparecem com maior frequência ou que são mais significativas para classificar, sintetizar, integrar e organizar os dados; c) Codificação axial, fase em que são especificadas as propriedades e as dimensões de uma categoria e suas relações com as subcategorias a fim de reagrupar dados que foram fragmentados durante o processo de codificação inicial para dar coerência a análise e para a emergência dos dados.

Neste tipo de metodologia a base está na coleta e análise de dados, por isso, é comum o uso de abordagem primordialmente qualitativa, mas também quantitativa associada à qualitativa, além da combinação de métodos complementares e diversos instrumentos de coleta de dados para buscar compreender a realidade pesquisada por meio de entrevistas, questionários, grupos focais, estudos etnográficos e outros. Exatamente por isso a primeira coleta realizada no campo é tão importante, pois, esta será essencial para apontar quaisquer ajustes e os caminhos a serem seguidos posteriormente, e depois de todo o processo na confrontação dos dados com as bases teóricas.

O segundo método adotado, a Netnografia, irá complementar à *Grounded Theory*, considerando que hodiernamente, assim como a pesquisa pode ser empreendida no campo marcada pela presença física do pesquisador, também os estudos qualitativos não escapam aos efeitos da tecnologia e computadores, usados para coletar dados e estabelecer relações de pesquisa e interação entre os indivíduos.

A netnografia também foi adotada nesta pesquisa, e, este termo recebeu influência da etnografia, método estabelecido em 1922 por Bronislaw Malinowski que estuda e descreve a cultura de determinada comunidade a partir dos dados coletados por meio de observação participante durante o processo de pesquisa (CORRÊA; ROZADOS, 2017). A netnografia, por sua vez, é um método de pesquisa embasado no trabalho de campo online e na observação participante, surgiu nos anos 90 em pesquisas sobre marketing e consumo (KOZINETS, 2014).

Para Bentes Pinto *et al.* (2007, p. 5), a netnografia é “uma descrição escrita resultante do trabalho de campo que estuda as culturas e as comunidades online emergentes, mediadas por computador, ou comunicações baseadas na Internet [...]”. Nesse sentido, Kozinets (2014) considera esta metodologia oriunda da antropologia como o tipo de pesquisa que segue as regras da etnografia aplicada ao ambiente

virtual, adaptada às contingências específicas dos mundos sociais no ciberespaço e mediados por computadores.

De acordo com Braga (2013), o termo netnografia é a combinação das palavras '*net*' e '*ethnography*' e foi criado em 1995 pelos pesquisadores americanos Bishop, Star, Neumann, Ignacio, Sandusky e Schatz como forma de traduzir os desafios metodológicos relacionados à observação em estudo etnográfico via meio eletrônico. A pesquisa realizada pelos estudiosos visava testar novos equipamentos para a construção de uma biblioteca digital na Universidade de Illinois.

Alguns autores relatam que o termo 'netnografia' é mais utilizado por pesquisadores da área de Marketing e Administração, enquanto a Antropologia e as Ciências Sociais empregam o termo 'etnografia virtual' (AMARAL; NATAL; VIANA, 2008; CORRÊA; ROZADOS, 2017). A netnografia pode combinar diferentes instrumentos e técnicas de pesquisa por ser multimétodos, adaptável e flexível. Ademais, proporciona a imersão do pesquisador a fim de se obter uma compreensão aprofundada de seu objeto de estudo (CORRÊA; ROZADOS, 2017).

Neste estudo a netnografia será usada como o segundo método e empregaremos a netnografia mista, conforme Kozinets (2006), é uma pesquisa que se caracteriza pela combinação da coleta de dados e interações online com as que são coletadas por meio do contato face a face.

A netnografia é um método que possui muitos processos semelhantes aos que são utilizados na Teoria Fundamentada, tais como (KOZINETS, 2006):

- a) A utilização de dados arquivados que são registros copiados diretamente do ambiente online onde ocorrem as interações;
- b) A extração de dados por meio de interação e troca de informações com membros da cultura investigada;
- c) Utilização de notas de campo, de forma que o pesquisador também registre suas próprias observações a respeito da comunidade;
- d) Codificação: Estabelece códigos e categorias para os dados online extraídos;
- e) Utiliza memorandos, anotações e observações à margem dos dados;
- f) Faz uso da abstração e comparação a fim de classificar, filtrar e identificar relações, diferenças, semelhanças etc.;
- g) Verifica e refina os dados coletados, pois retorna ao campo para verificar e refinar a compreensão do que os dados apresentam;



- h) Se vale da generalização para elaborar explicações e dar maior consistência ao conjunto de dados;
- i) Teoriza, a partir das generalizações, forma-se um conjunto de conhecimentos, teorias e análises para construir novas teorias e incorporação ao conhecimento existente.

Metodologicamente, Kozinets (2014) afirma que a netnografia funda-se na observação participante mesmo sem o contato face a face. Ainda assim, a participação e o envolvimento do pesquisador são intensos, pois, os ambientes onde ocorrem as interações e as comunicações mediadas por computador podem acontecer mesmo durante a ausência física do pesquisador (observador).

Nesse sentido, por razões culturais, na pesquisa qualitativa em que se lida com seres humanos, a condição social, a faixa etária ou qualquer outra classificação estão imbricadas com o investigador, isto é, “Numa ciência, onde o observador é da mesma natureza que o objeto, [...] ele mesmo, é uma parte de sua observação.” (LÉVI-STRAUSS, 1975, p. 215).

Essencialmente, como **esta metodologia utiliza a técnica da observação participante** para a coleta e análise dos dados que foram gerados no ambiente virtual, no caso desta investigação, aplicou-se ao ambiente virtual de aprendizagem escolhido para a realização do minicurso online (*Google Classroom*), proposta a ser detalhada em seção posterior. Optou-se pela abordagem à distância para a realização da capacitação pela dificuldade em conciliar a agenda, dias e horários dos sujeitos pesquisados, assim, escolhemos aplicar o curso neste formato a fim de dar maior liberdade, flexibilidade e fluidez aos pesquisadores participantes desta investigação.

Na observação participante, Angrosino (2009) afirma que o observador possui envolvimento com o contexto social investigado. Assim, na pesquisa de campo, essa técnica é uma maneira de pesquisar que insere o investigador no meio da comunidade estudada, abordagem fortemente defendida por antropólogos influentes como Franz Boas e Bronislaw Malinowski. Assim sendo, destacamos:

[...] a observação participante se resume a uma importante técnica de coleta de dados, empreendida em situações especiais e cujo sucesso depende de certos requisitos que a distinguem das técnicas convencionais [...], tais como o questionário e a entrevista. [...] não só como instrumento de captação de dados, mas também como instrumento de modificação do meio pesquisado, ou seja, de mudança social. [...] representa um processo de interação entre teoria e métodos dirigidos pelo pesquisador na sua busca de conhecimentos

não só da “perspectiva humana” como da própria sociedade. (HAGUETTE, 2010, p. 64, grifo da autora).

Nesse aspecto, “[...] o pesquisador precisa fazer uma imersão no mundo dos seus sujeitos; ele não pode ser um observador neutro das atividades deles, mas precisa subjetivamente tornar-se um deles.” Por esse prisma, entrelaçam-se as relações entre pesquisador e pesquisado, uma vez que há a interferência do bibliotecário. “Não só a interferência é salutar, como é, também imprescindível. Os limites entre a interferência e a manipulação, no entanto, são frágeis e passíveis de não serem percebidos.” (ALMEIDA JÚNIOR, 2007, p. 34). Baseado nessa característica, as ações de mediação durante o minicurso online seguiram os ditames da pesquisa netnográfica, com a aplicação da observação participante e o devido registro das anotações e interações entre os indivíduos, e, por fim coletamos suas opiniões e impressões sobre o curso.

Diante do exposto, justifica-se a combinação dos métodos adotados porque a Teoria Fundamentada em Dados envolve um conjunto rigoroso de procedimentos que conduzem às categorias conceituais a partir de dados empíricos. Estas categorias estão relacionadas umas com as outras. Em relação a netnografia, método que adapta as técnicas etnográficas para estudar culturas e comunidades, usa os mesmos preceitos, porém, aplicado ao ambiente virtual mediado por computador. Na Netnografia, Kozinets (2002) afirma que são utilizadas informações disponíveis publicamente na internet a fim de identificar e desenvolver a compreensão das necessidades e influência de grupos online, sendo este um método exclusivamente utilizado para observações e inserir participação. Ambas as metodologias ressaltam o papel de conceitos sensibilizadores, não são definitivos, ganham utilidade e importância por meio das correlações elaboradas, conforme o quadro 8.

Quadro 8 - Características da Teoria Fundamentada e da Netnografia

Teoria Fundamentada em Dados	Netnografia
<b>Origem e definição</b>	
Método geral de análise comparativa e que institui um conjunto de procedimentos com a finalidade de gerar uma teoria fundamentada nos dados obtidos em uma pesquisa.	Tem origens na etnografia, porém, é aplicada aos ambientes virtuais. É um método que consiste em observar indivíduos em condições reais, assim como as interações, comunicações etc., buscando compreender sua cultura, valores e crenças.
<b>Objetivos</b>	
O desenvolvimento de conceitos e teorias a partir dos dados obtidos em pesquisas e fenômenos estudados, tais como as palavras e ações dos indivíduos.	A identificação das interações, comunicações, comportamentos ligados ao contexto cultural de um grupo no ciberespaço.
<b>Coleta de dados</b>	
Combinação de diferentes métodos necessários para a coleta de dados e levantamento de dados.	O pesquisador tem participação prolongada no ambiente de pesquisa a fim de registrar o comportamento das pessoas.
<b>Análises</b>	
Realizam-se de maneira flexível e aberta por meio da comparação constante realizada simultaneamente com a coleta de dados.	Interação entre o pesquisador e os pesquisados. Sua análise dos dados obtidos é subjetiva e de caráter interpretacionista.

Fonte: Elaborado pela autora, baseado em Mendonça (2001).

Baseado no quadro anterior em que se apresentaram as características dos métodos escolhidos, justifica-se o uso da triangulação neste estudo para a validação da pesquisa, assim como o rigor científico, a qualidade, a riqueza e a complexidade do trabalho será assegurada, considerando que serão estabelecidas relações complementares e não dicotômicas, conforme reforçam Jensen e Jankowski (1993, p. 78): “a debilidade de cada método simples se compensará com o contrapeso da força

do outro.” Por isso, a triangulação confirma a validade dos processos, servindo aos objetivos da pesquisa e contribuindo para que os resultados alcançados possam ser verificados em perspectivas e aspectos variados. Assim, demonstra-se que a pesquisa qualitativa recorre a metodologias diversas para investigação, ou seja, a triangulação trata-se da combinação de abordagens metodológicas. A seguir, serão apresentados os instrumentos de coletas de dados.

### 3.2 Instrumentos de coleta de dados

Visando atingir os objetivos específicos desta pesquisa, as seções seguintes versarão sobre os instrumentos de coletas de dados utilizados para o desenvolvimento desta pesquisa. Assim, o objetivo geral se concentra em investigar as práticas e necessidades informacionais concernentes à Ciência Aberta e ao gerenciamento de dados de pesquisa, conseqüentemente, trabalha-se com a proposta da criação de serviços de dados e um programa de suporte e apoio ao pesquisador na Universidade Federal do Ceará (UFC). Para tal, os objetivos específicos estão alinhados com os seguintes instrumentos (Quadro 9):

Quadro 9 - Objetivos específicos, Instrumentos e fontes de coleta

(continua)

OBJETIVOS ESPECÍFICOS	INSTRUMENTOS	FONTE DE COLETA DE DADOS
Analisar a percepção dos pesquisadores sobre os princípios da Ciência Aberta e o gerenciamento de dados de pesquisa;	Aplicação de Questionário e Entrevista	Pesquisadores da Pós-Graduação <i>Stricto Sensu</i> da UFC
Averiguar quais as práticas e as necessidades informacionais destes pesquisadores referentes a estas temáticas;	Aplicação de Questionário e Entrevista	Pesquisadores da Pós-Graduação <i>Stricto Sensu</i> da UFC
Elaborar uma proposta de Política de Gestão de Dados com a finalidade de incentivar a adoção de práticas em gerenciamento de dados de pesquisa;	Análise documental e Fontes formais e informais de pesquisa	Sites das instituições ligadas aos serviços de gestão de dados de pesquisa e Ciência Aberta, materiais diversos disponibilizados, além de artigos e livros recuperados durante o processo de busca no Portal da Capes e <i>Twitter</i> .
Propor um Programa de Gestão de Dados (PGDP) com o objetivo de sugerir a criação de serviços de dados na UFC;	Análise documental e Fontes formais e informais de pesquisa e Diário de Campo Eletrônico	Sites das instituições ligadas aos serviços de gestão de dados de pesquisa e Ciência Aberta, materiais diversos disponibilizados, além de artigos e livros recuperados durante o processo de busca no Portal da Capes e <i>Twitter</i> .

(conclusão)

OBJETIVOS ESPECÍFICOS	INSTRUMENTOS	FONTE DE COLETA DE DADOS
Desenvolver ações de informação com vistas a testar um piloto de curso online como parte integrante do programa.	Aplicação do minicurso online e registro no Diário de Campo Eletrônico e em Caderno Eletrônico de Laboratório	Minicurso online e participantes

Fonte: Elaborado pela autora (2019).

Esta pesquisa usou elementos que auxiliam a assegurar a qualidade do estudo e coleta de dados, tais como a adoção da triangulação de dados. A seguir, serão apresentados os instrumentos de coleta de dados aplicados na pesquisa.

### 3.2.1 Questionário

Na etapa inicial deste estudo foi utilizado o questionário com o objetivo de fazer um levantamento e diagnóstico institucional sobre a Ciência Aberta, o compartilhamento e a gestão de dados de pesquisa. Para a elaboração do questionário (Apêndice A) a ser aplicado com os professores e alunos de pós-graduação *stricto sensu* da UFC (mestrado e doutorado), nos baseamos em modelos validados, sendo um questionário e um survey internacional, e um questionário nacional, a saber:

- A. Questionário intitulado “*What do you think about Open Science? 10 questions about your views on Open Science*”<sup>130</sup>, elaborado pelo *Centre de Recerca per a l'Educació Científica i Matemàtica* (CRECIM) da *Universitat Autònoma de Barcelona* (UAB) em parceria com a iniciativa ORION e com financiamento da União Europeia;
- B. Survey denominado “*Evaluation of Research Careers fully acknowledging Open Science Practices: Rewards, incentives and/or recognition for researchers practicing Open Science*”<sup>131</sup>, criado pela União Europeia;

<sup>130</sup> Disponível em: [https://www.orion-openscience.eu/public/2019-01/ORION\\_Questionnaire\\_RPFO-CRECIIM.pdf](https://www.orion-openscience.eu/public/2019-01/ORION_Questionnaire_RPFO-CRECIIM.pdf).

<sup>131</sup> Disponível em: [http://ec.europa.eu/research/openscience/pdf/os\\_rewards\\_wgreport.pdf](http://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport.pdf).

C. Questionário nomeado como “Práticas e percepções sobre acesso aberto a dados de pesquisa”<sup>132</sup>, elaborado pela Rede de Dados de Pesquisa (RDP Brasil) e pela iniciativa Acesso Aberto a Dados de Pesquisa (AADP).

Mesmo tendo sido baseado em modelos já validados na literatura científica, foram feitos dois pré-testes do questionário para validação. O primeiro pré-teste foi respondido por oito pessoas: Mestrado em Ciência da Informação (quatro alunos), Doutorado em Engenharia Elétrica (um professor), Mestrado em Psicologia (dois alunos) e Mestrado em Letras (um aluno).

No questionário foi disponibilizado um espaço para comentários, mas, apenas dois participantes deixaram suas contribuições e sugestões. Após a entrega das respostas, a pesquisadora indagou aos respondentes se havia algo que eles gostariam de sugerir, se tinham compreendido as perguntas ou se acrescentariam alguma informação a esse respeito. Em seguida, foram realizados alguns ajustes no questionário, como mudança de layout em algumas perguntas, reformulação e reescrita de outras questões consideradas redundantes. Também foi retirado o glossário com alguns conceitos antes das seções de algumas perguntas, pois, poderiam influenciar nas respostas dos participantes, além de deixar o questionário com a aparência de ser maior do que realmente é.

Assim, foi feito o segundo pré-teste a fim de validar o questionário. Quatro pessoas responderam, sendo dois bibliotecários (mestrados em Ciência da Informação), um professor em Ciência da Informação, e, por fim, um mestrando em Psicologia. Dois respondentes informaram que esta última versão estava mais fluida e fácil de responder do que a anterior. Outros comentários foram feitos a respeito de sugestões de substituição de termos e pequenos ajustes. Incluiu-se nessa segunda versão um texto de apresentação da pesquisa e os termos e condições de aceitação.

Para o envio e divulgação do questionário solicitamos a ajuda da Pró-Reitoria de Pesquisa e Pós-Graduação (PRPPG), entretanto, não obtivemos retorno sobre a solicitação. Assim, solicitou-se via e-mail às coordenações de cursos com endereços de contato expostos no site da UFC que repassassem o questionário aos alunos de Pós-Graduação. Alguns professores que mantêm contato constante com a Biblioteca de Ciências Humanas/UFC também responderam ao questionário e divulgaram entre seus alunos e grupos de *WhatsApp* do mestrado e doutorado. Um dos sujeitos de

---

<sup>132</sup> Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/185195/001082283.pdf?sequence=1&isAllowed=y>.

pesquisa entrevistado neste estudo divulgou o questionário no grupo de *WhatsApp* da Associação Nacional de Pós-Graduandos (ANPG) composto por participantes do Estado do Ceará. Ainda como procedimento para obter sujeitos para esta investigação, foram enviados 1.121 e-mails para pesquisadores dos cursos de programas de pós-graduação da UFC com o convite para responder ao questionário.

Atualmente, a UFC tem 2.297 docentes ativos e 6.992 discentes vinculados aos programas de pós-graduação *stricto sensu* (UNIVERSIDADE FEDERAL DO CEARÁ, 2019)<sup>133</sup>. Responderam ao questionário 135 pessoas, entretanto, três questionários foram descartados por terem sido respondidos por estudantes de graduação e um indivíduo externo, perfazendo o total de 132 questionários respondidos.

O questionário intitulado “Gestão de Dados de Pesquisa e Ciência Aberta: Percepção dos Pesquisadores da Universidade Federal do Ceará” foi elaborado e executado por meio da ferramenta Formulários Google. O instrumento divide-se em quatro seções:

- Seção 1: Apresentação inicial da pesquisa, objetivos, informações gerais e termos e condições de participação. Não há perguntas formuladas nesta seção, apenas as opções de aceitação dos termos da pesquisa;
- Seção 2 (Perfil dos usuários): Composta de seis questões que visam informar o perfil dos respondentes;
- Seção 3 (Delineamento das pesquisas e tipos de dados produzidos): Parte formada por nove perguntas das quais uma é aberta;
- Seção 4 (Práticas e princípios da Ciência Aberta): São dezesseis perguntas, incluindo um espaço adicional para comentários, perfazendo o total de 31 perguntas. O tempo médio de resposta para este questionário é de 15 a 20 minutos.

Em suma, no questionário, as questões foram arranjadas de forma a se encaixar nos seguintes aspectos: Perfil dos respondentes, Delineamento das Pesquisas, Gerenciamento dos Dados, Ciência Aberta e Responsabilidades

---

<sup>133</sup> Dados do Anuário Estatístico da UFC, Ano base 2019, disponível em: <https://is.gd/anuarioufc2019>.

institucionais e da biblioteca relacionadas aos serviços de dados. Os dados sobre o perfil permitiram conhecer melhor algumas características do público atendido, enquanto a seção de delineamento das pesquisas ajudou a identificar o tipo de pesquisa e abordagem desenvolvida pelos pesquisadores da UFC. As questões relacionadas ao gerenciamento dos dados de pesquisa serviram para identificar os tipos de documentos produzidos pelos pesquisadores, como eles organizam, armazenam, disponibilizam e preservam os dados científicos oriundos de suas investigações, qual a estratégia de backup, armazenamento e casos de perdas dos dados, qual a forma de documentação das pesquisas e se utilizam o Plano de Gestão de Dados, e ainda uma pergunta sobre a percepção a respeito do conceito de ‘dados de pesquisa’. Sobre essa seção, perguntou-se ainda sobre a ciência da existência de um setor da universidade responsável por disponibilizar um ambiente remoto de processamento de grande volume de dados utilizando máquinas de alto desempenho (Centro Nacional de Processamento de Alto Desempenho – Cenapad/UFC)<sup>134</sup>. A pesquisadora também visitou o Cenapad a fim de conhecer os serviços e ambientes oferecidos para os pesquisadores da UFC. O referido setor está vinculado à Pró-Reitoria de Pesquisa e Pós-Graduação (PRPPG-UFC) e foi fundado em 1994 como um projeto entre a Secretaria da Ciência, Tecnologia e Educação Superior (Secitece), Financiadora de Estudos e Projetos (Finep) e o Ministério de Ciência e Tecnologia. O Cenapad integra o consórcio do Sistema Nacional de Processamento de Alto Desempenho (Sinapad) e reúne cerca de dez centros em todo o Brasil com o objetivo de fornecer formação de uma grade computacional e ofertar recursos de processamento de alto desempenho (PAD). (CENTRO NACIONAL DE PROCESSAMENTO DE ALTO DESEMPENHO, 2019). O Sinapad também é vinculado à Rede Nacional de Ensino e Pesquisa (RNP).

Sobre a seção de Ciência Aberta, apesar de este trabalho dar maior ênfase aos dados de pesquisa, quer estes sejam ou não abertos, o objetivo era identificar se existem pesquisadores na UFC que já adotaram as práticas e valores da *Open Science*, inclusive o compartilhamento dos dados de pesquisa. Outros aspectos são evidenciados em relação aos dados de pesquisa e a Ciência Aberta: há tendências de avaliação de *datasets* para as revistas científicas, o envio e compartilhamento dos dados em repositórios e outros canais formais e informais da comunicação científica,

---

<sup>134</sup> Site do CENAPAD-UFC: <https://cenapad.ufc.br/>.



quer seja num livro, artigo ou até mesmo disponibilizando os dados em redes sociais acadêmicas como o *ResearchGate*. Também se sondou a respeito da percepção dos respondentes sobre a Ciência Aberta. Os dados gerados por meio dos Formulários Google permitiram a criação de vários gráficos e tabelas com informações numéricas a fim de subsidiar o apoio às análises dos dados qualitativos. As perguntas abertas sobre a percepção dos dados de pesquisa e a Ciência Aberta, assim como o espaço destinado aos comentários, foram processados com o auxílio do software Atlas.ti.

### **3.2.2 Entrevista Semiestruturada**

A fim de construir uma estratégia metodológica na coleta de dados que possa expressar o objetivo deste estudo, não se esquecendo de respeitar seus limites e a heterogeneidade de dados e sujeitos, empregou-se o questionário na primeira etapa como uma forma de fazer um levantamento e verificar as práticas e processos em gestão de dados de pesquisa e Ciência Aberta na UFC. Em seguida, concomitantemente a uma primeira análise dos resultados desse questionário, elaborou-se um roteiro de entrevista individual semiestruturada. Esse tipo de instrumento é flexível e permite que o entrevistador tenha a liberdade de fazer outras perguntas a fim de obter mais informações e aprofundamento (HERNÁNDEZ SAMPIERI; FERNÁNDEZ COLLADO; BAPTISTA LUCIO, 2013).

A entrevista visou obter informações em maior grau de granularidade, algumas nuances, reações e a percepção dos sujeitos de pesquisa (professores e alunos de pós-graduação) que não poderiam ser captadas pelo questionário, e como explicitado por Hernández Sampieri, Fernández Collado e Baptista Lucio (2013), a entrevista qualitativa é mais íntima, flexível e aberta, isto é, permite manter uma comunicação e ao mesmo tempo a construção de significados sobre um tema.

Pré-testes foram realizados com três mestrandos (um do curso de Psicologia e dois da área de Ciência da Informação) e um professor (curso de Educação Física). Esse procedimento foi fundamental para alinhar as perguntas, fazer ajustes, reescrever e reformular questões e até para identificar espaço para novas perguntas. Inclusive, durante os pré-testes já se tornou possível sentir as primeiras impressões e diferentes tipos de reação sobre o tema, preparando assim a pesquisadora para o que poderia encontrar no campo de pesquisa.

Após o pré-teste, passou-se a etapa de execução das entrevistas agendadas com os sujeitos de pesquisa. O roteiro de entrevista (Apêndice B) contém dezoito questões abertas que versam desde a abordagem da pesquisa desenvolvida pelos indivíduos; formas de armazenamento e backup a fim de identificar práticas de gerenciamento de dados de pesquisa; buscou-se descobrir características de compartilhamento dos pesquisadores e de suas respectivas áreas; sondou-se a respeito do conhecimento sobre as formas de documentação de uma pesquisa (como a elaboração de um Plano de Gestão de Dados); além de levantar opiniões e percepções sobre a gestão de dados e Ciência Aberta, seus princípios e práticas. A última pergunta é um espaço para comentários adicionais, assim os entrevistados poderiam se expressar livremente caso desejassem.

Na entrevista, o roteiro semiestruturado permitiu que outras perguntas pudessem ser feitas caso fosse necessário, e o encadeamento e a ordem foram semelhantes às do questionário. Todavia, algumas questões foram elaboradas para a confirmação ou o aprofundamento dos primeiros achados de pesquisa, e, a partir disso, com a finalidade de extrair as categorias de análise deste estudo.

Uma das perguntas visava sondar o interesse dos entrevistados em participar do minicurso online, sendo essa mais uma estratégia adotada como forma de recrutamento das pessoas que quisessem participar dessa última etapa da pesquisa. Entretanto, durante as primeiras entrevistas realizadas, dois entrevistados da área de Ciências Sociais alertaram que nem todas as pessoas se sentiriam à vontade para recusar a participação, ou seja, correria o risco de se sentir pressionado a participar, enquanto o outro pesquisador alertou que nem todos têm o mesmo ritmo, alguns já estão em uma fase da vida que não estão mais dispostos a fazer cursos, a não ser que queiram ou despertem o interesse, então, foi sugerido que a pergunta fosse reformulada e suavizada, de maneira a deixar qualquer respondente mais confortável para aceitar ou recusar a participação no curso, e, caso a resposta fosse positiva, a pesquisadora poderia então falar da última etapa do estudo e fazer o convite, sugestões substanciais que foram acatadas.

De acordo com o que foi explicitado na seção sobre o universo e amostra de pesquisa, buscou-se por um representante de cada uma das principais grandes áreas do conhecimento que constam na Tabela de Área do Conhecimento da Capes, a saber: Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, Ciências Agrárias, Ciências Sociais Aplicadas, Ciências Humanas, Linguística,

Letras e Artes, e Multidisciplinar. Esse foi o critério aplicado a fim de contemplar a representatividade de todas as áreas do conhecimento. Realizaram-se doze entrevistas contemplando as áreas supramencionadas.

### 3.2.3 Diário de Campo Eletrônico e Caderno de Laboratório Eletrônico

No início da pesquisa de campo, o aplicativo gratuito *OneNote* foi usado como diário de campo eletrônico. A opção por este aplicativo se deu pelo fato de a pesquisadora tê-lo instalado em seu computador e celular, além de ter uma interface intuitiva, possibilita criar vários cadernos divididos em seções e notas dentro do mesmo sistema, permite a criação de senha para proteger o acesso aos cadernos de anotações, além de ser possível a sincronização de dados entre vários dispositivos (computador, celulares, *tablets* e outros dispositivos móveis), pode ser utilizado mesmo *offline* e posteriormente sincronizar as anotações, pode ser instalado no computador, está disponível na nuvem e possui aplicativo para celular, tem pesquisa instantânea por anotações, além de possuir recursos extras, o que não seria possível com aplicativos concorrentes semelhantes, os quais exigem a assinatura de um plano anual para ter acesso a mais recursos e espaço de armazenamento (Quadro 10).

Quadro 10 - Características e funcionalidades do OneNote

(continua)

Características e Funcionalidades	OneNote	Concorrente
Disponível para Windows, Mac, iOS, Android e na Web.	✓	✓
Sincronização de notas em todos os seus dispositivos.	✓	Limitado a 2 dispositivos para a versão gratuita. Requer um plano pago para sincronizar múltiplos dispositivos.
Acesso offline a notas em dispositivos móveis.	✓	Necessita de uma assinatura paga.
Carregamentos mensais ilimitados	✓	60 MB/mês (Gratuito) 1 GB/mês (Plano pago)

Quadro 10 - Características e funcionalidades do OneNote

(conclusão)

<b>Características e Funcionalidades</b>	<b>OneNote</b>	<b>Concorrente</b>
Escrever em qualquer local na página com a tela de forma livre	✓	✗
Compartilhamento de conteúdo com outras pessoas	✓	✓
Ferramenta de recorte / captura de conteúdo da Web	✓	✓
Guarda mensagens de e-mail nas suas notas	✓	Necessita de plano pago.
Digitaliza cartões de visita	✓	Necessita de plano pago.
Permite a inclusão de imagens, vídeos, áudios, gravações em áudio e vídeo e afins	✓	✓

Fonte: Elaborado pela autora, baseado no site oficial do OneNote (2019).

No entanto, pouco tempo após o uso do *OneNote*, a pesquisadora buscou por recursos mais robustos, em consonância com a Ciência Aberta e o gerenciamento de dados de forma responsiva por meio de aplicativos ou softwares direcionados à academia e capazes de armazenar certas informações de pesquisa, como metadados melhor estruturados e com a capacidade de oferecer interoperabilidade, atribuição de etiquetas (*tags*), integração com diversas plataformas e repositórios de dados. Após uma vasta busca por diários de campos eletrônicos, percebeu-se que há uma infinidade de opções oferecidas para essa função, mas, na verdade, constatou-se que existe a carência de ferramentas especializadas para esta finalidade, portanto, sendo terreno fértil para o uso adaptado de recursos variados como diário de campo eletrônico. Talvez, seja por que muitos softwares de análise qualitativa já têm

incorporados alguns recursos de anotações semelhantes ao diário e de construção de memorandos e outros elementos.

Por outro lado, é grande a oferta por *Electronic Laboratory Notebooks* – ELN (Cadernos de Laboratório Eletrônicos) – alguns gratuitos e outros pagos, que são softwares e ferramentas direcionadas, especializadas e inclusive contam com mais recursos integrados. Assim, depois de comparar os prós e contras das ferramentas usadas como diário de campo eletrônico e os cadernos de laboratório, escolheu-se combinar o uso do *OneNote* com o *ResearchSpace*<sup>135</sup> (*RSpace*) para registrar alguns dos dados de pesquisa selecionados e anonimizados, mantendo-os integrados e seguros em uma mesma plataforma. O *RSpace* é uma ferramenta de pesquisa multifacetada que incorpora em sua estrutura os seguintes elementos:

- a) Caderno de Laboratório Eletrônico: permite capturar e organizar dados;
- b) Software de colaboração: ideal para o compartilhamento e organização de dados em grupos de trabalho;
- c) Ferramenta de gerenciamento de fluxos de trabalho;
- d) Plataforma que permite capturar, publicar e arquivar dados.

Criado em 2012, o *RSpace* foi desenvolvido por professores, profissionais de TI, bibliotecários de dados e administradores da Universidade de Wisconsin em parceria com a Universidade de Edimburgo e Universidade de Goettingen. O *RSpace* tem API<sup>136</sup> (*Application Programming Interface*) e armazenamento ilimitado, possui ainda integração com diversas ferramentas, tais como o *Box*, *Dataverse*, *DSpace*, *Dropbox*, *Egnyte*, *Figshare*, *GitHub*, *Google Drive*, *Google Hangouts*, *Mendeley*, *Microsoft Teams*, *ORCID*, *OneDrive*, *OwnCloud*, *Protocols.io*, *Slack*, *SURFdrive* e *SwitchDrive*.

O *RSpace* (2019) segue as boas práticas de laboratório e estratégias de proteção de patentes em conformidade com o padrão 21CFR11, considerado o padrão ouro internacional importante para o controle de versões e o gerenciamento dos dados científicos. Esse padrão estabelece regras para registros eletrônicos,

---

<sup>135</sup> Disponível em: <https://www.researchspace.com/>.

<sup>136</sup> API RSpace: <https://community.researchspace.com/public/apiDocs>.

segurança da informação, controle de acesso aos dados, assinatura digital, auditoria, entre outros elementos<sup>137</sup>.

Apesar de existirem certas diferenças entre os registros de um diário de campo eletrônico e um caderno de laboratório eletrônico, adotou-se também o *RSpace* (Caderno de Laboratório Eletrônico) para registrar alguns dados selecionados referentes a este estudo, isto é, a estratégia combinada do diário de campo eletrônico e o caderno de laboratório eletrônico. Então, alguns dos conteúdos produzidos inicialmente que constavam no *OneNote* foram refinados, selecionados e incorporados ao *RSpace* ao longo do período da pesquisa.

Para Falkembach (1987) o diário de campo é um instrumento de anotações, comentários e reflexões de **uso individual do pesquisador** em seu cotidiano que facilita a escrita e a observar com atenção, auxiliando na precisão da descrição e reflexão sobre os acontecimentos. Podem ser anotadas observações de fatos concretos, fenômenos sociais, acontecimentos, experiências pessoais, comentários e reflexões do pesquisador, além das verificações feitas no trabalho de campo. A justificativa para isso se deve pelo fato de que um diário de campo, por tradição, não costuma ser aberto, compartilhado ou objeto de verificação por parte dos pesquisadores e áreas que costumam usar esse instrumento, tendo em vista o seu caráter subjetivo e intimista, que abriga anotações de nomes, diálogos, descrição de pessoas, narrativas e histórias de vida, notas complementares de observações sobre gravações de áudio e vídeo para registrar as impressões e vivências do trabalho de campo.

Nesse sentido, o diário de campo documenta a pesquisa, porém, pode conter informações que comprometam a identidade dos sujeitos de pesquisa e informantes, e, assim sendo, não costumam ser objeto de auditoria. O caderno de laboratório eletrônico, por sua vez, é mais relacionado à descrição de procedimentos, experimentos e observações que podem ser partilhadas, são auditáveis, e, geralmente não envolvem a revelação de dados que comprometem a identificação de pessoas ou a exposição de dados sensíveis. Em resumo, o caderno eletrônico de laboratório:

[...] é um sistema para criar, armazenar, recuperar e compartilhar registros eletrônicos. Ao invés de registrar informações no papel, os esboços, textos, equações, imagens, gráficos e outros dados são gravados eletronicamente.

---

<sup>137</sup> Ver também: <http://www.fda.gov/ICECI/EnforcementActions/default.htm>.

[...] pode envolver dados de entrada, outra saída de programa, equipamento de imagem, microfone e diretamente de instrumentos científicos. Os cadernos eletrônicos podem variar em capacidade e complexidade, dos tipos mais simples, que usam software comum (processamento de texto, planilha, gráficos) no computador para anotar e acompanhar arquivos de dados, para sistemas mais parecidos com cadernos, até software comercial especial para autenticação. [...] Se estiverem envolvidos dados pessoalmente identificáveis e / ou sensíveis, devem ser cumpridos os padrões apropriados da Lei de Privacidade e segurança de TI. (NATIONAL INSTITUTES OF HEALTH, 2008, p. 8, tradução nossa).

Sobre os cadernos de laboratório, Ryan ([200-?], p. 3, tradução nossa), esclarece que “Um mal-entendido comum sobre o caderno de laboratório é que ele é um diário para as suas considerações científicas ou pessoais.” Na visão do referido autor, existem diferenças entre o registro de uma hipótese a ser testada por um experimento específico e anotar ideias no caderno eletrônico. Reforça ainda que é desaconselhável o uso do caderno de laboratório para manter registros de conversas ou comunicações.

Por fim, o uso do caderno de laboratório eletrônico nesta pesquisa teve a intenção de: a) registrar alguns dados selecionados referentes a este estudo que não envolvem a revelação de dados que comprometem a identificação de pessoas ou a exposição de dados sensíveis; b) descrever procedimentos de pesquisa e notas que podem ser compartilhadas; c) documentar a pesquisa e dotá-la de elementos técnicos necessários para gerar uma boa documentação da pesquisa; d) cumprir com as boas práticas da pesquisa científica exigidas no cenário vigente. Ressalta-se ainda que o uso do *OneNote* (diário de campo eletrônico) e *RSpace* (caderno de laboratório eletrônico) ocorreu em todas as etapas da pesquisa.

### **3.3 Ação de Informação: proposta de minicurso online**

A última etapa dessa pesquisa vislumbrou a possibilidade de promover uma ação de informação (também conhecida como ação educacional) por meio de um minicurso online na modalidade Educação a Distância e (EaD) junto a uma pequena parcela da comunidade estudada. O público-alvo para essa intervenção são os professores e alunos de pós-graduação *stricto sensu* da UFC (dentre estes, também os técnico-administrativos em educação que estejam cursando ou tenham concluído o mestrado ou doutorado).

Conforme será explicado na seção de universo e amostra da pesquisa presente neste manuscrito, alguns dos sujeitos entrevistados foram convidados para

participar do minicurso online. Esclareceu-se que a participação seria livre, voluntária e se fosse de acordo com a vontade do participante. Dos 12 entrevistados, cinco aceitaram o convite e participaram do minicurso. Inclusive, as quatro pessoas que participaram do pré-teste da entrevista também se inscreveram. Duas pessoas disseram que preferiam cursos presenciais, e os demais declinaram o convite por estarem muito ocupados durante o período de realização do curso ou por não terem interesse em participar.

Sobre a Educação a Distância, a título de informação, Junqueira (2018, p. 21) explica que

Convencionou-se chamar de educação a distância (EaD) os cursos e processos de aprendizagem formal e informal caracterizados por uma separação física (distância geográfica) entre o professor e o aluno, que não estão juntos por todo o tempo ou pela maior parte do tempo. Devido a essa separação física, geográfica, a prática de estudo apresenta grande flexibilidade, permitindo ao aluno estudar onde e quando melhor lhe convier, não necessitando seguir horários estipulados pela instituição no que se refere a uma parte significativa das atividades em estudo.

A história da EaD revela que existiram outros modelos predecessores, como os cursos por correspondência fortemente baseado em apostilas e livros em 1900, mais tarde foi criado o Instituto Universal Brasileiro, em 1941, que também oferecia cursos por correspondência. Em 1946, o Senac criou a Universidade do Ar e atendia diversas localidades em 1950. Outra modalidade de EaD surgiu com o rádio, em 1923, com a fundação da Rádio Sociedade do Rio de Janeiro e sua Rádio Escola-Postal. Até mesmo a Igreja Católica enveredou pela EaD em 1961 e idealizou o Movimento de Educação de Base (MEB) e era voltado para a educação rural e popular fundamentada nos ensinamentos de Paulo Freire.

Em 1970, o Ministério da Educação (MEC) deu início ao projeto Minerva e ofertava orientações e ensino supletivo aos jovens e adultos. Também a televisão apostou na EaD, a exemplo do Telecurso 2000. Outra infinidade de modelos de EaD foram surgindo ao longo do tempo até chegar em modelos mais atuais e parecidos com os que temos hoje, tal como a teleaula, videoaula, educação online ou *e-learning*, educação móvel ou *m-learning*. (BELLONI, 2001; JUNQUEIRA, 2018; LITTO; FORMIGA, 2009).

Dando continuidade ao assunto sobre a ação de informação, para a idealização dessa intervenção, seguiu-se a teoria do equilíbrio da sabedoria relatada por Sternberg (1998), a qual afirma que uma ação sábia leva em consideração o

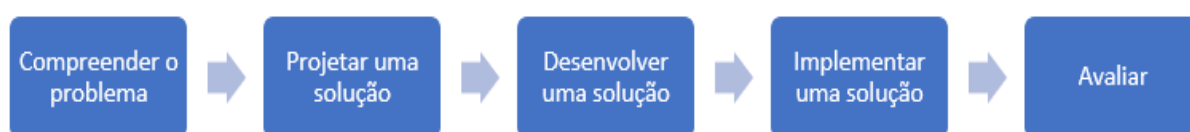


interesse individual, o do próximo e o de sua comunidade. Portanto, “Agir sabiamente também envolve saber lidar com as novidades que surgem no exercício da profissão, com o momento histórico atual, com o advento das novas tecnologias e com a capitalização do conhecimento [...]” (ANTONELLI-PONTI *et al.*, 2018, p. 185).

Desse modo, a etapa de planejamento e esboço do modelo do curso se fundamentou a partir dos primeiros resultados por meio dos dados coletados no questionário de levantamento e nas primeiras entrevistas, isto é, em baseado em evidências. A inspiração para elaborar e delinear o curso advém de Filatro e Piconez (2004), adotando para isso o Design Instrucional (DI), definido pelas autoras como um planejamento do ensino-aprendizagem. Inclui os tipos de atividades, as estratégias empregadas, o sistema de avaliação, os métodos e materiais utilizados para aplicação em curso online ou em Educação a Distância, mas tradicionalmente é uma abordagem incorporada na produção de materiais didáticos analógicos.

Basicamente, o tipo de DI aplicado ao curso é o Design Instrucional Contextualizado, também denominado DI 4.0 (FILATRO *et al.*, 2019). A escolha se deve pelo fato de que é um modelo centrado nas pessoas e por integrar todas as potencialidades dos modelos tradicionais de design instrucional (DI). É composto pelas seguintes etapas (Figura 23):

Figura 23 - Etapas do DI 4.0



Fonte: Adaptado de Filatro *et al.* (2019).

Em suma, o modelo acima se enquadrava bem a proposta de minicurso online, considerando que “[...] a aprendizagem é social, ou seja, se dá na e pela interação com outras pessoas, que aprendem de forma menos estruturada e também em espaços não formais de aprendizagem [...]”. (FILATRO *et al.*, 2019). Ademais, é uma proposta que combina com os métodos adotados, potencializa a capacidade de entender o fenômeno em estudo e seus processos, ao mesmo tempo que dá um retorno aos participantes e contribui para modificar as suas estruturas de conhecimento.

Para a consecução da proposta, a plataforma EaD escolhida foi o *Google Classroom*, pelo fato de ser relativamente conhecida e adotada na universidade, além de ser bastante intuitiva e fácil de utilizar, tanto para os sujeitos de pesquisa quanto para a instrutora/tutora. Além disso, o *Google Classroom* é multiplataforma e acessível em tablets e celulares, o que oferece ao usuário a oportunidade de fazer o curso totalmente pelo celular, ou combinando o uso em diferentes equipamentos (computador ou tablet).

Para os fins deste estudo, o minicurso na modalidade EaD foi restrito ao público-alvo da pesquisa. O curso segue os moldes de acompanhamento tutorial semipresencial por meio da plataforma educacional utilizada, e como formas de interação citamos a síncrona (interação em tempo real via chat, hangouts e videochamadas) e assíncrona (não exige a participação simultânea, por meio de mensagens nas postagens, mensagens enviadas particularmente ou por e-mail). A carga horária estipulada foi de 15h para o conteúdo completo do curso introdutório e o período em que o curso ficou disponível para os participantes compreende o período de 11 de setembro de 2019 a 10 de outubro de 2019 (30 dias).

Cada curso possui uma particularidade e um desenho único, sua própria maneira de ser construído, embora possam ser encontrados elementos em comum em diferentes cursos. Desse modo, o modelo construído e documentado para este minicurso pode ser consultado no Apêndice C<sup>138</sup>.

Os materiais didáticos elaborados para entregar os conteúdos aos participantes são os seguintes:

- a) Apresentações de slides;
- b) Videoaulas inéditas e exclusivamente produzidas para este minicurso, gravadas e capturadas pelo software gratuito *Atube Catcher*, posteriormente editadas no software proprietário *Camtasia* e salvas ao final em arquivo mp4;
- c) Documentos de anotações resumidas (Notas de aula) com o conteúdo de cada aula, com informações e links extras;
- d) *Podcast*;
- e) Indicações de materiais externos, tais como links, sugestões de leitura, documentários e vídeos diversos, entre outros recursos audiovisuais.

---

<sup>138</sup> Informação verbal, sugestão proferida pela Profa. Dra. Andréa Soares, em ocasião da banca de qualificação desta dissertação. Modelo disponível em: <https://is.gd/plcursodadosres20>.

Para os slides e notas de aula, tanto quanto possível, aplicou-se o uso do QR Code, a fim de ampliar a experiência de acesso aos conteúdos, participação no curso e imersão na plataforma virtual de aprendizagem. As videoaulas foram gravadas e editadas para a aplicação dos efeitos visuais, assim como edições no áudio das videoaulas a fim de melhorar a qualidade do som.

A divulgação das inscrições foi realizada por e-mail enviado aos pesquisadores da UFC, conforme solicitação da pesquisadora às coordenações de cursos de graduação e pós-graduação. No texto do e-mail, também foram incorporados links para o formulário de inscrição<sup>139</sup> e outro para um vídeo promocional de um minuto de duração<sup>140</sup> elaborado pela pesquisadora na ferramenta Biteable. Esse mesmo vídeo foi incorporado no formulário de inscrição.

Faz-se necessário explicar que a ideia de ministrar um minicurso online não se deve apenas à proposta de uma intervenção que visa uma ação de informação educacional, mas também, foi uma das maneiras encontradas para conciliar as agendas e participação dos pesquisadores, uma vez que alguns dos entrevistados alegaram falta de tempo para participar de um curso presencial, portanto, o curso EaD se tornou uma opção viável e dinâmica, pois pode ser acessado em qualquer lugar, em diferentes equipamentos (celular, computador e tablet), em qualquer horário e estar disponível os sete dias da semana e 24 horas por dia.

Por se tratar de um teste piloto, espera-se que o produto final desse minicurso online possa ser ajustado, incorporado à proposta de criação de um programa de apoio e suporte ao pesquisador, e também servir para que outros pesquisadores da universidade possam aproveitá-lo ao máximo. Nas palavras de Freire (2017, p. 4), “As ações de informação criam uma sinergia para o trabalho a ser empreendido na rede de projetos, além de gerar comprometimento nos sujeitos sociais envolvidos [...], com a efetiva construção das condições para sua realização na comunidade acadêmica.”

Apesar de já existirem cursos sobre Ciência Aberta, gestão ou compartilhamento de dados de pesquisa, acredita-se que as instituições precisam criar e promover seus próprios cursos, afinal, estes serão mais voltados para o seu público-alvo e o conteúdo será elaborado de maneira personalizada. Fazer um estudo e levantamento das necessidades dos usuários se tornou primordial nesse ponto

---

<sup>139</sup> Formulário de inscrição do minicurso online: <https://forms.gle/yqF3H4U8PvdtgBfWA>.

<sup>140</sup> Vídeo de divulgação do curso: <https://is.gd/gdados1minpromo>.

porque também ajuda a traçar a melhor estratégia de como desenvolver essas capacitações, quer sejam presenciais, online ou mista.

Corroborando o exposto acima, o documento “*Providing researchers with the skills and competencies they need to practice Open Science*” da European Commission e do Working Group on Education and Skills under Open Science (2017) traz recomendações diversas sobre a Ciência Aberta e uma delas é sobre o treinamento de pesquisadores acerca do tema e o incentivo às instituições de ensino para que produzam suas próprias capacitações voltadas para os seus usuários. Tennant (2018) reforça que promover treinamentos em Ciência Aberta ajuda a mudar culturas, empodera e capacita novas lideranças, impulsiona uma dinâmica voltada para reduzir os abusos e erros, além de incentivar a criação de uma comunidade global baseada na colaboração e no compartilhamento. É com esse espírito que seguimos com uma proposta plural e integrativa de investigação do campo e dos sujeitos de pesquisa, mas também da aplicação de uma ação de informação educacional capaz de ser uma prática dinamizadora de mediação da informação com vistas a apoiar o caminho a ser trilhado para a gestão da informação relacionada aos dados de pesquisa na universidade.

### 3.4 Campo de Pesquisa

A pesquisa de campo será desenvolvida na Universidade Federal do Ceará (UFC). A escolha do ambiente se deu pelo fato de que é uma instituição universitária de prestígio na região Nordeste e que está presente em rankings internacionais como o *Best Global Universities*<sup>141</sup>; o *Center for World University Rankings*<sup>142</sup>; o *QS Top Universities (World University Rankings)*<sup>143</sup> e no *CWTS Leiden Ranking*<sup>144</sup>, caracterizando a sua excelência no tripé ensino, pesquisa e extensão. Ademais, a exigência por produção científica é uma constante, uma realidade que faz parte do cotidiano da universidade.

A Universidade Federal do Ceará foi criada em 16 de dezembro de 1954 pela Lei nº 2.373 e instalada no ano seguinte, em 25 de junho de 1955. A UFC é uma autarquia vinculada ao Ministério da Educação. O fundador da UFC, Antônio Martins

---

<sup>141</sup> Disponível em: <https://www.usnews.com/education/best-global-universities/universidade-federal-do-ceara-503817>.

<sup>142</sup> Disponível em: <https://cwur.org/2018-19/brazil.php>.

<sup>143</sup> Disponível em: <https://www.topuniversities.com/university-rankings/world-university-rankings/2016>.

<sup>144</sup> Disponível em: <http://www.leidenranking.com/ranking/2019/list>.

Filho, dirigiu a universidade e sua estrutura inicial era formada pela Escola de Agronomia, a Faculdade de Direito; a Faculdade de Medicina e a Faculdade de Farmácia e Odontologia. Sediada em Fortaleza, atualmente é composta por sete *campi*: três na capital cearense (Benfica, Pici e Porangabuçu) e quatro no interior (Sobral, Quixadá, Crateús e Russas).

A UFC possui um Sistema de Bibliotecas composto por 19 bibliotecas (quatorze em Fortaleza e cinco no interior) e atende aos cursos de graduação, pós-graduação e Casas de Cultura Estrangeira da instituição. Os serviços oferecidos são direcionados para alunos, professores e servidores técnico-administrativos. Oferece diversos serviços tradicionais como empréstimo, devolução, renovação, reserva, entre outros; além de disponibilizar um site em que se encontram todas as informações sobre seus serviços e produtos. Entretanto, não oferece e nem dispõe de quaisquer informações direcionadas ao público-alvo relacionadas à Ciência Aberta, compartilhamento e gerenciamento de dados de pesquisa.

O repositório institucional da UFC está sob a responsabilidade do Sistema de Bibliotecas, no entanto, não há auto arquivamento, contrariando as boas práticas de incentivo ao acesso aberto e motivação para os autores realizarem o depósito de seus materiais na instituição a qual estão vinculados. Há ainda a restrição para a disponibilização de alguns tipos de materiais, embora testes tenham sido realizados por parte da Biblioteca Universitária (BU) com a criação de comunidades de arquivos audiovisuais e de eventos após a contribuição dada no estudo de Santos (2018) com a pesquisa desenvolvida no Mestrado em Ciência da Informação do Programa de Pós-Graduação em Ciência da Informação da UFC (PPGCI-UFC) sobre repositórios e acervos audiovisuais. A referida pesquisa deu o passo inicial para demonstrar a importância e a necessidade de adotar o autoarquivamento e de incorporar outros tipos de materiais que não sejam exclusivamente aqueles revisados por pares no Repositório Institucional da UFC, trazendo à tona também a reflexão a respeito dos tipos de documentos, e, talvez, dados de pesquisa, ou sobre a viabilidade de se criar ou sugerir a adoção de repositórios que incorporem dados de pesquisa. Indubitavelmente, essas são questões a serem discutidas futuramente pela instituição e em alguns pontos deste estudo.

O repositório institucional utilizado pela UFC é o *DSpace* e não dispõe, até o momento, de infraestrutura necessária para abrigar dados de pesquisa ou mesmo planos de implantação para repositórios de dados científicos especializados mais

robustos e com vistas à preservação digital, segurança da informação e auditoria dos dados como já existem em repositórios como o *Dataverse*, *Figshare*, *Zenodo* e outros que contam com o selo internacional *Core Trust Seal*, entre outras certificações internacionais para plataformas de infraestrutura de dados.

### 3.5 Universo e amostra de pesquisa

Inicialmente, faz-se necessário esclarecer que exclusivamente para a finalidade dessa pesquisa qualitativa e por conta da metodologia adotada (*Grounded Theory*), a composição do tipo de amostragem é totalmente distinta da amostragem probabilística estatística, afinal, não se refere ao número de participantes em relação ao universo da população em estudo, isto porque a amostra não se forma *a priori*, mas sim ao longo da pesquisa como parte do processo analítico e vai aumentando a sua extensão progressivamente conforme as exigências do trabalho de conceituação teórica no decorrer da análise, do número de participantes e suas características até a saturação dos dados, isto é, trata-se de uma **amostragem teórica**. (TAROZZI, 2011).

Dessa maneira, com o intuito de iniciar a composição da amostragem de participantes para as entrevistas, uma das estratégias empregadas se deu em contato *vis-à-vis*, presencialmente, durante os atendimentos realizados no ambiente de trabalho da bibliotecária e em suas observações informais empreendidas no Serviço de Referência da Biblioteca de Ciências Humanas. Em seguida, era acertado o agendamento prévio da visita da pesquisadora para comparecer aos locais escolhidos pelos sujeitos de pesquisa. Para outros indivíduos, o convite para a entrevista foi enviado por e-mail, e, posteriormente, acertou-se a marcação de data, horário e local.

Um dos fatores que favoreceu o acesso e o contato com os sujeitos para convidá-los a participar do estudo se deve ao fato de esta pesquisadora trabalhar diretamente com o atendimento ao usuário na referida biblioteca da UFC, além de já ter ministrado treinamentos presenciais e à distância para uma boa parte dos professores e alunos recrutados. Desde o início, os docentes e discentes convidados foram bastante cordiais e receptivos, aceitaram contribuir com a pesquisa seja em entrevista ou participando das etapas ulteriores do estudo. Alguns professores indicaram alunos seus para participar da pesquisa, formando assim a parcela de alunos de pós-graduação selecionados para a investigação.

Durante o recrutamento, alguns contatos se estabeleceram de maneira diferente, conforme o caso de um aluno do Mestrado em Psicologia que se ofereceu espontaneamente para participar da pesquisa por interesse próprio (este contribuiu como pré-teste nos instrumentos de coleta – questionário e entrevista); outro aluno do Mestrado em Geografia foi apresentado e encaminhado por um colega de profissão para contribuir com o estudo, pois o referido discente demonstrou interesse em discutir o assunto e ajudar, prontificando-se a participar da entrevista; e o caso da discente do Mestrado em Linguística (também possui habilitação em Letras Espanhol) que foi encaminhada por uma professora da graduação em Letras Espanhol e também por sua orientadora da Pós-graduação em Linguística a buscar a ajuda da bibliotecária desse estudo para disponibilizar o seu *corpus* de análise textual em um repositório de dados. O referido *corpus* textual produzido pela discente é o produto de sua dissertação ainda a ser defendida.

A discente supracitada só descobriu que poderia obter esse tipo de ajuda e suporte bibliotecário por meio de suas professoras que a encaminharam para a biblioteca, especificamente no setor de atendimento ao usuário, e isto só ocorreu porque uma das professoras também foi apresentada a temática deste estudo pela pesquisadora em ocasião de um treinamento sobre normalização e metodologia da pesquisa, demonstrando assim, evidências reais da necessidade latente por esse tipo de serviço na universidade.

Explicitando mais um dos critérios de seleção dos indivíduos para essa pesquisa, a escolha recaiu sobre os pesquisadores da Universidade Federal do Ceará. Para os fins deste estudo, entende-se como pesquisador cada um dos atores envolvidos na pós-graduação (alunos, professores e técnico administrativos da UFC com pós-graduação *stricto sensu*), afinal, a intenção é de que essa pesquisa seja a mais inclusiva possível, levando em consideração uma das bandeiras da Ciência Aberta, abrangendo assim públicos e atores heterogêneos, com interesses diferentes, afins e/ou convergentes. Deste modo, fazer uma pesquisa só com os pesquisadores de produtividade da UFC não seria inclusivo de fato, ademais, poderiam escapar alguns achados relevantes no campo de pesquisa que envolvem não apenas o professor pesquisador de produtividade, mas também certos anseios e práticas dos alunos de pós-graduação que igualmente desenvolvem pesquisa na UFC. Ainda assim, ressalta-se que esse estudo tem limitações nesse sentido, tendo em vista que não foram englobados aqui os estudantes de graduação, meramente por um critério

de exclusão para afunilar o recorte da pesquisa, o que não significa dizer que a graduação não possa ser incluída em pesquisas futuras.

Isto posto, decidiu-se adotar a amostra não probabilística e intencional qualitativa, que, de acordo com Fragoso, Recuero e Amaral (2011), é constituída por elementos selecionados de acordo com critérios que derivam do problema de pesquisa, das características do ambiente observado, das condições e métodos de observação e análise, ou seja, em consonância com os princípios de composição da **amostragem teórica** na Teoria Fundamentada. Charmaz (2009) explica que a amostragem teórica é uma amostragem intencional, porém, assim se constitui de acordo com as categorias que alguém desenvolve a partir de suas análises, portanto, baseadas em preocupações teóricas e não em cotas.

Com a finalidade de garantir a diversidade dos tipos de sujeitos e de áreas do conhecimento, optou-se também pelo subtipo de amostra de máxima heterogeneidade, que diz respeito à seleção de elementos de todas as variações identificadas previamente de modo que represente toda a variedade do universo de pesquisa. Para garantir esse aspecto, consideramos que a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) possui 49 (quarenta e nove) áreas de avaliação que são agregadas em dois níveis e por critério de afinidade<sup>145</sup>. Note-se que, o subtipo da amostra pode ser identificado previamente, isso não quer dizer que tenha sido formada previamente sem antes analisar os achados no campo de estudo e suas características em constante comparação com os dados, conforme preconiza a Teoria Fundamentada.

O primeiro nível é constituído por três colégios (Ciências da Vida, Humanidades, Ciências Exatas, Tecnológicas e Multidisciplinar) e o segundo nível é formado por nove grandes áreas do conhecimento. Assim, foi estabelecido que seriam selecionados para as entrevistas individuais semiestruturadas (Apêndice B) pelo menos um representante de cada uma das principais grandes áreas do conhecimento que constam na Tabela de Área do Conhecimento da Capes<sup>146</sup>, a saber: Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, Ciências Agrárias, Ciências Sociais Aplicadas, Ciências Humanas, Linguística, Letras e Artes, e Multidisciplinar, conforme está distribuído no quadro 11 a seguir.

---

<sup>145</sup> Disponível em: <https://www.capes.gov.br/avaliacao/sobre-as-areas-de-avaliacao>

<sup>146</sup> Disponível em: <https://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>.



Quadro 11 - Distribuição dos sujeitos selecionados para a entrevista

(continua)

<b>COLÉGIO DE CIÊNCIAS DA VIDA</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
20000006	Ciências Biológicas	Oceanografia / Ciências Marinhas Tropicais	Discente (Doutorado)	1
40000001	Ciências da Saúde	Odontologia	Docente	1
50000004	Ciência Agrárias	Agronomia / Ciência do Solo	Docente	1
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>3</b>
<b>COLÉGIO DE HUMANIDADES</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
60000007	Ciências Sociais Aplicadas	Administração	Docente	1
		Finanças	Docente	1
70000000	Ciências Humanas	Geografia <sup>147</sup>	Discente	1
		Psicologia	Docente	1
80000002	Linguística, Letras e Artes	Letras Inglês	Docente	1
		Letras Espanhol	Discente (Mestrado)	1
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>6</b>

<sup>147</sup> A Geografia é uma ciência ampla e pode estar presente nas Geociências (Geografia Física) e ainda estar contida nas Ciências Humanas, conforme a Tabela de Área do Conhecimento da Capes.

Quadro 11 - Distribuição dos sujeitos selecionados para a entrevista

(conclusão)

<b>COLÉGIO DE CIÊNCIAS EXATAS, TECNOLÓGICAS E MULTIDISCIPLINAR</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
10000003	Ciências Exatas e da Terra	Ciência da Computação	Técnico-Administrativo em Educação (Mestrado)	1
30000009	Engenharias	Engenharia Civil / Engenharia Sanitária / Recursos Hídricos	Discente (Mestrado)	1
		Engenharia de Produção	Docente	1
90000005	Multidisciplinar	-	-	-
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>3</b>
<b>TOTAL GERAL DE PARTICIPANTES</b>				<b>12</b>

Fonte: Elaborado pela autora (2019).

A grande área denominada como 'Multidisciplinar' não tem nenhum representante, entretanto, considera-se que todas as grandes áreas supracitadas já foram contempladas dentro desse universo, então, desconsideramos esta última, apesar de ter feito diversas buscas ativas na captação de pesquisadores da área de Biotecnologia e Materiais que poderiam ocupar essa classe. Na próxima seção, serão apresentados os procedimentos de identificação das categorias de análise e os resultados encontrados.

#### 4 ANÁLISE DOS DADOS E DISCUSSÃO DOS RESULTADOS

Nesta seção serão apresentados os resultados obtidos durante a investigação que envolveu a coleta por meio de questionário, seguida da realização de entrevistas e a montagem e criação de um minicurso online sobre dados de pesquisa.

Os resultados dessas etapas são produtos dos instrumentos que procuraram captar a percepção dos pesquisadores da UFC, sobre os dados de pesquisa e verificar a situação institucional em relação ao tratamento dos dados pela instituição. Esse foi o pontapé inicial para fazer um diagnóstico prévio sobre o tema e também teve como objetivo auxiliar na identificação de lacunas institucionais relacionadas aos serviços e conhecimentos sobre o assunto por parte da comunidade pertencente à Pós-graduação.

Prosseguindo para a segunda fase, a pesquisa de campo teve início com algumas visitas informais realizadas nos *campi* Benfica, Pici e Porangabuçu, para buscar informações e tentar identificar alguns ambientes da UFC produtores de dados de pesquisa. O campus Labomar não foi visitado, entretanto, um dos entrevistados é vinculado ao referido local. Informalmente, houve a descoberta de alguns ambientes produtores de dados de pesquisa: identificou-se através do site da UFC alguns laboratórios existentes na instituição, e também por meio de conversas com usuários e sujeitos de pesquisa. Um desses espaços já é conhecido pela Biblioteca de Ciências Humanas da UFC, como é o caso de um dos laboratórios do curso de Psicologia que, inclusive, permitiu a entrada da pesquisadora e aceitou colaborar em todas as etapas, conforme será descrito adiante. As visitas ocorreram nos mesmos dias marcados com os sujeitos de pesquisa que aceitaram ser entrevistados. Assim, por uma questão de facilidade, foi útil unir a ocasião da realização da entrevista com a visita a fim de ganhar tempo e coletar o máximo de dados possíveis no mesmo dia.

Na entrevista, o roteiro semiestruturado permitiu que outras perguntas pudessem ser feitas caso fosse necessário, e o encadeamento e a ordem foi semelhante as do questionário, todavia, algumas questões foram elaboradas para a confirmação ou o aprofundamento dos primeiros achados de pesquisa, e, a partir disso, com a finalidade de extrair as categorias de análise deste estudo, procedeu-se ao processo de codificação que será descrito a seguir.

#### 4.1 Identificação das categorias de análise

Esta subseção descreve a síntese da extração dos códigos e escolha das categorias. Para isso, cada um dos instrumentos de coleta – questionário e entrevista, tiveram propósitos específicos: enquanto um visava fazer um levantamento acerca da percepção sobre os dados de pesquisas, da noção e do conhecimento das práticas relacionadas, como o gerenciamento e o compartilhamento, o outro buscou um aprofundamento e investigação mais acurada acerca das práticas e dos comportamentos diante das questões levantadas. Por isso, os instrumentos foram elaborados pensando na dinamização da relação de extração e estabelecimento das categorias *a posteriori*.

Passado o período de aplicação do questionário, os dados coletados foram transpostos para uma planilha do Google Formulários a fim de iniciar o processo de análise inicial, codificação e preparação para a extração das primeiras categorias, simultaneamente à aplicação das primeiras entrevistas. Deste modo, as respostas do questionário foram tratadas e geraram dados estatísticos apenas para apoiar a análise qualitativa. Além disso, as respostas abertas sobre a percepção dos dados de pesquisa, Ciência Aberta e os comentários foram processadas no software Atlas.ti com o intuito de atribuir os primeiros códigos e subsidiar a extração das categorias relacionadas às respostas do questionário.

A utilização do Atlas.ti (versão 8) se deu exclusivamente com o objetivo de construir as categorias à luz da Teoria Fundamentada em Dados. Trata-se de um processo rigoroso e detalhado, exige a quebra dos dados (microanálise) por meio da análise linha a linha, atribuição de códigos, organização e agrupamento da codificação, e finalmente a extração das categorias.

Rememorando o referido conceito, a **codificação** é o procedimento em que os dados são divididos, conceitualizados e nos quais se estabelecem suas relações. Trata-se de um processo analítico e tem como objetivo construir teorias, estabelecer o devido rigor metodológico ao processo científico, auxiliando no desenvolvimento do fundamento, da densidade, da sensibilidade e da integração necessárias à criação de uma teoria. (CORBIN; STRAUSS, 1990; CASSIANI; CALIRI; PELÁ, 1996).

Após a codificação de todos os dados, passou-se a próxima etapa de redução das categorias e a delimitação da teoria. Para Charmaz (2009, p. 70), “A codificação



Figura 25 – Relação dos códigos do questionário



Fonte: Dados da Pesquisa (2019).

Ressalta-se que a atribuição dos códigos e o tratamento dos dados poderia ter sido feito também manualmente, mas, houve preferência pelo uso deste software por ter sido indicado em algumas publicações científicas sobre a *Grounded Theory* (FERNANDES; MAIA, 2001; STRAUSS; CORBIN, 2008; TAROZZI, 2011), inclusive

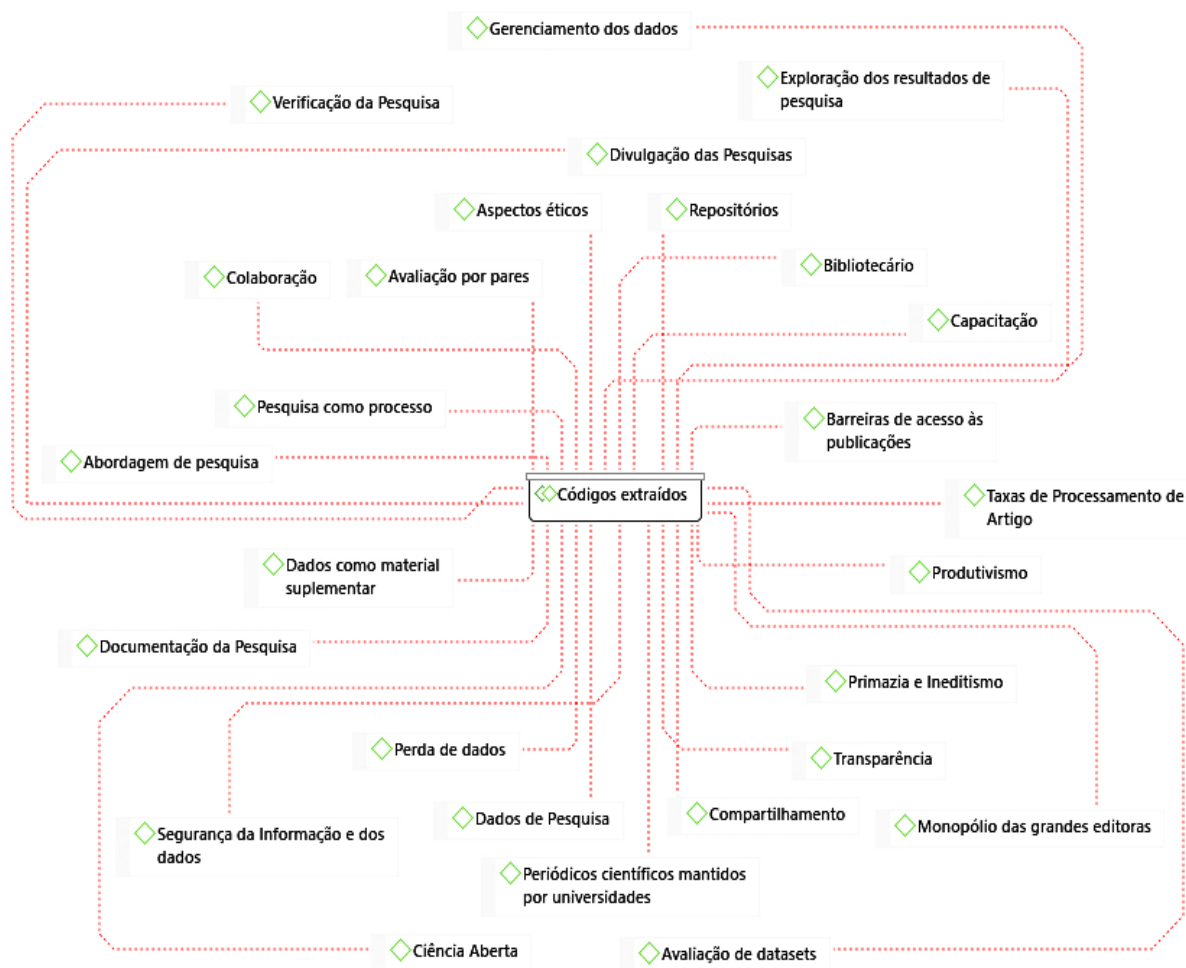
pelo fato de reunir todas as funções necessárias para manipular e processar os dados de acordo com o referido método, tais como: a) a extração manual ou automática das categorias; b) a codificação – incluindo a codificação *in vivo*, marca inconfundível da Teoria Fundamentada –; c) ferramentas de anotações, comentários e construção de memorandos; d) ótima usabilidade e facilidade de recursos voltados para este tipo de método; entre outros benefícios.

O processo de codificação não se concluiu nessa etapa, considerando que ainda se torna necessário proceder à aplicação das entrevistas, transcrevê-las e em seguida iniciar o mesmo processo de análise, codificação, categorização e agrupamento dos códigos em categorias, sejam elas novas ou existentes e realizar a comparação constante e o confronto com os dados coletados no questionário.

Na sequência, semelhante tratamento foi empregado às entrevistas com os pesquisadores: todas as respostas foram integralmente transcritas, anonimizadas, codificadas manualmente, salvas em arquivos individuais de extensão \*.txt e com a atribuição de um código para cada um dos sujeitos, e, finalmente, processadas no Atlas.ti para a etapa de codificação e formação das categorias.

Assim, procedeu-se ao carregamento dos arquivos das entrevistas no Atlas.ti a fim de se extrair os códigos e compará-los com os códigos extraídos do questionário, para somente então fazer emergir as categorias de análise. Inicialmente, foram extraídos das entrevistas 96 (noventa e seis) códigos, mas, posteriormente foram sendo agrupados e refinados com o auxílio do software, perfazendo o total de 27 (vinte e sete) códigos. O software gera uma lista desses primeiros códigos (Figura 26).

Figura 26 - Primeiros códigos extraídos das entrevistas



Fonte: Dados da Pesquisa (2019).

Após o processo de extração dos códigos em ambos os instrumentos (questionário e entrevista) procedeu-se ao refinamento dos códigos com o auxílio do Atlas.ti, o que pode ocorrer ao longo do processo de redução. O referido software também permite ter uma visão das relações das categorias com os códigos relacionados (Figura 27). Esse tipo de funcionalidade também é útil para que se possa enxergar visualmente o esquema das categorias.



Figura 27 - Relação das categorias e seus códigos



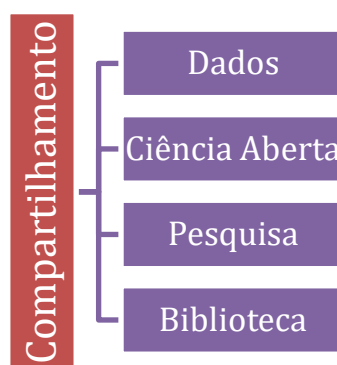
Fonte: Dados da Pesquisa (2019).

É confirmado na literatura científica por Strauss e Corbin (2008) que nas primeiras análises dos dados coletados geram-se muitos dados, mas, a partir da evolução da análise e do método da comparação constante, os conceitos vão se aglutinando e formando as categorias: essas categorias representam **temas**, conceitos mais abstratos e de nível superior que agrupam palavras (conceitos) que representam eventos, acontecimentos, situações e problemas descritos nos dados.

Dando prosseguimento ao processo, é preciso esclarecer que essa etapa ocorreu pelo menos três vezes, tendo em vista que a primeira coleta foi pelo questionário, seguido da primeira série de entrevistas, e, por fim a consecução de mais uma série de entrevistas, pois, como recomendação do método, aconselha-se não realizar todas as entrevistas de uma vez, afinal, é preciso ter em mente o constante retorno ao campo para maximizar a amostragem teórica de diferentes grupos, investigar semelhanças e diferenças entre as informações (RICHARDSON, 2017; STRAUSS; CORBIN, 2008).

Ao longo do percurso da pesquisa, do processo de refinamento das categorias fundamentadas nos dados e após diversas análises, percebeu-se que o “compartilhamento” se repetia constantemente ou mantinha algum elo entre todas as categorias estabelecidas, assim sendo, portanto, considerada como a categoria principal e emergente desta pesquisa (Figura 28). Segundo Creswell (2014), a categoria central é o conceito mais abstrato e de nível mais elevado do estudo, sendo esta a fibra teórica que une as demais categorias num todo.

Figura 28 – Categoria principal e categorias (Temas)



Fonte: Dados da Pesquisa (2019).

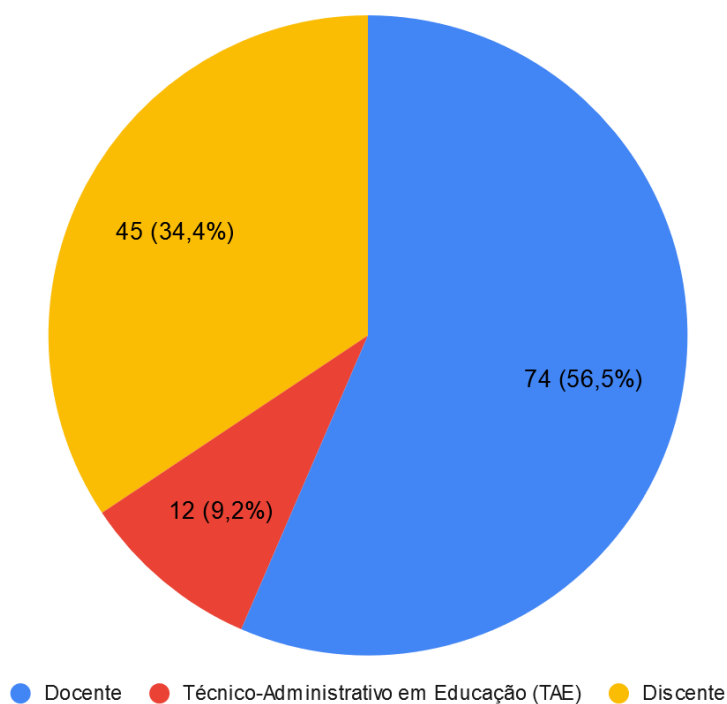
Assim, apesar de ser a gestão de dados de pesquisa o tema central dessa dissertação, através do processo de comparação constante e refinamento das categorias fundamentadas nos dados, percebeu-se que as categorias emergentes com base no *corpus* das respostas dos pesquisadores (questionário e entrevista), apresentam forte interesse na questão do compartilhamento, portanto, representam de maneira fidedigna o fenômeno em estudo. Em face ao exposto, a etapa de elaboração das categorias foi finalizada, passando ao próximo passo da pesquisa correspondente a apresentação dos resultados e análises.

## 4.2 Sondagem e investigação das práticas e necessidades dos pesquisadores

A etapa inicial da pesquisa se deu por meio da coleta de dados via questionário, intitulado “Gestão de Dados de Pesquisa e Ciência Aberta: Percepção dos Pesquisadores da Universidade Federal do Ceará” foi elaborado e executado por meio dos Formulários Google. O questionário foi destinado aos pesquisadores dos cursos de Pós-Graduação da UFC (docentes e discentes).

Dando prosseguimento a apresentação dos resultados, com base no gráfico 1, houve porcentagem expressiva de respostas dos docentes (56,5%), seguidos de 34,4% referente aos discentes e com apenas 9,2% da categoria relacionada aos Técnico-Administrativos em Educação (TAE). Em relação ao gênero, responderam ao questionário 58,3% homens e 41,7% mulheres, distribuídos nas respectivas faixas etárias: 21 a 30 anos (26,5%), 31 a 40 anos (31,1%), 41 a 50 anos (22%) e 51 anos ou mais (20,5%).

Gráfico 1 - Categoria de respondentes do questionário



Fonte: Dados da Pesquisa (2019).

A tabela 2 a seguir condensa as informações detalhadas sobre o perfil dos respondentes, agregando as informações de gênero, faixa etária, categoria, titulação, seguidas das respectivas quantidades.

Tabela 2 - Síntese das informações sobre o perfil dos respondentes

(continua)

<i>GÊNERO</i>	<i>FAIXA ETÁRIA</i>	<i>CATEGORIA</i>	<i>TITULAÇÃO</i>	<i>CATEGORIA (TOTAL)</i>	<i>TITULAÇÃO (TOTAL)</i>	<i>GÊNERO (TOTAL)</i>	<i>FAIXA ETÁRIA (TOTAL)</i>				
Feminino	21 a 30 anos	Discente	Especialização	4	4	4	4				
			Graduação	5	5	5	5				
			Mestrado	5	5	5	5				
	<b>Discente Total</b>				14	14	14	14			
		Docente	Doutorado		1	1	1	1			
				<b>Docente Total</b>				1	1	1	1
				<b>21 a 30 anos (Total)</b>				15	15	15	15
	31 a 40 anos	Discente	Especialização		1	1	1	1			
				Mestrado		2	2	2	2		
					<b>Discente Total</b>				3	3	3
		Docente	Doutorado		3	3	3	3			
				Mestrado		1	1	1	1		
					<b>Docente Total</b>				4	4	4
		Técnico-Administrativo em Educação (TAE)		Doutorado		1	1	1	1		
					Especialização		2	2	2	2	
Graduação						1	1	1	1		
Mestrado						2	2	2	2		
<b>Técnico-Administrativo em Educação (TAE) Total</b>				6	6	6	6				
<b>31 a 40 anos (Total)</b>				13	13	13	13				
41 a 50 anos	Discente	Especialização		1	1	1	1				
			Mestrado		2	2	2	2			
				<b>Discente Total</b>				3	3	3	3
	Docente	Doutorado		9	9	9	9				
			Mestrado	1	1	1	1				

Tabela 2 - Síntese das informações sobre o perfil dos respondentes  
(continuação)

<i>GÊNERO</i>	<i>FAIXA ETÁRIA</i>	<i>CATEGORIA</i>	<i>TITULAÇÃO</i>	<i>CATEGORIA (TOTAL)</i>	<i>TITULAÇÃO (TOTAL)</i>	<i>GÊNERO (TOTAL)</i>	<i>FAIXA ETÁRIA (TOTAL)</i>
		<b>Docente Total</b>		10	10	10	10
		Técnico-Administrativo em Educação (TAE)	Doutorado	1	1	1	1
			Mestrado	2	2	2	2
		<b>Técnico-Administrativo em Educação (TAE) Total</b>		3	3	3	3
	<b>41 a 50 anos (Total)</b>			16	16	16	16
	51 anos ou mais	Discente	Doutorado	1	1	1	1
		<b>Discente Total</b>		1	1	1	1
		Docente	Doutorado	9	9	9	9
		<b>Docente Total</b>		9	9	9	9
		Técnico-Administrativo em Educação (TAE)	Especialização	1	1	1	1
		<b>Técnico-Administrativo em Educação (TAE) Total</b>		1	1	1	1
	<b>51 anos ou mais (Total)</b>			11	11	11	11
<b>Feminino Total</b>				55	55	55	55
<b>Masculino</b>	21 a 30 anos	Discente	Graduação	8	8	8	8
			Mestrado	6	6	6	6
		<b>Discente Total</b>		14	14	14	14
		Docente	Doutorado	1	1	1	1
			Mestrado	4	4	4	4
		<b>Docente Total</b>		5	5	5	5
		Técnico-Administrativo em Educação (TAE)	Mestrado	1	1	1	1
		<b>Técnico-Administrativo em Educação (TAE) Total</b>		1	1	1	1
	<b>21 a 30 anos (Total)</b>			20	20	20	20

Tabela 2 - Síntese das informações sobre o perfil dos respondentes (conclusão)

GÊNERO	FAIXA ETÁRIA	CATEGORIA	TITULAÇÃO	CATEGORIA (TOTAL)	TITULAÇÃO (TOTAL)	GÊNERO (TOTAL)	FAIXA ETÁRIA (TOTAL)	
	31 a 40 anos	Discente	Doutorado	2	2	2	2	
			Especialização	1	1	1	1	
			Graduação	1	1	1	1	
			Mestrado	5	5	5	5	
		<b>Discente Total</b>		<b>9</b>	<b>9</b>	<b>9</b>	<b>9</b>	
		Docente	Doutorado	15	15	15	15	
			Mestrado	2	2	2	2	
		<b>Docente Total</b>		<b>17</b>	<b>17</b>	<b>17</b>	<b>17</b>	
		Técnico-Administrativo em Educação (TAE)	Doutorado		1	1	1	1
				Mestrado	1	1	1	1
	<b>Técnico-Administrativo em Educação (TAE) Total</b>			<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	
	<b>31 a 40 anos Total</b>		<b>28</b>	<b>28</b>	<b>28</b>	<b>28</b>		
	41 a 50 anos	Discente	Mestrado	1	1	1	1	
			<b>Discente Total</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	
		Docente	Doutorado	11	11	11	11	
	Mestrado		1	1	1	1		
	<b>Docente Total</b>		<b>12</b>	<b>12</b>	<b>12</b>	<b>12</b>		
	<b>41 a 50 anos (Total)</b>		<b>13</b>	<b>13</b>	<b>13</b>	<b>13</b>		
	51 anos ou mais	Docente	Doutorado	16	16	16	16	
			<b>Docente Total</b>	<b>16</b>	<b>16</b>	<b>16</b>	<b>16</b>	
	<b>51 anos ou mais (Total)</b>		<b>16</b>	<b>16</b>	<b>16</b>	<b>16</b>		
<b>Masculino Total</b>				<b>77</b>	<b>77</b>	<b>77</b>	<b>77</b>	
<b>TOTAL GERAL</b>				<b>132</b>	<b>132</b>	<b>132</b>	<b>132</b>	

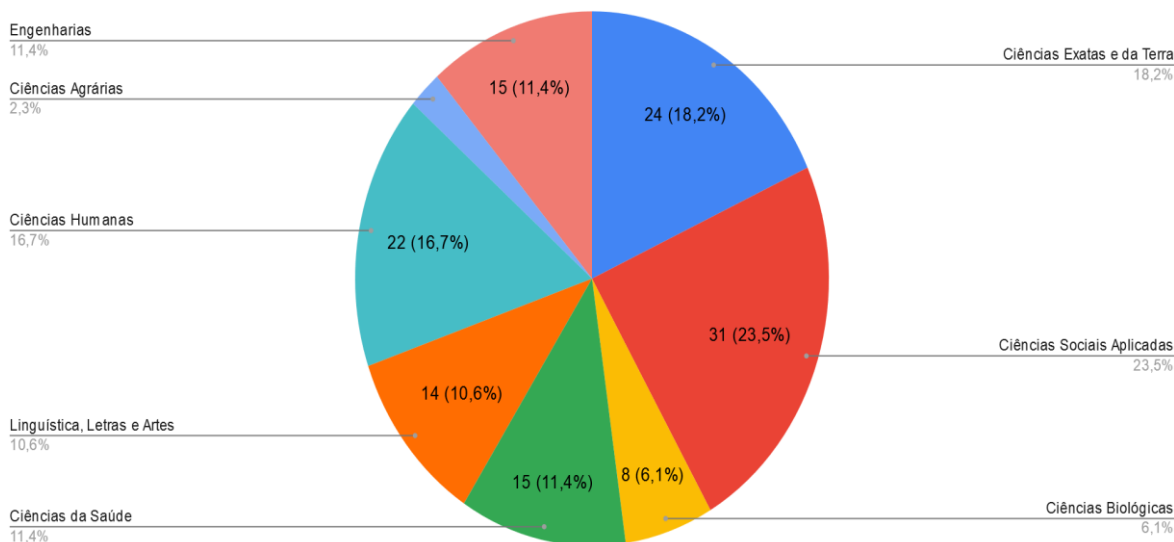
Fonte: Dados da Pesquisa (2019).

A partir do gráfico 2, figuram as porcentagens de respondentes de acordo com as áreas de conhecimento. Nota-se que o maior número de respostas

correspondem à área de Ciências Sociais Aplicadas (23,5%), seguidos das respectivas áreas: Ciências Exatas e da Terra (18,2%), Ciências Humanas (16,7%), Ciências da Saúde (11,4%), Engenharias (11,4%), Linguística, Letras e Artes (10,6%), Ciências Biológicas (6,1%) e Ciências Agrárias (2,3%).

Gráfico 2 - Porcentagem de participação por Área do Conhecimento

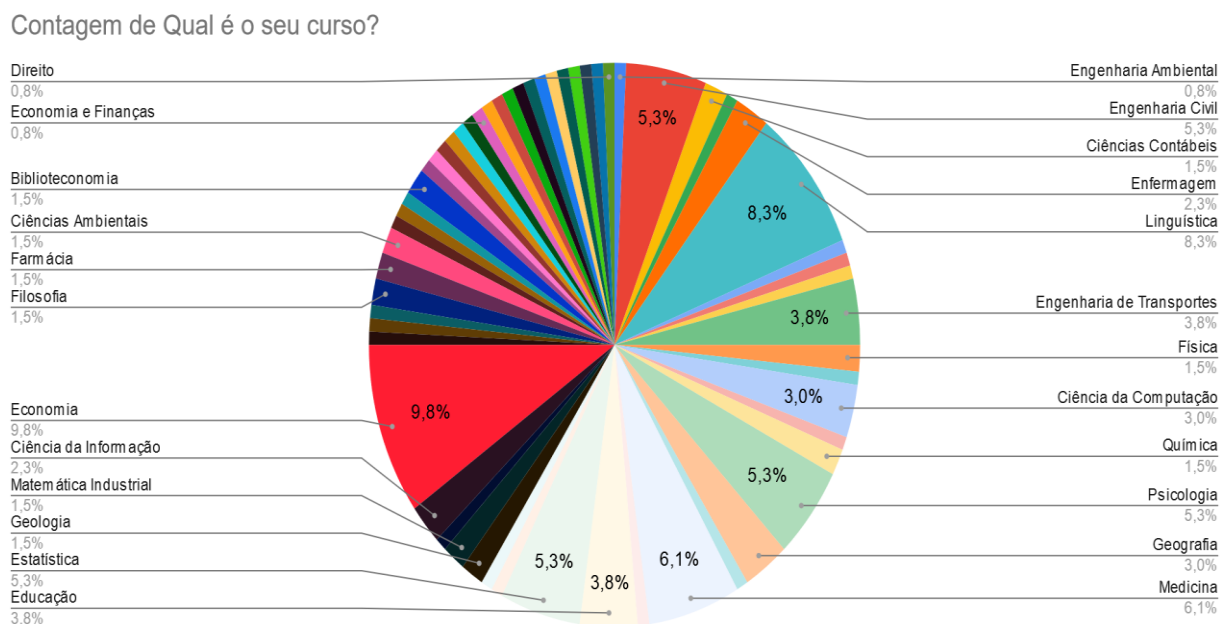
Contagem de Área do Conhecimento



Fonte: Dados da Pesquisa (2019).

Vários cursos responderam ao questionário, conforme o gráfico 3 abaixo. O maior percentual corresponde ao curso de Pós-graduação em Economia corresponde a 9,8%, sendo um dos maiores colaboradores desta pesquisa, seguidos dos cursos de Letras (Linguística) (8,3%), Medicina (6,1%), Engenharia Civil (5,3%), Estatística (5,3%) e Psicologia (5,3%).

Gráfico 3 – Porcentagem de participação dos cursos



Fonte: Dados da Pesquisa (2019).

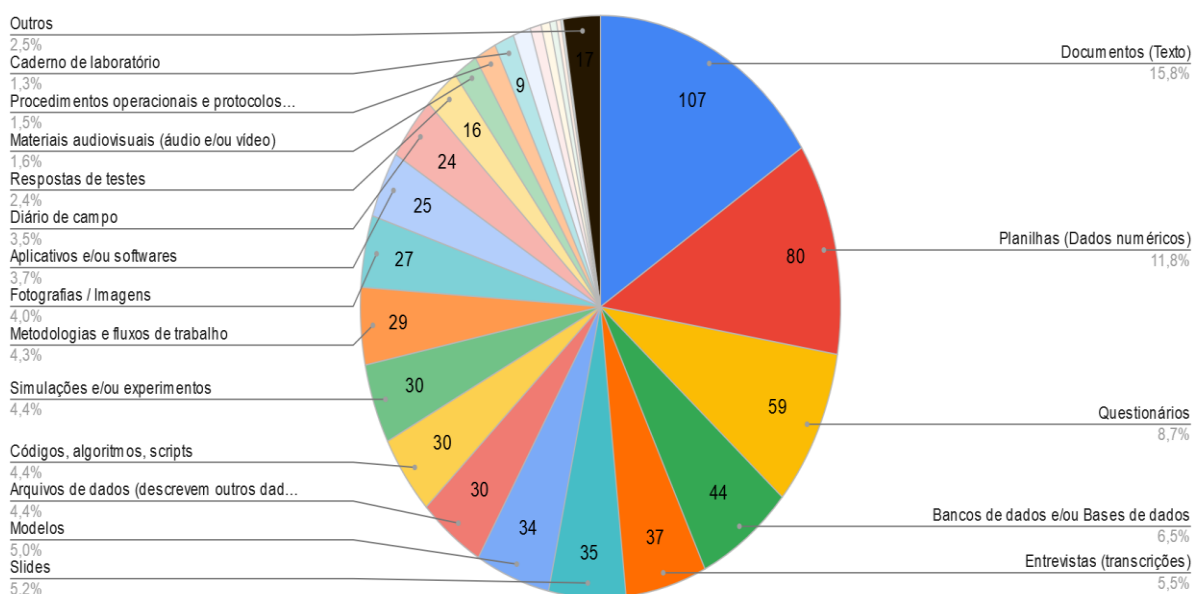
Perguntou-se aos usuários qual é o tipo de abordagem usado em suas pesquisas (quantitativa, qualitativa ou mista). De acordo com as respostas, a maioria das pesquisas desenvolvidas na UFC são de abordagem mista (quantitativa e qualitativa) e correspondem a 48,9%. Em segundo lugar, com 29,8%, são de pesquisas quantitativas e em terceiro, com 21,4% as pesquisas qualitativas.

Para sondar quais os tipos de dados produzidos nas pesquisas desenvolvidas na UFC (Gráfico 4), dedicamos uma questão do questionário a esse respeito. A quantidade e variedade de dados foi impressionante, inclusive, outros tipos foram relatados pelos respondentes através da opção outros, na qual era possível marcar e escrever qual o tipo de documento não constava nessa lista. Sobre esses tipos que integraram a alternativa 'outros', destacam-se as cartografias sociais, os mapas georreferenciais e o diário de aprendizagem.



Gráfico 4 - Tipos de dados gerados em pesquisas na UFC

## Tipos de dados gerados em pesquisas na UFC



Fonte: Dados da Pesquisa (2019).

O questionário continha uma pergunta aberta que visava captar a percepção e o entendimento dos pesquisadores sobre os dados de pesquisa. As respostas foram as mais variadas possíveis, o que é perfeitamente compreensível, principalmente se levar em consideração que a própria definição de 'o que é um dado?' é inesgotável até na literatura científica, além de não ser uma tarefa simples definir ou expressar o que é um dado ou dado de pesquisa, a exemplo de uma das respostas registradas que se limitou a informar "o próprio nome já diz". A partir desse comentário, podemos inferir que parece difícil a esse respondente exprimir o que entende sobre dado de pesquisa.

Outras respostas classificaram dados de pesquisas como 'resultados de pesquisa', 'variáveis matemáticas', 'conjuntos de informações', 'observações', 'fontes de informação', '*corpus* e *corpora*' (coletânea ou reunião de textos e documentos, pode ser formado por um assunto ou tema e tem tamanhos variados), sendo estes últimos bastante utilizados por pesquisadores da linguística no desenvolvimento de suas pesquisas. Foram selecionados alguns exemplos dessas percepções sobre os dados de pesquisa. Escolhemos separar e classificar de forma simples as respostas (Dado, Informação, Conhecimento e Insumo), conforme se apresenta no quadro 12.

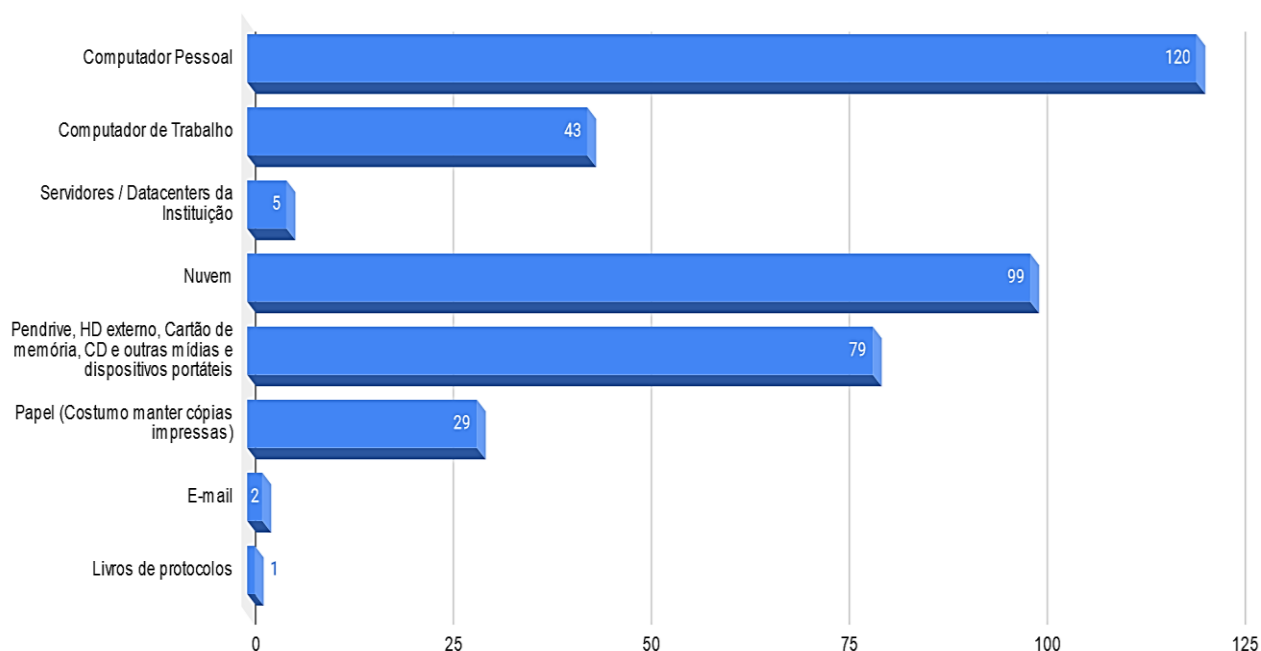
Quadro 12 - Exemplos de algumas das percepções e conceitos sobre dados de pesquisa informados pelos respondentes

O QUE VOCÊ ENTENDE POR 'DADOS DE PESQUISA'?			
DADO	INFORMAÇÃO	CONHECIMENTO	INSUMO
<b>Dados</b> oriundos da coleta efetuada relacionada a pesquisa.	Todas as <b>informações</b> do seu trabalho	Todos os elementos informativos relativos à construção da pesquisa e do <b>conhecimento</b> sobre o objeto de estudo e suas interrelações com a realidade complexa na qual existe.	<b>Produtos</b> de coleta obtidos a partir de uma metodologia aplicada, geralmente com instrumentos de pesquisa.
<b>Dados</b> gerados a partir de uma investigação com objetivos bem definidos.	Todas aquelas <b>informações</b> usadas como base ou provenientes da pesquisa	-	São os <b>insumos</b> da pesquisa.
Os valores referentes à pesquisa, sejam preliminares ou definitivos	<b>Informações</b> coletadas durante a pesquisa	-	Corpora adquirida para defender uma hipótese de pesquisa
Aqueles gerados em função de pesquisas científicas realizadas.	<b>Informações</b> geradas a partir da exploração científica de um tema de interesse	-	São os diversos tipos de <b>produtos</b> advindos de uma pesquisa. Podem ser desde dados quantitativos ou qualitativos e possuem diferentes destinações e públicos. Ao fim e ao cabo, é a resposta de determinado problema de investigação que é oferecido para a sociedade.
<b>Dados</b> coletados da realidade empírica para elaborar ou testar hipóteses.	<b>Informações</b> brutas do objeto da pesquisa	-	Os <b>insumos</b> utilizados para viabilizar a pesquisa em sua dimensão conceitual.
<b>Dados</b> de coletas primárias (pesquisa de campo) e coleta secundária (bases já existentes de acesso público)	São as <b>informações</b> básicas para realizar a análise ou construir o modelo pretendido.	-	Entendo que é o <b>produto</b> obtido de questionário respondidos, entrevistas realizadas, observações e outros meios, onde o pesquisador usará esses dados para embasar sua pesquisa.
<b>Dados</b> observados a partir de experimentos, para fins de validação da teoria.	Todas as <b>informações</b> coletadas, processadas e geradas durante a pesquisa	-	<b>Insumos</b> quantitativos que, após processados, passam a gerar informação
-	-	-	Todo material que é produzido em uma investigação científica e que serve como elementos para análises e informações sobre o problema de pesquisa.

Fonte: Dados da Pesquisa (2019).

Com o objetivo de conhecer quais são as estratégias de armazenamento e que fazem parte das boas práticas na gestão de dados de pesquisa, perguntou-se de que forma os pesquisadores costumam organizar, armazenar, manter e preservar os seus dados. As alternativas apresentadas foram as seguintes: a) Computador pessoal (90,9%); b) Computador de trabalho (32,6%); c) Servidores e Datacenters da instituição (3,8%); d) Nuvem (75%); e) Pendrive, HD externo, Cartão de memória, CD e outras mídias e dispositivos portáteis (59,8%); f) Papel (para aqueles que costumam manter uma cópia impressa extra) (22%). Nessa pergunta também havia a opção 'outros' para registro de alguma informação adicional que não fosse uma das alternativas discriminadas. Então, foram informadas ainda como estratégias de armazenamento o e-mail (1,6%) e o livro de protocolo (0,8%).

Gráfico 5 - Estratégias de armazenamento dos dados de pesquisa



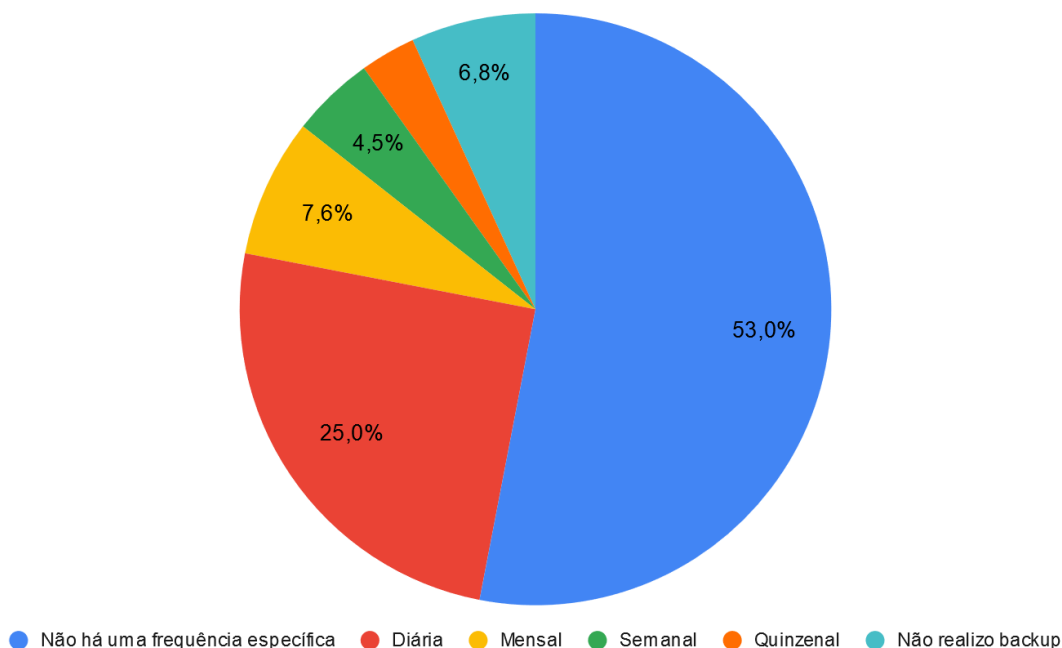
Fonte: Dados da Pesquisa (2019).

Sobre os tipos de armazenamento, percebe-se que a maioria das pessoas usa a estratégia de armazenar seus dados no computador pessoal, na nuvem e em dispositivos portáteis. O computador de trabalho é pouco utilizado para esse fim, e, o número de pessoas que salva os dados de sua pesquisa em servidores ou datacenters institucionais é ainda menor. A respeito disso, vale ressaltar que em algumas instituições internacionais que já implantaram serviços de gestão de dados, como a

Universidade de Edimburgo, por exemplo, é oferecido pela instituição a provisão do espaço de armazenamento em seus datacenters para os pesquisadores, utilizando adicionalmente medidas de controle de acesso, *backups* frequentes, entre outras medidas de segurança da informação e dos dados. Sobre o e-mail, apesar de não ser o meio mais confiável para guardar os dados de pesquisa, ainda é usado por algumas pessoas.

Concernente à regularidade da realização de backups, a maioria declarou que não há uma frequência específica, mas, executam a tarefa (53%), e, outros informaram que fazem diariamente (25%), seguidos das frequências mensais (7,6%), semanais (4,5%) e quinzenais (3%). Entre os respondentes, há a parcela de 6,8% que não fazem cópias de segurança. Claramente, os dados demonstram que os pesquisadores precisam repensar sua conduta a respeito dessa atividade, afinal, garantir que os dados serão devidamente armazenados e acessados também ensejam a garantia de que outras cópias desses materiais poderão ser recuperadas no caso de qualquer sinistro.

Gráfico 6 - Frequência de backup dos dados

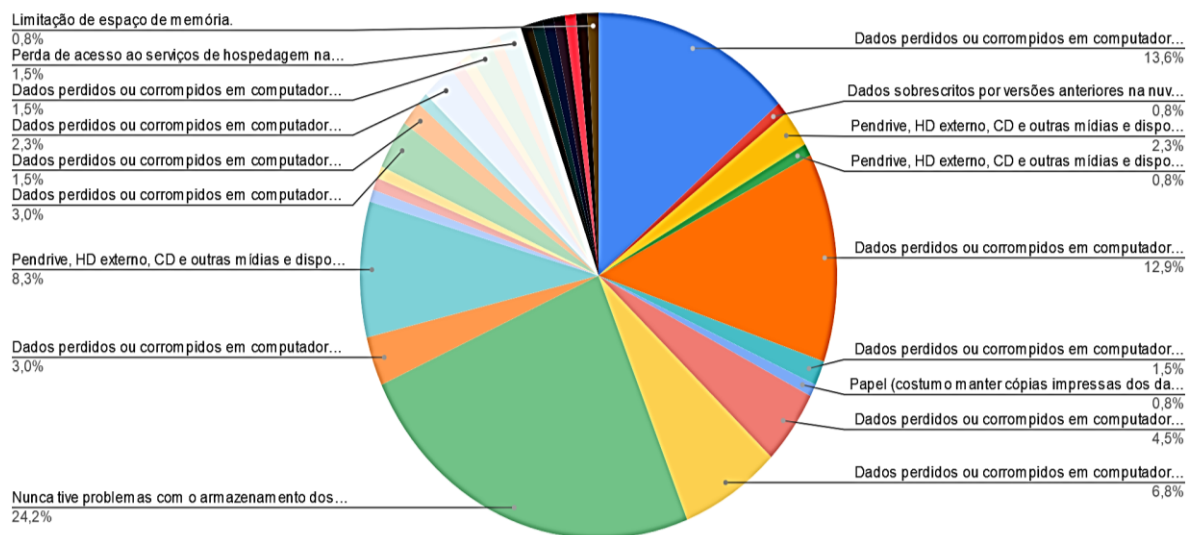


Fonte: Dados da Pesquisa (2019).

Outra pergunta feita diz respeito às experiências de perda de dados e quais os possíveis motivos. Aproximadamente 24% dos pesquisadores respondeu que

nunca tiveram problemas em relação a isso, enquanto os demais declararam já ter tido algum problema de perda dos dados em computadores danificados ou arquivos corrompidos, entre outras causas.

Gráfico 7 - Experiências de perda dos dados de pesquisa



Fonte: Dados da Pesquisa (2019).

Uma das perguntas feitas no questionário desta pesquisa tem relação direta com um setor da UFC visitado durante a pesquisa de campo realizada no Campus do Pici. A pesquisadora foi informada sobre um setor da universidade que oferece máquinas de alto desempenho e um ambiente de acesso remoto para que os pesquisadores possam rodar scripts, testar modelos, realizar simulações e processar grandes massas de dados.

Diante do exposto, a pesquisadora buscou informações detalhadas no site do Cenapad/UFC e descobriu que qualquer pessoa poderia agendar uma visita para conhecer o local, inclusive, o público externo. Assim, foi preenchido um formulário de solicitação de agendamento e enviado por e-mail, conforme as instruções constantes no site do Cenapad. Em poucas horas a resposta de confirmação foi dada e a visita pôde ser realizada no dia e horário marcado. Durante a visita, a pesquisadora informou ao servidor que a recebeu qual o motivo da visita e solicitou autorização para tirar fotos e gravar o áudio das conversas.

Figura 29 - Exemplos dos computadores de alto desempenho do Cenapad-UFC



Fonte: Dados da Pesquisa (2019).

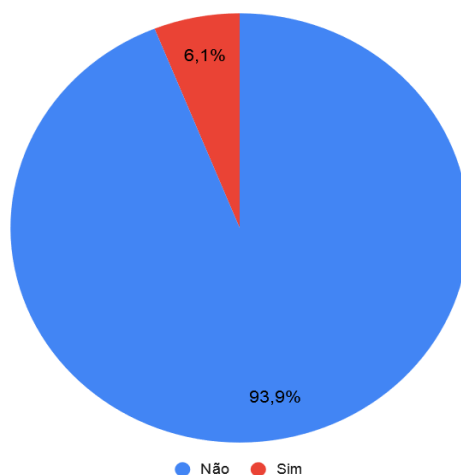
Durante a visita, o servidor e demais bolsistas presentes falaram sobre os serviços e suporte oferecido, mas, informaram que poucos pesquisadores da UFC procuravam o setor. Pesquisas relevantes e de grande porte das mais diversas áreas tinham dados armazenados nos servidores e máquinas do Cenapad/UFC, mas, isso tem um período determinado e segue as regras e políticas estabelecidas<sup>148</sup>. Além disso, deve ser realizado um cadastro antes, e, o espaço de armazenamento possui alguns limites.

O objetivo da visita era descobrir se o Cenapad/UFC também oferecia espaço de armazenamento para qualquer tipo de pesquisa, com diferentes tipos de dados e de tamanhos variados, inclusive, por quanto tempo. Contudo, o servidor esclareceu que a estrutura não comportava a oferta de espaço de armazenamento para os pesquisadores e que o objetivo do setor era oferecer esses ambientes computacionais para que os pesquisadores pudessem processar grandes volumes de dados em pouco tempo por meio dos recursos oferecidos. Um dos respondentes utilizou o espaço dos comentários para se expressar a respeito do armazenamento pela UFC: “A universidade deveria fornecer uma plataforma para que professores pesquisadores armazenassem seus dados de pesquisas.”

<sup>148</sup> Políticas do CENAPAD-UFC: <https://cenapad.ufc.br/quem-somos/politicas/>.

Constatou-se na visita que o setor não é procurado pelos pesquisadores, e, pela natureza das respostas do questionário (93,9% desconhecem o setor), pouquíssimos conhecem o Cenapad. O setor também carece de uma maior mão de obra, tendo em vista que apenas um servidor público estava respondendo pelo setor e os demais membros da equipe de trabalho eram bolsistas ou estagiários. O servidor relatou ainda que é preciso dar treinamento sobre os ambientes computacionais, computação em grid e demais softwares e recursos para alguns pesquisadores que procuram o Cenapad/UFC, mas, isso também estava comprometido, tendo em vista que muitos professores que tinham domínio desses conhecimentos tinham se desligado do setor.

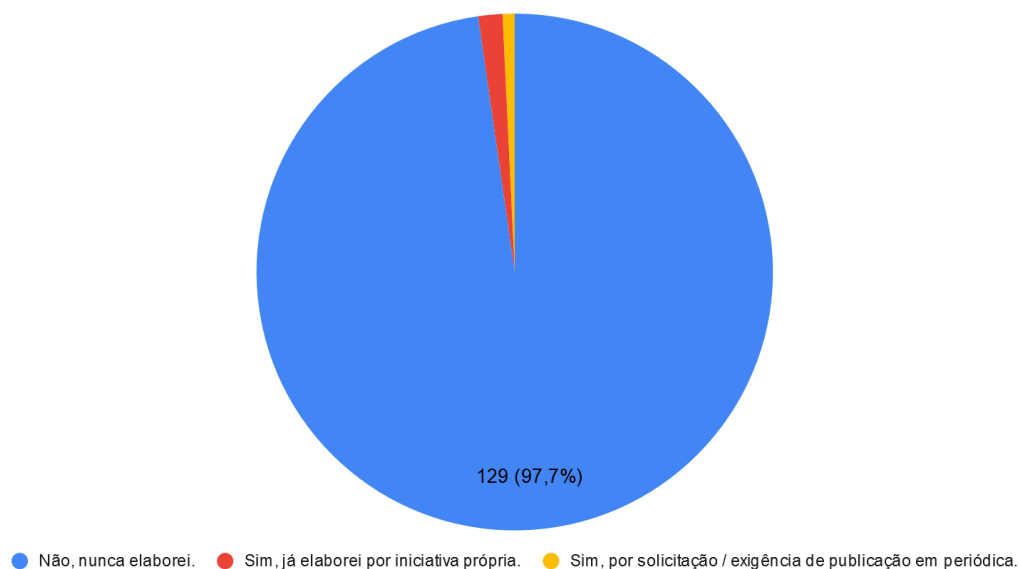
Gráfico 8 - Porcentagem de pesquisadores que conhecem o CENAPAD-UFC



Fonte: Dados da Pesquisa (2019).

Sobre os procedimentos de documentação da pesquisa científica, perguntou-se aos usuários se já elaboraram um Plano de Gestão de Dados (PGD). A maioria dos pesquisadores da UFC declarou nunca ter elaborado um PGD (97,7%), enquanto 1,5% declarou ter elaborado por iniciativa própria e 0,8% informou que fez por solicitação ou exigência de publicação periódica.

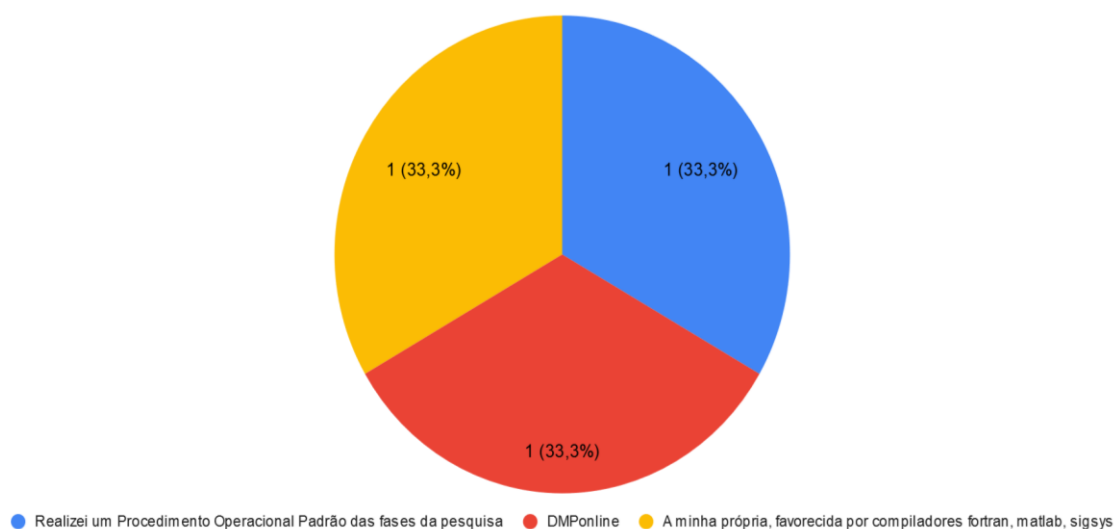
Gráfico 9 - Elaboração de um Plano de Gestão de Dados (PGD)



Fonte: Dados da Pesquisa (2019).

Para os pouquíssimos respondentes que elaboraram o PGD, solicitou-se que indicassem qual foi a ferramenta utilizada para isso. Entretanto, a partir das respostas fica claro que apenas uma pessoa de fato elaborou um PGD por meio do DMPonline. Os outros dois que declararam ter elaborado um PGD, na verdade, criaram outro tipo de documento para documentar as suas pesquisas.

Gráfico 10 - Ferramenta utilizada na elaboração do PGD



Fonte: Dados da Pesquisa (2019).



Os próximos dados e gráficos a seguir estão relacionados à subseção do questionário que se dedicou a sondar a percepção e o conhecimento da comunidade acadêmica da pós-graduação da UFC a respeito da Ciência Aberta. Como já foi explicado anteriormente neste trabalho, a decisão de também abordar alguns aspectos da Ciência Aberta se deu porque estes se relacionam de forma direta ou indireta, especialmente, no que diz respeito ao compartilhamento, afinal, é preciso investigar e comprovar, por exemplo, quais são os motivos que os pesquisadores alegam para adotar ou não esta prática. Algumas pesquisas de âmbito geral já foram feitas e os motivos que trazem as vantagens, desvantagens e as razões pelas quais os pesquisadores não compartilham são conhecidos (DIGITAL SCIENCE, 2017; ORION OPEN SCIENCE, 2018; REDE DE DADOS DE PESQUISA BRASILEIRA, 2019), mas, para os fins deste estudo, é essencial investigar o fenômeno e esse processo na UFC, identificar a situação institucional e assim delinear os planos e estratégias de abordagem sobre o assunto na instituição. Se o pesquisador irá compartilhar ou não, de fato, isso compete apenas a ele mesmo. De acordo com as palavras de Sayão e Sales (2015), a decisão é do pesquisador.

O primeiro grande bloco de perguntas de múltipla escolha traziam as práticas e processos comuns na *Open Science*. A partir da tabela 3, evidencia-se que o conhecimento sobre as práticas e processos da Ciência Aberta ainda são incipientes.

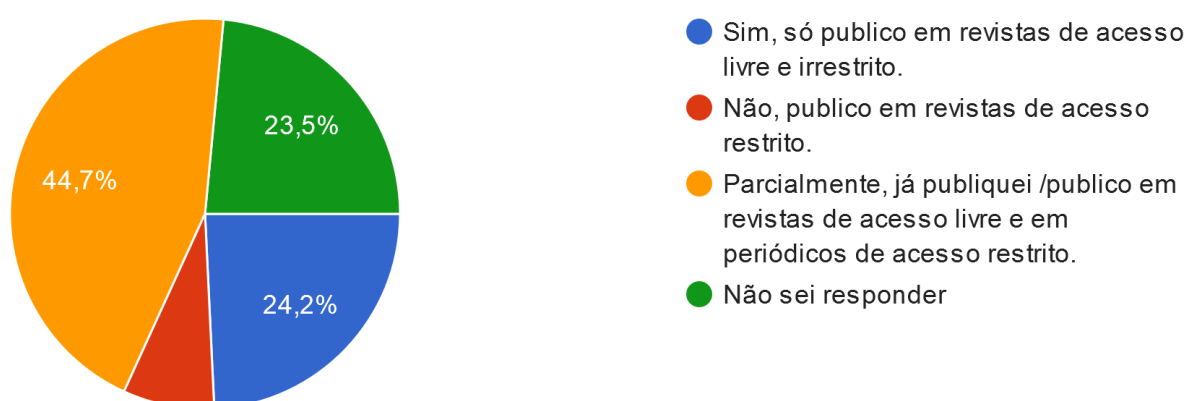
Tabela 3 - Familiaridade com as práticas e processos da Ciência Aberta (Porcentagem)

FAMILIARIDADE COM AS PRÁTICAS E PROCESSOS DA CIÊNCIA ABERTA															
	Compartilhamento de dados abertos	Gestão de dados	Dados Governamentais abertos	Reprodutibilidade	Códigos e softwares abertos	Hardware abertos	Acesso Aberto	Cadernos de laboratório abertos	Recursos Educacionais abertos	Colaboração	Advocacy	Ciência cidadã	Promoção de Políticas	Inovação aberta	Artes (OpenGLAM)
Conhecem pouco	51,50%	45,50%	50,80%	31,10%	43,20%	35,60%	28,80%	30,30%	47,70%	38,60%	22%	26,50%	33,30%	28%	22,70%
Conhecem bastante	21,20%	9,10%	26,50%	18,90%	22,70%	6,80%	65,90%	3,00%	3,00%	8,30%	2,20%	5,30%	1,50%	3,10%	1,50%
Desconhecem	27,30%	45,50%	22,70%	43,90%	34,10%	57,60%	5,30%	66,70%	66,70%	53,00%	75,80%	68,20%	65,20%	68,90%	75,80%

Fonte: Dados da Pesquisa (2019).

Ainda sobre a questão da familiaridade com os processos e práticas na Ciência Aberta, o Acesso Aberto às Publicações Científicas é um dos mais conhecidos pelos respondentes (65,9%). Fazendo o cruzamento com outras questões, relaciona-se à pergunta a respeito dos tipos de revistas que os pesquisadores da UFC costumam publicar. A porcentagem de pessoas que não soube especificar se publica em revistas de acesso aberto ou acesso restrito foi de 23,5%, maior do que o percentual de pesquisadores que afirmaram publicar apenas em periódicos de acesso restrito (7,6%). Os que publicam em revistas exclusivamente de acesso aberto, livre e irrestrito são 24,2%, e, por fim, 44,7% informaram que publicam tanto em revistas de acesso aberto e livre como em periódicos de acesso fechado.

Gráfico 11 - Sondagem sobre os tipos de revistas em que os pesquisadores publicam



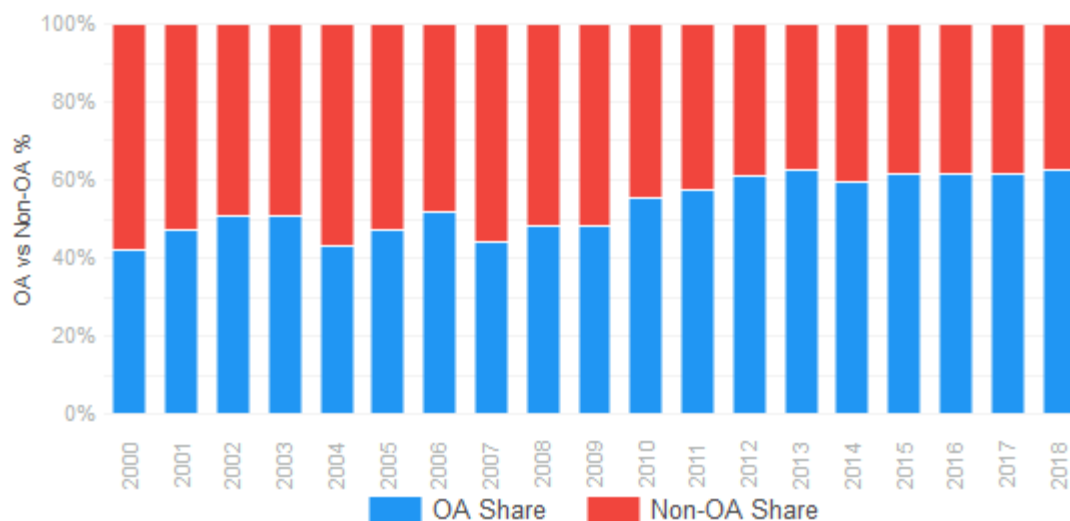
Fonte: Dados da Pesquisa (2019).

Isso se reflete nos dados apresentados pelo Wizdom.ai<sup>149</sup> sobre a produção institucional da UFC, especificamente, no gráfico da relação de artigos publicados em revistas de acesso aberto *versus* acesso restrito (OA share x Non-OA share) desde o ano 2000 a 2018, confirmando os dados coletados nesta pesquisa.

<sup>149</sup> Confira os dados sobre a produção da UFC que estão disponíveis neste link: [https://www.wizdom.ai/institution/universidade\\_federal\\_do\\_cear%C3%A1/grid.8395.7](https://www.wizdom.ai/institution/universidade_federal_do_cear%C3%A1/grid.8395.7).

Gráfico 12 - Produções da UFC em Acesso Aberto x Acesso Restrito

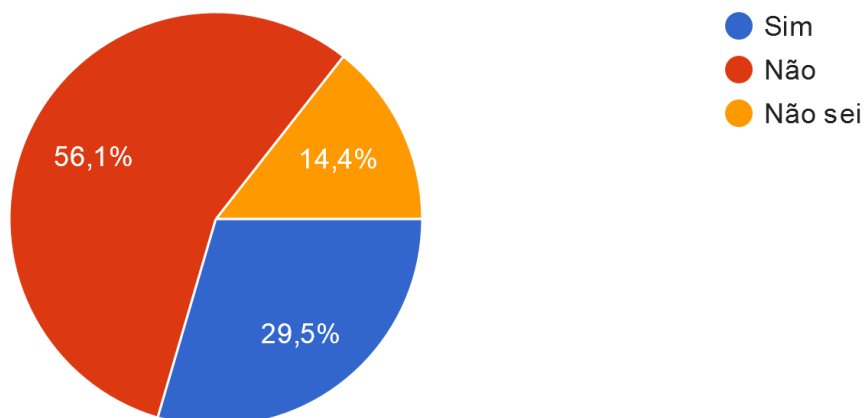
## Open Access vs Non-Open Access



Fonte: Wizdom.ai (2019).

Uma das questões levantadas pelo Acesso Aberto às publicações científicas é a disponibilização das versões pré-prints dos artigos. Por isso, indagou-se aos pesquisadores se estes costumam acessar e ler versões pré-print. A maioria afirma não ler ou acessar (56,1%), 29,5% declaram ler e acessar e 14,4% não souberam responder.

Gráfico 13 - Acesso e leitura de pré-prints por parte dos pesquisadores



Fonte: Dados da Pesquisa (2019).

Dando continuidade, também foi perguntado aos pesquisadores o motivo de lerem ou não os pré-prints, assim, uma série de afirmações foi disposta e os respondentes poderiam marcar as opções que conviessem. A maioria dos usuários

indicou que não publica versões pré-print de seus artigos (64,4%); os 12,9% apontam não confiar em revisões que não foram avaliadas por pares e apenas 3% publica pré-prints.

Para investigar qual a noção ou entendimento que os usuários ligados à pós-graduação têm sobre a Ciência Aberta, perguntamos: “O que você entende por Ciência Aberta?”. Curiosamente, dos 132 respondentes, 22 declararam que não sabem responder. Isso corrobora com os dados apresentados acima a respeito da familiaridade com as práticas e processos na Ciência Aberta. As demais respostas foram registradas, algumas foram selecionadas como exemplo e estão elencadas abaixo.

Quadro 13 - Percepção sobre a Ciência Aberta

(continua)

PERCEPÇÃO DOS USUÁRIOS SOBRE A CIÊNCIA ABERTA
Acessível
Acesso à Ciência
Acesso irrestrito a dados
Aquela em que o pesquisador compartilha o acesso aos seus dados de pesquisa, por ele coletado.
Aquela que é acessível inclusive a quem não está realizando a prática em questão.
Aquela que produz conhecimento sistematizado ao alcance de quem tiver interesse. Por exemplo, com o acesso ao banco de dados original no intuito de reproduzir a pesquisa sob as mesmas condições iniciais ou condições diferentes.
Aquela que se baseia em transparência, repetibilidade, acesso gratuito a banco de dados e softwares
Aqueles trabalhos científicos com acesso aberto
Área do conhecimento que privilegia maior clareza na obtenção, comprovação e análise dos dados de pesquisa, bem como dos indivíduos envolvidos.
As bases de dados das pesquisas científicas, bem como as publicações possuem acesso aberto a todos
Atividades científicas que podem ser compartilhadas pela comunidade científica e pela sociedade
Ciência acessível
Ciência acessível a toda a população, inclusive leigos.
Ciência colaborativa
Ciência com acesso a todos
Ciência democrática
Ciência disponível, cujos resultados (de forma entendível seja para população ou comunidade científica) estão disponíveis para toda a sociedade.
Ciência em que há partilha de conhecimento afetivo, cognitivo e psicomotor
Ciência praticada sem o sigilo dos dados, em uma matriz colaborativa ampla.
Ciência que pode ser acessada ou distribuída sem pagamento de direitos autorais.
Ciência transparente, de livre acesso a todos.
Colaboração em pesquisa no desenvolvimento e disseminação do conhecimento em prol da inovação em diversas áreas de atuação.
Material compartilhado para domínio público.
Movimento atual, advindo de uma "crise da reprodutibilidade/replicação", que advoga pela divulgação pública e aberta dos dados da pesquisa bem como dos scripts e métodos de análise para que outros pesquisadores possam trabalhar os dados e replicar o estudo; advoga também pela publicação pré-print (para recebimento de feedback de pares), pelo parecer não-cego, com interação real e significativa entre autores e pareceristas, pelas publicações abertas e gratuitas, por exemplo.

## Quadro 13 - Percepção sobre a Ciência Aberta

(conclusão)

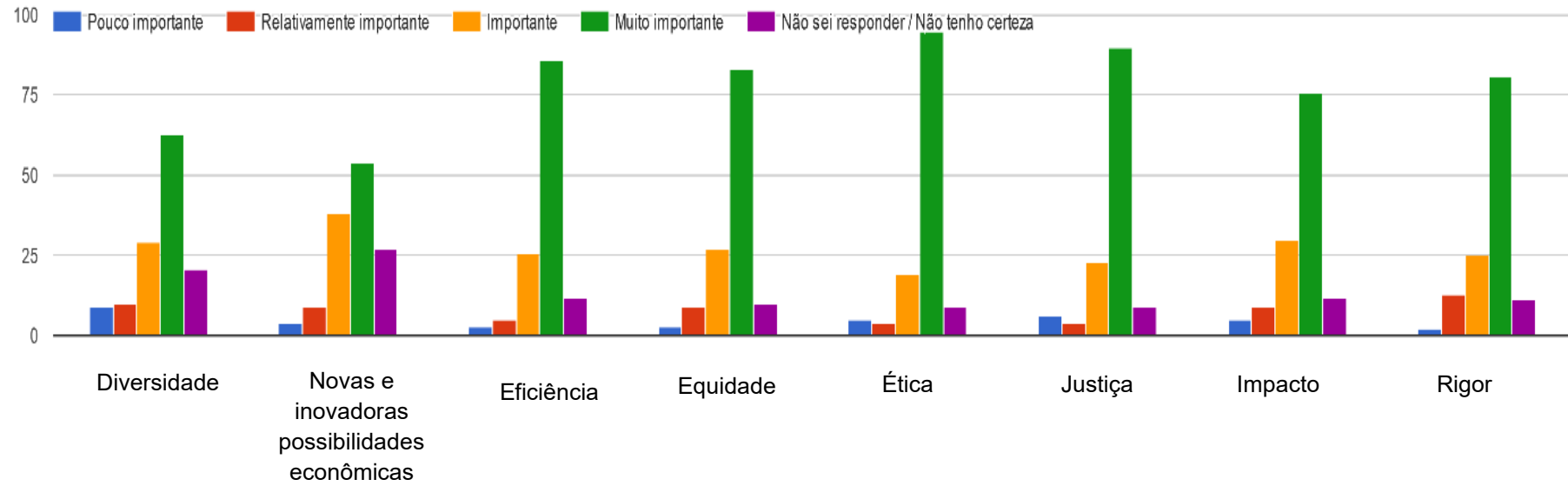
PERCEPÇÃO DOS USUÁRIOS SOBRE A CIÊNCIA ABERTA
Movimento que visa a publicização de pesquisas e dados de pesquisas em rede de livre acesso. Uma forma de compartilhar conhecimento e tornar a ciência mais acessível. Uma ferramenta de acessibilidade informacional.
Não se trata de algo específico, como um programa do governo. Pelo que entendo, é algo relacionado à abertura ao acesso de dados de pesquisas anteriores.
O compartilhamento de resultados de pesquisas sobre determinado assunto ou fenômeno relacionado a alguma área do conhecimento.
Ocorre colaboração de diversos pesquisadores e/ou instituições.
Paradigma baseado em disponibilização de dados provenientes de pesquisas científicas.
Pesquisas que utilizam ferramentas livres (não pagas) e que têm seus resultados compartilhados gratuitamente.
Primeira vez que vejo o termo. Suponho que seja uma forma de compartilhar ideias estudos e pesquisas científicas em colaboração com o fim de melhorar e manter a circulação de informações e descobertas científicas.
Publicação e livre acesso a material publicado
Reprodutibilidade
Sem proprietário, mantendo a autoria de fontes
Sem restrição de patente é voltada para solução de uma questão coletiva e aberta ao acesso de todo público
Transparência
Uma ciência cujos produtos e processos sejam conhecidos publicamente, sem barreiras de acesso às informações.

Fonte: Dados da Pesquisa (2019).

Em relação aos valores defendidos pela Ciência Aberta, foi questionado o quanto os pesquisadores consideram importante adotar os princípios da diversidade, novas e inovadoras possibilidades econômicas (*crowdfunding*, economia criativa etc.), eficiência, equidade, justiça, ética, impacto e rigor. De acordo com o gráfico 14, grande parte dos respondentes foram favoráveis a esses princípios.

Gráfico 14 - Valores e princípios da Ciência Aberta

Na sua opinião, o quanto você considera importante que os pesquisadores adotem os princípios da ciência aberta?

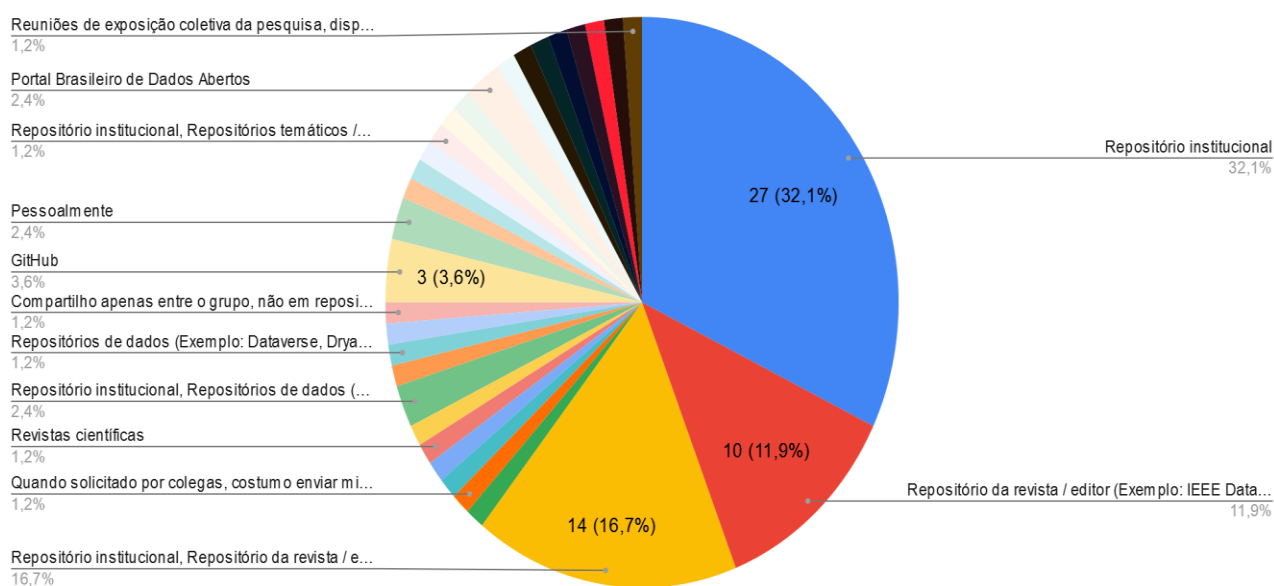


Fonte: Dados da Pesquisa (2019).

Há uma pergunta para investigar se os pesquisadores costumam acessar outros repositórios ou sites para obter dados de pesquisa de outros pesquisadores ou grupos de pesquisa. Boa parte (75,8%) respondeu que sim e 24,2% afirmou que não acessam.

Sobre o compartilhamento de dados, uma das perguntas foi a seguinte: Você compartilha seus dados de pesquisa? Cerca de 25,8% responderam que nunca compartilharam os dados, 19,7% compartilha somente entre os pares e alunos e 18,9% compartilha livremente e em acesso aberto, entre outras formas de compartilhamento. Ainda sobre esse aspecto, foi indagado àqueles que afirmaram ser compartilhadores se eles poderiam indicar quais as plataformas e repositórios que costumam utilizar. A maioria dos pesquisadores que declararam compartilhar seus dados fazem isso em repositórios institucionais, enquanto outros compartilham em repositórios de dados oferecidos pelas publicações periódicas e afins. Silva (2019) também identificou essas formas de compartilhamento em seus estudos sobre a gestão de dados científicos.

Gráfico 15 - Formas de compartilhamento de dados



Fonte: Dados da Pesquisa (2019).

Os motivos para não compartilhar os dados são variados, mas, a maioria dos respondentes declarou que o principal motivo é por desconhecimento e por não saber como fazer isso (56,9%). As demais razões estão relacionadas a dúvidas de como os



dados serão usados; outros são porque se tratam de dados confidenciais; por serem objeto de patente (1,5%); porque os dados estão em formato proprietário (6,2%); por não saberem quais os tipos de licenças aplicar aos seus datasets; receio de perder o ineditismo da pesquisa e deixar de publicar; medo de ser plagiado; por não ser uma exigência institucional ou do financiador; falta de tempo, infraestrutura e de financiamento; e falta de padronização. Sobre esse fato, há enorme verossimilhança na exposição de motivos relatados pelos pesquisadores da UFC com o que consta em estudos sobre o compartilhamento de dados e em Ciência Aberta. (GIARETTA, 2010; NELSON, 2009, SALES; SAYÃO, 2018; SCHELIGA; FRIESIKE, 2014; SILVA, 2019; TENOPIR *et al.*, 2011), porém, chama a atenção que boa parte dos respondentes justificaram que não compartilham porque desconhecem e não sabem como fazer isso. Portanto, vê-se que a literatura científica evidencia que “São muitas as razões para que a opacidade se instale nesse segmento da ciência e criem obstáculos para que os dados sejam razoavelmente integrados a outros ativos, compartilhados e reusados.” (SALES; SAYÃO, 2018, p. 4184).

A respeito do compartilhamento, outro bloco de perguntas de múltipla escolha com diversas afirmações foram elencadas no questionário para que os respondentes pudessem marcar as alternativas de acordo com a escala (Concordo, Discordo, Não sei responder), conforme a tabela 4 a seguir. De maneira geral, os pesquisadores da UFC parecem concordar com algumas proposições a respeito do compartilhamento, mas, de acordo com Nelson (2009), Sales e Sayão (2018) o fato de os pesquisadores concordarem com o compartilhamento e acesso aos dados proposto pela Ciência Aberta não significa que eles compartilhem, e quando o fazem, trata-se de uma exceção ao invés de ser uma regra, especialmente em algumas áreas.

Tabela 4 - Sobre o compartilhamento

(continua)

COMPARTILHAMENTO			Não souberam responder
	Concordam	Discordam	
Os dados devem ser compartilhados livremente em repositórios e plataformas especializadas	69,70%	15,20%	15,20%
Os dados devem ser compartilhados entre pares e comunidade acadêmica	82,60%	9,10%	8,30%
Compartilhamento apenas se for obrigatório	14,40%	75,80%	9,80%

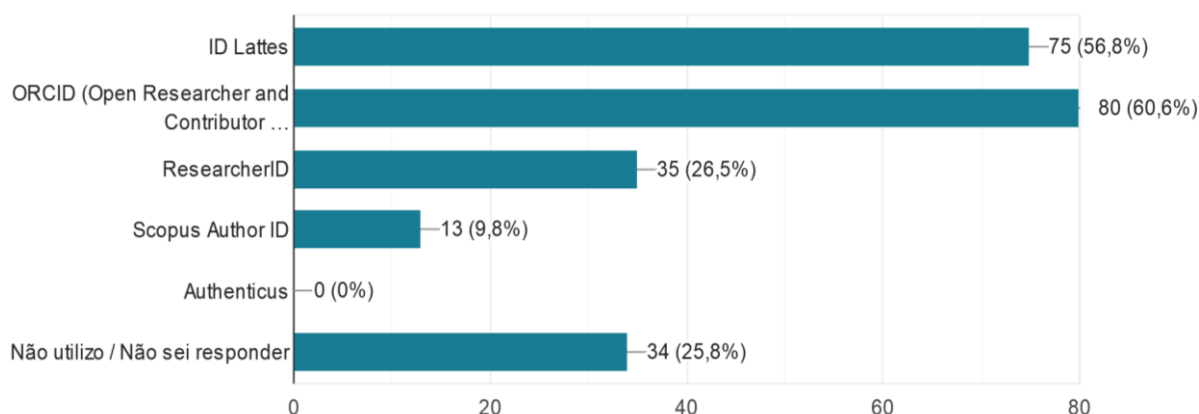
Tabela 4 - Sobre o compartilhamento

COMPARTILHAMENTO	(conclusão)		
	Concordam	Discordam	Não souberam responder
Compartilhamento é uma prática comum em minha área/disciplina	36,40%	40,20%	23,50%
Disponibilização dos dados para verificação e reprodutibilidade	91,70%	3,00%	5,30%
Disposição do pesquisador em se dedicar aos processos de organização e disponibilização de seus próprios dados de pesquisa	65,90%	15,90%	18,20%
Receio em compartilhar por mau uso ou má interpretação dos dados	34,10%	50,00%	15,90%
Receio em compartilhar por medo de cópia dos dados e de ideias	22,70%	58,30%	18,90%
Receio de perder vantagens em processos de inovação e de patentes	32,60%	46,20%	21,20%
Receio de perder oportunidade e vantagens em publicar sobre o assunto futuramente	40,20%	43,20%	16,70%
Compartilhar contribui para a divulgação, melhora o impacto e reconhecimento da pesquisa	78,00%	10,60%	11,40%
Compartilhar conjuntos de dados rende citações	75,80%	7,60%	16,70%
Uso e reuso dos dados é apropriado, inclusive para criar novos conjuntos de dados	81,10%	6,80%	12,10%
Compartilhamento de dados desde que estejam sob uma licença e termos apropriados para o acesso e a reutilização	78,80%	9,80%	11,40%
Compartilhamento de dados deve ser exigência institucional e governamental	61,40%	18,90%	19,70%

Fonte: Dados da Pesquisa (2019).

O uso de identificadores digitais para autores é importante para auxiliar nos processos de desambiguação e na atribuição correta das produções científicas, incluindo nesse rol os dados de pesquisa. Recomenda-se o uso de identificadores digitais para os datasets (Digital Object Identifier – DOI) e Orcid, por exemplo, para os autores. Deste modo, uma das perguntas do questionário buscou conhecer quais os identificadores utilizados pelos pesquisadores na UFC.

Gráfico 16 - Identificadores digitais para autores usados pelos pesquisadores



Fonte: Dados da Pesquisa (2019).

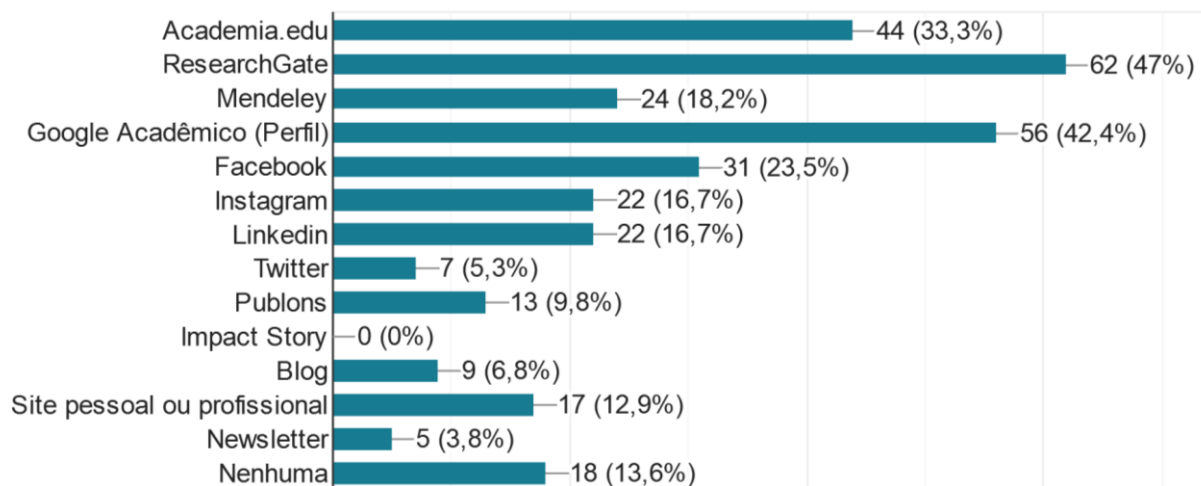
Os identificadores digitais para autores mais utilizados pelos pesquisadores são o OrcidRCID (60,6%), seguidos do ID Lattes (56,8%), ResearcherID (26,5%) e Scopus Author ID (9,8%). Não souberam responder ou não utilizam os identificadores para autor 25,8% dos respondentes.

Muitos dados de pesquisa não são compartilhados apenas em repositórios institucionais, repositórios de dados e plataformas de infraestrutura de dados, mas também em redes sociais acadêmicas, a exemplo do ResearchGate, ou em sites dos pesquisadores. Dessa forma, uma das perguntas foi direcionada para sondar quais são os canais informais de divulgação da produção científica dos pesquisadores, tais como as mídias sociais, redes sociais acadêmicas, blogs ou outros recursos usados para ampliar o impacto de suas pesquisas.

A rede social acadêmica ResearchGate é utilizada por 47% dos usuários, enquanto 33,3% também usa a Academia.edu. O módulo rede social do gerenciador de referências Mendeley é usado por 18,2% das pessoas. O Publons, uma rede social para editores e pareceristas, tem 9,8% de usuários entre o grupo investigado. Outra forma de divulgar as produções e perfis dos pesquisadores é o Google Acadêmico, adotado por 42,4% deles. As redes sociais Facebook (23,5%), Instagram (16,7%), LinkedIn (16,7%) e Twitter (5,3%) também são utilizadas. A divulgação também ocorre em blogs (6,8%), sites (12,9%) e por meio do envio de Newsletter (3,8%). Nenhum dos pesquisadores que responderam a pesquisa declararam ter perfil cadastrado no Impact Story, uma ferramenta de código-fonte aberto que fornece métricas alternativas (Altmetrics) para auxiliar os pesquisadores a mensurar o impacto e menções de suas produções disponibilizadas na Web. Com o Impact Story, podem

ser mapeadas desde uma menção a pesquisa no Twitter, como também contabiliza artigos, publicações em blogs, conjuntos de dados (datasets), softwares, todavia, alguns informaram não usar nenhum desses canais (13,6%).

Gráfico 17 - Canais informais de divulgação científica usados pelos pesquisadores



Fonte: Dados da Pesquisa (2019).

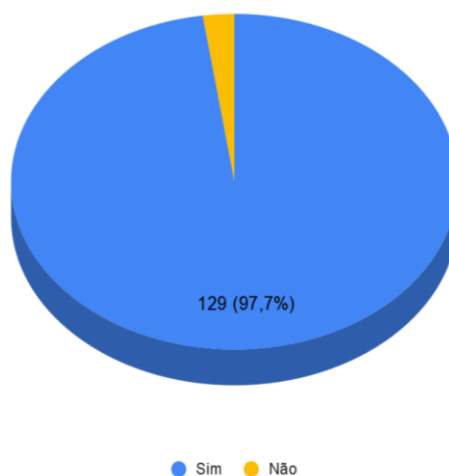
Um dos pesquisadores escreveu o seguinte comentário a respeito de uma das redes sociais acadêmicas por ele utilizada:

A utilização do ResearchGate foi um divisor de águas na minha vida acadêmica, pois consigo ter um panorama geral do que está sendo produzido na minha área e sobre o meu tema de pesquisa, posso opinar na pesquisa de colegas e receber dicas, opiniões, críticas e elogios acerca das minhas pesquisas, mais do que uma ferramenta para publicizar a referida rede é uma ferramenta de interação entre pesquisadores.

Para sondar sobre a opinião dos usuários a respeito da oferta de serviços de dados e apoio ao pesquisador, direcionamos as perguntas finais. Uma das perguntas versava sobre a responsabilidade institucional da UFC em oferecer tais serviços voltados para a coleta, organização, arquivamento, preservação e compartilhamento de dados. A expressiva maioria dos pesquisadores concordou que sim (97,7%) e apenas 2,3% consideram que não é necessário.

Gráfico 18 - Opinião dos pesquisadores sobre a responsabilidade institucional para a criação de serviços de dados

Você considera que a sua instituição deveria oferecer serviços de apoio ao pesquisador sobre coleta, organização, arquivamento, preservação e compartilhamento de dados?

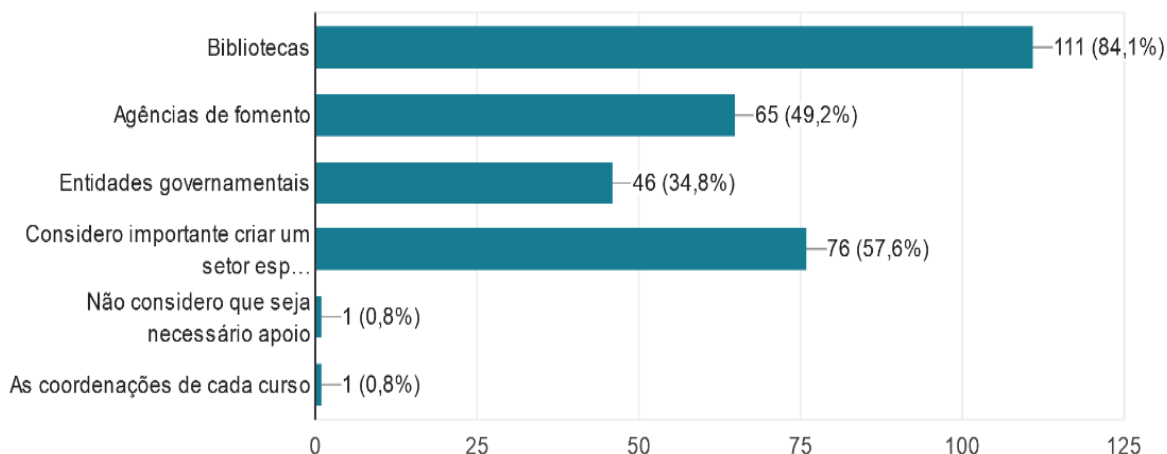


Fonte: Dados da Pesquisa (2019).

Outra questão visava saber a opinião dos pesquisadores sobre quem ou qual setor estaria mais qualificado para a prestação desses serviços. Entre as opções estavam as Bibliotecas (84,1%), as Agências de fomento (49,2%), Entidades governamentais (34,8%). Alguns responderam que seria importante considerar a criação de um setor especializado (57,6%) e um dos pesquisadores indicou as coordenações de cada curso (0,8%). Uma pessoa (0,8%) não julga necessário apoio para isso.

Em relação ao setor especializado, um dos respondentes usou o espaço dos comentários para informar que existe na universidade um setor que ele considera ter funções semelhantes a proposta de serviço de apoio ao pesquisador: “A UFC possui um prédio denominado Ambulatórios, em que há um andar totalmente dedicado ao apoio ao pesquisador, desde bibliotecária a estatístico e salas para estudo e revisão de documentos.”

Gráfico 19 - Opinião dos pesquisadores sobre quem ou qual setor está qualificado para oferecer serviços de apoio ao pesquisador

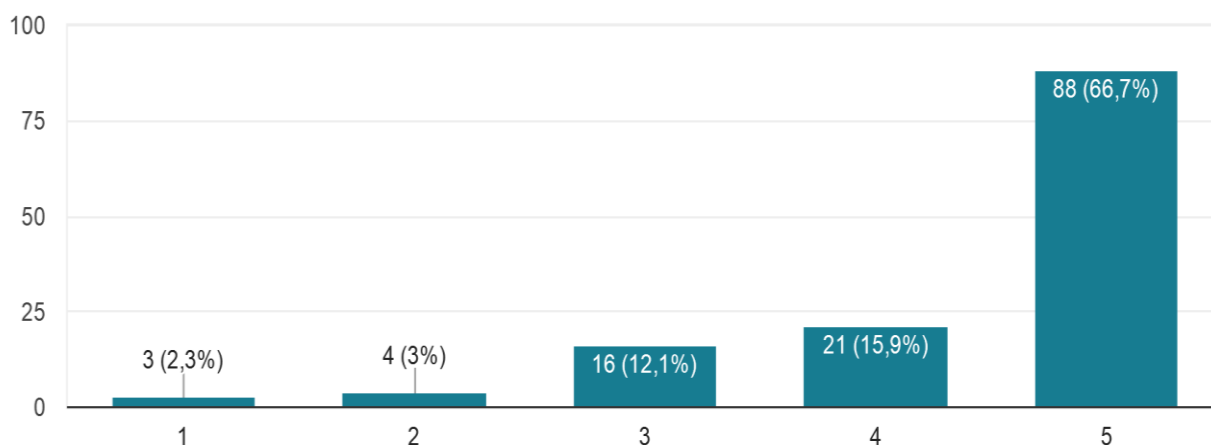


Fonte: Dados da Pesquisa (2019).

A última questão tinha o intuito de sondar o quanto os usuários consideram importante a assistência e apoio de um bibliotecário nesses serviços, tendo em vista o papel que o profissional e as bibliotecas vêm desempenhando na gestão de dados de pesquisa, a exemplo da University of Edimburgh, University of Queensland, University of Leeds e University of Bristol.

Gráfico 20 - Opinião dos pesquisadores sobre a importância do apoio bibliotecário em serviços de dados

Informe o quanto você considera importante a assistência e apoio ao pesquisador por parte do bibliotecário sobre orientações em pesquisa, ciê...mento, preservação e compartilhamento de dados: 132 respostas



Fonte: Dados da Pesquisa (2019).

Alguns respondentes usaram o espaço dos comentários para exprimir suas considerações a respeito desta pesquisa. Um dos pesquisadores comentou: “Nunca consegui acesso ao repositório institucional UFC para adicionar meus trabalhos publicados e de meus orientandos.” Um das pessoas afirmou que o tema é uma novidade: “Muito do que foi apresentado no levantamento é absolutamente novo para mim.”

Outra pessoa comentou ao final: “Fiquei confusa sobre o compartilhamento de dados, se são somente os dados ou no formato de artigos. Respondi como se fosse artigo”. Isso demonstra como os pesquisadores estão acostumados apenas a apresentar seus resultados de pesquisa em forma de artigo ou outra publicação acadêmica formal, dificultando assim o próprio entendimento para as possibilidades de utilização dos dados e do leque de opções disponíveis.

Por fim, alguns apontamentos, conforme segue:

Compartilhar dados de pesquisa é uma forma de documentar o conhecimento científico em uma resolução e com uma qualidade que antes não eram possíveis. Podemos partilhar grandes volumes de dados planilhados, fotos vídeos, e oferecer um registro mais rico para o futuro. (Comentário 1).

Fico impressionada ao descobrir o quanto ignoro sobre Ciência Aberta! Por que razão não discutimos mais esse assunto na universidade, é o que me pergunto agora. (Comentário 2).

Através do questionário foi possível ter uma noção sobre o entendimento e as práticas dos participantes em relação a gestão de dados de pesquisa e dos temas correlatos. Comprovou-se que é preciso sensibilizar, conscientizar e deixar os pesquisadores cientes sobre o tema. Por isso, são fundamentais as ações de informação promovidas pelos bibliotecários. Desse modo, as informações coletadas foram os insumos necessários que tornaram palpáveis a construção do minicurso, conforme será apresentado em seção exclusiva para essa finalidade mais adiante. Antes disso, passaremos para a apresentação dos resultados da pesquisa de campo e das entrevistas.

#### **4.3 Aspectos da pesquisa de campo, entrevistas e aprofundamento das investigações**

A segunda etapa da investigação foi composta pela pesquisa de campo e contou com a realização de visitas nos *campi* e a realização das entrevistas com os pesquisadores da UFC. Dificuldades ocorreram para convencer participantes da área

de Ciências Exatas e da Terra. Apesar dos entraves, um técnico-administrativo em educação da UFC formado na área de Computação, mestre e professor em outra instituição de ensino superior da rede privada, se prontificou a participar desta pesquisa como entrevistado.

Encontrar participantes vinculados às Ciências Biológicas também se constituiu em uma adversidade, pois, ao longo da procura por membros dessa área, a pesquisadora havia conseguido duas pessoas, mas, depois estas pessoas desistiram. Tentou-se contato com um professor do Campus Porangabuçu (cursos da área da saúde com profissionais que possuem formação às áreas ligadas às Ciências Biológicas, tais como morfologia, genética etc.) e com outro no Campus do Pici (curso de Ciências Biológicas), porém, sem sucesso.

Nesse ínterim, durante a divulgação do minicurso online, um professor do Campus do Labomar entrou em contato por e-mail elogiando a iniciativa, solicitou o acesso ao material do curso e à ementa, pois o mesmo tem pretensão de propor e ministrar uma disciplina sobre gestão de dados para seus alunos. A partir disso, a pesquisadora lançou o convite para entrevistá-lo como representante da área, mas, o referido docente declinou o convite por estar atarefado no referido período de aplicação da pesquisa e indicou os contatos de outros colegas e alunos. O docente não foi entrevistado, mas, se inscreveu e participou ativamente do curso online.

Após o início do curso, a pesquisadora identificou que alguns pesquisadores da área de Ciências Biológicas estavam participando do curso, então, como última tentativa, lançou uma postagem no curso convidando qualquer um dos profissionais que se dispusessem a ajudar em sua pesquisa. Após alguns dias, a busca chegou ao fim e um representante se prontificou para a entrevista.

Apesar de a área de Artes estar vinculada ao Colégio de Humanidades e fazer parte da grande área Linguística, Letras e Artes na tabela de área do conhecimento da Capes, a pesquisadora tentou inúmeras vezes contato com os professores e alunos dos cursos ligados às Artes, todavia, apenas uma professora do curso de Dança se colocou à disposição, porém, a data em que ela estaria disponível ultrapassaria muito o período da coleta de dados, o que infelizmente inviabilizou a sua participação.

Houve dificuldades em se obter adesão dos pesquisadores dos cursos ligados às Artes para participação na pesquisa, embora já tivesse esquematizado uma estratégia de coleta de dados diferenciada para a referida área, e se esperava contemplar de maneira mais fidedigna e completa possível os dados produzidos no



campo das Artes a fim de expor os achados sobre os dados de pesquisa e as particularidades das Artes e Artefatos produzidos, inclusive os digitais, e assim relacioná-los com as Humanidades digitais e com o conceito de *OpenGLAM* ou simplesmente *GLAM*, acrônimo para “*Galleries, Libraries, Archives and Museums*” (Galerias, Bibliotecas, Arquivos e Museus), iniciativa que tem como proposta oportunizar o acesso aberto ao patrimônio cultural, promover o intercâmbio e a colaboração entre instituições culturais que apoiam o acesso aberto ao conhecimento também para esse seguimento. (OPEN GALLERIES, LIBRARIES, ARCHIVES AND MUSEUMS – OpenGLAM, 2019).

Lamentavelmente, essa meta não foi alcançada tendo em vista a indisponibilidade de sujeitos do campo das Artes. Apesar do entrave, considera-se que não é exclusividade desse estudo a dificuldade em conseguir adentrar o universo das Humanidades digitais, uma vez que há pesquisas como a de Knöchelmann (2020) que relatam o quanto é difícil adentrar nessa área, pois, muitas vezes não há sequer acesso às produções, publicações e dados abertos em humanidades digitais, entre outras questões. Também a universidade holandesa TUDELFT enfrentou dificuldades para conseguir pesquisadores voluntários nessa área para colaborar nos serviços de gestão e curadoria de dados

Outro ponto a ser destacado é que a partir do contato com um dos sujeitos da área de Psicologia durante a fase da coleta de dados e pesquisa de campo, a pesquisadora foi apresentada à Professora Zulmira Bomfim e para a equipe responsável por um dos laboratórios do curso sobre estudos em Psicologia Ambiental. A Psicologia Ambiental tem como tema central estudar o ser humano em seu contexto, ou seja, as interrelações entre a pessoa e o meio ambiente físico e social (MOSER, 1998). Esse campo de estudo visa analisar como as condições ambientais afetam as capacidades cognitivas, mobilizam comportamentos sociais que impactam na saúde mental dos indivíduos e contribuem para a análise das percepções e interpretações das pessoas sobre o meio ambiente. (GÜNTHER; ROZESTRATEN, 2004).

Nesse contexto, o Laboratório de Pesquisa em Psicologia Ambiental (LOCUS) possui uma série de documentos e arquivos impressos e eletrônicos sobre dados oriundos de pesquisas realizadas anteriormente pelo grupo de estudo, tais como os Mapas Afetivos, mantidos dentro dos armários e computador da professora responsável, acontecimento este que revela mais uma vez como a demanda existe na prática, porém, ainda latente, e conforme relatos presentes na literatura científica

sobre a situação desses documentos guardados em salas, gabinetes e laboratórios, trancados em gavetas e armários sem o devido tratamento e tão importantes para o conhecimento científico (FERGUSON *et al.*, 2014; SAYÃO; SALES, 2015, 2019), ou seja, como mencionado por Sayão e Sales (2019), a cauda longa da ciência.

Explicando brevemente o que são os Mapas Afetivos, também denominado Instrumento Gerador dos Mapas Afetivos (IGMA), conforme Augusto, Feitosa e Bomfim (2016), “[...] foi construído por Bomfim (2010) em sua tese de doutorado, tendo como finalidade ser um método de investigação dos afetos em relação ao ambiente.” Portanto, os Mapas Afetivos são conceituados pela autora como “[...] um instrumento que facilita torná-los tangíveis, por meio de imagens, palavras e da formulação de sínteses ligadas aos sentimentos, [...] de forma menos elaborada e [...] mais sensível.” (BOMFIM, 2010, p. 137). Em outras palavras:

A composição original do instrumento, direcionada a abordar os afetos dos sujeitos em relação à cidade, abrange o Desenho elaborado pelo respondente; o Significado do desenho atribuído também pelo participante; os Sentimentos; as Palavras-síntese; O que pensa da cidade; Categorias da Escala de tipo Likert; Comparação da cidade; Caminhos percorridos; Participação em associação; Participação eventual em movimentos sociais; e as Características sociodemográficas. (AUGUSTO; FEITOSA; BOMFIM, 2016, p. 149).

Tão logo este ambiente foi identificado durante o andamento da pesquisa de campo, a pesquisadora e sua chefia imediata da biblioteca em que trabalha estabeleceram contato com a professora coordenadora do LOCUS e prestaram alguns treinamentos solicitados. Inclusive, construíram planos de parcerias futuras que englobam não somente o pequeno acervo bibliográfico do referido laboratório, mas, também o tratamento e curadoria dos mapas afetivos do LOCUS, dando toda a assessoria necessária. Contudo, devido a certos entraves institucionais em torno de permissões e recursos para prosseguir com o planejamento idealizado, ainda não houve avanço significativo dessas ações.

Aconteceram reuniões de orientações gerais entre a professora, o grupo de alunos ligados ao LOCUS, a pesquisadora e seu chefe imediato (ambos bibliotecários) sobre como iniciar a organização dos Mapas Afetivos, sua digitalização e descrição para agilizar o tratamento e armazenamento no repositório da instituição e outro de dados futuramente. O impasse daquele momento, até então, era conseguir permissão para inserir os mapas afetivos no Repositório Institucional da UFC, contudo, em acesso restrito aos membros do LOCUS, pois, tratam-se de documentos que possuem

dados confidenciais que não podem ser expostos ou revelados ao público, mas, que merecem o devido tratamento e preservação, inclusive para estudos futuros e novas produções do LOCUS ou parcerias de pesquisa com outras instituições nacionais e internacionais que venham a se formar, exatamente por ser uma produção importante para a área de Psicologia e para a universidade, além disso, os Mapas Afetivos são uma criação da Professora Zulmira Bomfim, pesquisadora pertencente a UFC.

Não foi possível durante a pesquisa de campo iniciar os trabalhos de curadoria e gestão de dados dos Mapas Afetivos do LOCUS, afinal, apesar da abertura e receptividade por parte da professora e seu grupo, diversas questões vieram à tona, principalmente no tocante ao acesso aos Mapas Afetivos e de ordem ética da Psicologia. Uma das dúvidas e que não se chegou a um denominador comum é se os bibliotecários poderiam ou não ter acesso ao material, afinal, este é um instrumento utilizado apenas pelos psicólogos, assim como acontece com os testes psicológicos e que é acessível apenas os profissionais da área. Por ser uma espécie de material que abriga dados considerados sensíveis, foi discutida a possibilidade de esses mapas afetivos serem digitalizados e armazenados, porém, não poderiam estar disponíveis em rede e abertos ao acesso público, enquanto as cópias em papel desses mapas afetivos deveriam receber o tratamento documental adequado e serem organizados.

Outro problema encontrado está intimamente relacionado à falta de infraestrutura adequada e a inexistência de políticas ou diretrizes para o armazenamento desse tipo de material no Repositório Institucional (RI) da UFC, haja vista que nem autoarquivamento o referido RI possui. Evidentemente, isso se deve ao fato de que a gestão de dados de pesquisa é algo ainda inexistente, inédito e bastante atual na universidade, constituindo-se na prova cabal de que é premente a criação desses serviços por parte da biblioteca, conforme uma das propostas desta pesquisa.

A UFC tem inúmeros ambientes que abrigam uma infinita diversidade de dados e materiais produzidos, entretanto, neste estudo, tivemos acesso e abertura apenas do LOCUS. A título de informação, estão discriminados no apêndice G alguns dos ambientes identificados. É preciso ressaltar que mesmo tendo sido realizado um levantamento prévio sobre os ambientes produtores de dados de pesquisa, a referida lista pode não ter contemplado outros locais que porventura não tenham sido mapeados. Ademais, a própria universidade como um todo é produtora nata de conhecimento, portanto, produz dados de pesquisa.

Após relatar sobre a pesquisa de campo, o foco volta-se para as entrevistas. Atribuiu-se um código para cada um dos entrevistados, assim expresso: Entrevistado 1 = E1, e assim sucessivamente até completar os 12 pesquisadores entrevistados (7 docentes, 4 discentes e 1 técnico administrativo). A numeração dos códigos adotada foi sorteada e atribuída aleatoriamente, portanto, não corresponde necessariamente à ordem de ocorrência das entrevistas.

Quadro 14 - Distribuição dos entrevistados de acordo com a área de conhecimento

(continua)

<b>COLÉGIO DE CIÊNCIAS DA VIDA</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
E10	Ciências Biológicas	Oceanografia / Ciências Marinhas Tropicais	Discente (Doutorado)	1
E6	Ciências da Saúde	Odontologia	Docente	1
E9	Ciência Agrárias	Agronomia / Ciência do Solo	Docente	1
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>3</b>
<b>COLÉGIO DE HUMANIDADES</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
E5	Ciências Sociais Aplicadas	Administração	Docente	1
E1	Ciências Sociais Aplicadas	Finanças	Docente	1
E3	Ciências Humanas	Geografia <sup>150</sup>	Discente	1
E8		Psicologia	Docente	1
E4	Linguística, Letras e Artes	Letras Inglês	Docente	1
E2		Letras Espanhol	Discente (Mestrado)	1
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>6</b>

<sup>150</sup> A Geografia é uma ciência ampla e pode estar presente nas Geociências (Geografia Física) e ainda estar contida nas Ciências Humanas, conforme a Tabela de Área do Conhecimento da Capes.

Quadro 14 - Distribuição dos entrevistados de acordo com a área de conhecimento  
(conclusão)

<b>COLÉGIO DE CIÊNCIAS EXATAS, TECNOLÓGICAS E MULTIDISCIPLINAR</b>				
<b>CÓDIGO</b>	<b>GRANDE ÁREA</b>	<b>SUBÁREA</b>	<b>CATEGORIA</b>	<b>QUANTIDADE</b>
E12	Ciências Exatas e da Terra	Ciência da Computação	Técnico-Administrativo em Educação (Mestrado)	1
E11	Engenharias	Engenharia Civil / Engenharia Sanitária / Recursos Hídricos	Discente (Mestrado)	1
E7		Engenharia de Produção	Docente	1
<b>TOTAL DE PARTICIPANTES/COLÉGIO</b>				<b>3</b>
<b>TOTAL GERAL DE PARTICIPANTES</b>				<b>12</b>

Fonte: Dados da pesquisa (2019).

A partir do quadro 14 de distribuição dos entrevistados, como representantes do Colégio de Ciências da Vida, temos os entrevistados E6, E9 e E10. Representando o Colégio de Humanidades, os sujeitos E1, E2, E3, E4, E5 e E8. No Colégio de Ciências Exatas, Tecnológicas e Multidisciplinar, temos os pesquisadores E7, E11 e E12. Dos 12 participantes, sete são docentes, quatro são discentes e um é servidor Técnico-Administrativo em Educação da UFC que possui Mestrado. As entrevistas com os indivíduos ocorreram nos meses de agosto a outubro de 2019.

De acordo com o que já foi mencionado na seção sobre a metodologia, após fazer o primeiro levantamento de informações com o uso do questionário e extrair os dados para a análise inicial, prosseguiu-se com a entrevista. Isso foi fundamental para aprofundar o entendimento do fenômeno estudado e para fazer o cotejamento dos dados coletados no questionário com os das entrevistas realizadas.

Assim como no questionário, a maioria dos entrevistados pertence ao gênero masculino (quatro docentes, três discentes e um técnico-administrativo), e apenas quatro mulheres participaram das entrevistas (três docentes e um discente).

Perguntados sobre o tipo de abordagem que costumam desenvolver em suas pesquisas, adotam a abordagem quantitativa os sujeitos E1 (Ciências Sociais Aplicadas), E6 (Ciências da Vida), E7 (Engenharias) e E9 (Ciências Agrárias), enquanto os sujeitos E2 (Linguística, Letras e Artes), E3 (Ciências Humanas) e E5 (Ciências Sociais Aplicadas) realizam pesquisas de natureza predominantemente qualitativa. Utilizam a abordagem mista (quantitativa e qualitativa) os indivíduos E4 (Linguística, Letras e Artes), E8 (Ciências Humanas), E10 (Ciências Biológicas), E11 (Engenharias) e E12 (Ciências Exatas, Tecnológicas e Multidisciplinar).

#### 4.3.1 Dados

Considerando os tipos de arquivos produzidos e as estratégias de organização, armazenamento, manutenção e preservação dos dados de pesquisa, investigou-se junto aos entrevistados acerca do assunto. Evidentemente, os tipos de materiais produzidos correspondem aos tipos de abordagem utilizadas nas pesquisas declaradas pelos entrevistados, e, as estratégias de organização e backup dos dados dos indivíduos revela que a maioria deles não possui uma frequência específica de backup dos dados, pois costumam usar a sincronização e o salvamento automático na nuvem, em serviços como o Google Drive, Dropbox e OneDrive, ou, preferem salvar os arquivos de acordo com a conveniência e manipulação dos dados (Quadro 15).

Quadro 15 - Estratégias de organização e armazenamento dos dados

(continua)

<b>COMO VOCÊ ORGANIZA, ARMAZENA, MANTÉM E PRESERVA OS SEUS DADOS PARA USO FUTURAMENTE?</b>			
<b>ENTREVISTADO</b>	<b>ÁREA</b>	<b>TIPOS DE DADOS</b>	<b>ESTRATÉGIAS UTILIZADAS</b>
E1	Ciências Sociais Aplicadas (Finanças)	Planilhas, números, Scripts	- Armazenamento: Computador de trabalho, HD externo. - Organização: No computador, pastas separadas por ano e artigo ou livro + uma pasta extra conectada ao gerenciador de referências...

Quadro 15 - Estratégias de organização e armazenamento dos dados

(continuação)

COMO VOCÊ ORGANIZA, ARMAZENA, MANTÉM E PRESERVA OS SEUS DADOS PARA USO FUTURAMENTE?			
ENTREVISTADO	ÁREA	TIPOS DE DADOS	ESTRATÉGIAS UTILIZADAS
			(Mendeley) com os materiais usados para compor o referencial teórico das publicações. O e-mail é utilizado para acompanhamento das mensagens trocadas e como um registro histórico. - Frequência de backup: Semestral.
E2	Letras, Linguística e Artes (Letras Espanhol)	<i>Corpora</i> textuais, textos	- Armazenamento: E-mail, Nuvem e Pendrive. - Organização: Não há uma organização fixa. Utiliza o e-mail para salvar os dados de pesquisa e manter um histórico das atualizações e manipulações constantes dos arquivos. - Frequência de backup: Não há uma frequência específica, é de acordo com a conveniência e manipulação dos dados e constantes envios por e-mail e uso do salvamento automático da nuvem.
E3	Ciências Humanas (Geografia)	Textos, Quadros teóricos, Poligonais, Coordenadas, Mapas georreferenciais	- Armazenamento: Nuvem, Pendrive e HD externo. - Organização: Organiza os arquivos em pastas, realiza clipping de matérias jornalísticas. - Frequência de backup: Não há uma frequência específica, é de acordo com a conveniência e manipulação dos dados e constantes envios por e-mail e uso do salvamento automático da nuvem.
E4	Letras, Linguística e Artes (Letras Inglês)	<i>Corpora</i> textuais, arquivos TXT, Textos	- Armazenamento: OneDrive, Google One, Dropbox e HD externo. - Organização: Organiza pastas por assunto. - Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.

Quadro 15 - Estratégias de organização e armazenamento dos dados

(continuação)

<b>COMO VOCÊ ORGANIZA, ARMAZENA, MANTÉM E PRESERVA OS SEUS DADOS PARA USO FUTURAMENTE?</b>			
<b>ENTREVISTADO</b>	<b>ÁREA</b>	<b>TIPOS DE DADOS</b>	<b>ESTRATÉGIAS UTILIZADAS</b>
E5	Ciências Sociais Aplicadas (Administração)	Transcrições e Textos	<ul style="list-style-type: none"> <li>- Armazenamento: Computador pessoal, Pendrive e Nuvem.</li> <li>- Organização: Não especificada.</li> <li>- Frequência de backup: Quinzenal e Mensal. Também utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>
E6	Ciências da Saúde (Odontologia)	Planilhas, Números, arquivos SPSS, Cópias físicas impressas dos Termos de Consentimento Livre e Esclarecido (TCLE) aplicados em pesquisas	<ul style="list-style-type: none"> <li>- Armazenamento: Nuvem e Pendrive.</li> <li>- Organização: Organiza os arquivos em pastas. Os TCLE são guardados e organizados em caixas-arquivo, mantidos pelo período de cinco anos ou mais.</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>
E7	Engenharias (Engenharia de Produção)	Dados de entrada e saída, Planilhas, Números, arquivos TXT com matrizes de dados, Códigos e linguagem de programação (Julia e Python), Scripts	<ul style="list-style-type: none"> <li>- Armazenamento: Nuvem, Dropbox, Google Drive, ResearchGate.</li> <li>- Organização: Não há uma organização específica, é mais voltada para o tipo de pesquisa envidado (pesquisa operacional).</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>



Quadro 15 - Estratégias de organização e armazenamento dos dados

(continuação)

<b>COMO VOCÊ ORGANIZA, ARMAZENA, MANTÉM E PRESERVA OS SEUS DADOS PARA USO FUTURAMENTE?</b>			
<b>ENTREVISTADO</b>	<b>ÁREA</b>	<b>TIPOS DE DADOS</b>	<b>ESTRATÉGIAS UTILIZADAS</b>
E8	Ciências Humanas (Psicologia)	Transcrições, Textos, Bancos de dados, arquivos SPSS, Números	<ul style="list-style-type: none"> <li>- Armazenamento: HD externo, Nuvem, Google Drive e E-mail.</li> <li>- Organização: Não especificada.</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>
E9	Ciências Agrárias (Ciência do Solo)	Planilhas, Números, Planilhas físicas em papel, Diário de campo, Dados observacionais	<ul style="list-style-type: none"> <li>- Armazenamento: Nuvem, OneDrive.</li> <li>- Organização: Não especificada.</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência. Anteriormente, usava o HD externo para o backup semanal.</li> </ul>
E10	Ciências Biológicas (Ciências Marinhas e Tropicais)	Dados de medição, Dados observacionais, Diário de campo, Registros fotográficos, Planilhas, Números e Textos.	<ul style="list-style-type: none"> <li>- Armazenamento: Pendrive, Nuvem, E-mail e HD externo.</li> <li>- Organização: Arquivos organizados a partir da data de modificação.</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>
E11	Engenharias (Engenharia Civil / Sanitária / Recursos Hídricos)	Planilhas, Números e Textos. Já usou caderno de laboratório.	<ul style="list-style-type: none"> <li>- Armazenamento: Computador e Nuvem.</li> <li>- Organização: Pastas (Computador e Nuvem).</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento na nuvem de acordo com a conveniência.</li> </ul>

Quadro 15 - Estratégias de organização e armazenamento dos dados

(conclusão)

COMO VOCÊ ORGANIZA, ARMAZENA, MANTÉM E PRESERVA OS SEUS DADOS PARA USO FUTURAMENTE?			
ENTREVISTADO	ÁREA	TIPOS DE DADOS	ESTRATÉGIAS UTILIZADAS
E12	Ciências Exatas e da Terra (Ciência da Computação)	Textos e Slides	<ul style="list-style-type: none"> <li>- Armazenamento: HD externo e Nuvem. Dependendo da natureza e sensibilidade dos dados, costuma criptografá-los.</li> <li>- Organização: Não especificada.</li> <li>- Frequência de backup: Não há uma frequência específica, utiliza a sincronização e salvamento automático da nuvem ou de acordo com a conveniência.</li> </ul>

Fonte: Dados da Pesquisa (2019).

Conforme mencionado anteriormente e exposto no quadro acima, poucos são os pesquisadores que mantêm uma rotina de backup de seus arquivos, enquanto a maioria não estabelece um período porque utiliza o salvamento automático e a sincronização na nuvem, ou ainda de acordo com a conveniência. De acordo com Sayão e Sales (2015), o ideal é que o backup seja feito a cada nova alteração, além de ser recomendado que se tenham cópias dos backups em outros dispositivos e mídias de diferentes formatos. O backup automático é uma interessante alternativa, mas, vale lembrar que apenas ele não é suficiente para garantir a existência dos dados.

A pesquisa revela ainda que alguns dos entrevistados confiam mais no e-mail para salvar os seus arquivos porque consideram confiável e por manter um registro histórico das movimentações e alterações realizadas nos arquivos de acordo com as datas mais antigas e as recentes, como podemos notar no depoimento da Entrevistada 2 (2019, grifo nosso):

Pra uso futuro é complicado porque até pra uso presente eu não sou uma pessoa que se diga tão organizada, né? Eu tenho eles [os dados] salvos na nuvem, **salvos no e-mail, que eu ainda confio mais do que na nuvem, porque no e-mail ele fica como um arquivo mesmo, um registro já mais**

**duradouro, aparentemente.** [...] tenho em pendrives, não gravo mais CD, né, até porque os notebooks hoje em dia nem tem mais o espaço, e, **recentemente através dos repositórios, né, que você me apresentou, mas, a preferência sempre são a nuvem e os e-mails. Pela necessidade da pesquisa, de uma atualização, de uma manipulação constante, eu acabo salvando automaticamente no e-mail as cópias que são geradas, as edições que são feitas, e, na nuvem, ela salva automaticamente, né? Então, é o máximo que eu faço de segurança, mas, com relação à backups,** a atualizar, por exemplo, pendrives, eu não atualizo porque até você achar o pendrive que, às vezes tem muitos, perde muito fácil, e aí ficam várias versões, mas, o meu cuidado maior são com as salvaguardas online. [...] Eu tive dificuldade de encontrar o arquivo mais recente porque eu salvo as edições e aí até achar a mais atual... Às vezes, eu posso encontrar uma dificuldade, mas, geralmente, não, perca [dos dados], não.

O relato evidencia ainda que a pesquisadora não faz uso de nenhuma estratégia de organização de seus dados, e destaca a questão da obsolescência das mídias quando justifica o motivo pelo qual não salva mais arquivos em CD. Até mesmo o pendrive é utilizado sem uma sistemática de organização, afirma que pode ser perdido facilmente, o que de fato pode acontecer com qualquer um, e reforça que a sua preocupação de manutenção dos arquivos se restringe ao e-mail e à nuvem para o armazenamento dos dados que são constantemente manipulados. Sayão e Sales (2015) alertam que o envio de arquivos por e-mail não é recomendado e reforçam que é preciso que os dados sejam criptografados. Além disso, um e-mail pode ser invadido, esquecido aberto em um computador ou outros dispositivos por seu usuário e ainda ser alvo de vazamentos na internet. Exemplo de um site que verifica se um e-mail foi alvo de violações de dados da internet é o 'Have I Been Pwned?'<sup>151</sup>.

A referida entrevistada comenta ainda que a partir das indicações da dada pela bibliotecária (neste caso, a mestrandia responsável por esse estudo) em ocasião de um atendimento personalizado na biblioteca é que tomou conhecimento dos repositórios que poderiam abrigar os seus conjuntos de dados formados por *corpora* textuais. Apesar da recente descoberta dos repositórios de dados, a E2 continua preferindo a nuvem e o e-mail para disponibilizar seus dados de pesquisa. Naturalmente, isso se deve pelo fato de que é necessário incutir nos pesquisadores a cultura de gerenciar os próprios dados de pesquisa, fazer disso uma rotina, um padrão a ser seguido, assim, aos poucos o comportamento dos pesquisadores irá mudando e essa prática irá sendo cada vez mais internalizada na instituição.

<sup>151</sup> Have I Been Pwned? Disponível em: <https://haveibeenpwned.com/>.

Assim como no questionário, também foi solicitado aos pesquisadores entrevistados que falassem sobre qual seria a percepção que eles têm sobre o que são dados de pesquisa. A seguir, o quadro 16 apresenta a definição dada por cada um dos sujeitos.

Quadro 16 - Percepção sobre os Dados de Pesquisa

(continua)

O QUE VOCÊ ENTENDE POR DADOS DE PESQUISA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E1	Ciências Sociais Aplicadas (Finanças)	<p>“Pra mim, os dados da pesquisa são... [...] primeiro pensar o que é dado? <b>Dado é uma unidade que não informa muita coisa ainda, pura e simplesmente ele quantifica um fato [...] pra uma determinada observação [...]. [...]</b> por exemplo, no meu caso, se eu olhar pra uma tabela do Excel, uma célula é um dado, só isso. Quando eu olho para os meus scripts, o meu dado é a linha, que a linha de comando me diz alguma coisa. Às vezes eu faço um comentário, então, cada linha de comando é um dado. Na minha cabeça, dado é isso. <b>Se o script em si não é um dado, script pra mim é uma informação. A tabela é um dado, só que quando eu faço a estatística, eu tenho a informação.</b>”</p>
E2	Letras, Linguística e Artes (Letras Espanhol)	<p>“Dados de pesquisa <b>são as ferramentas utilizadas na pesquisa pra comprovar aquilo que você busca observar, analisar [...].</b> Os seus meios de observação da pesquisa.”</p>
E3	Ciências Humanas (Geografia)	<p>“Nas minhas palavras, eu explicaria que <b>os dados de pesquisa são os elementos que você tem pra subsidiar a sua pesquisa.</b> Podem ser dados numéricos. [...] podem ser dados quantitativos e podem ser esses dados qualitativos, que é mais ou menos como eu tava falando que eu organizei o meu, né? Por exemplo, textos, dissertações, teses, referentes aquele assunto, então, eu tento organizar todos os meus dados, que eu considero os meus dados de pesquisas ali, num arquivo só. Aí como tem, eu trabalho com essas várias abordagens, aí eu faço a separação, por exemplo, é... referente só à questão de justiça, só a questão do papel do Estado, aí eu considero que os dados são isso.”</p>

Quadro 16 - Percepção sobre os Dados de Pesquisa

(continuação)

O QUE VOCÊ ENTENDE POR DADOS DE PESQUISA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E4	Letras, Linguística e Artes (Letras Inglês)	“Eu não saberia... Eu não tenho assim uma definição muito clara. Pra mim, os <b>dados de pesquisa são o que eu gero a partir dos meus instrumentos</b> . Então, se eu tenho uma entrevista, eu vou ter uma transcrição, se eu tenho um formulário, um questionário escrito, eu vou ter o próprio questionário e a tabulação daqueles dados, se eu tenho uma análise computacional, eu tenho o resultado da análise que o software oferece num formato de arquivo que pode variar. E pra mim, dados são esses, são essas coisas. Eu não tenho nenhuma... Nunca parei pra pensar se haveria uma outra maneira de definir isso, não. É a isso que a gente se refere quando tá escrevendo e diz "os dados sugerem... os dados informam...". Quando eu falo os dados, é isso aí.”
E5	Ciências Sociais Aplicadas (Administração)	“Dados de pesquisa... Eu acho que tudo, né? Se a gente pensar em metodologia, é tudo. <b>Não só os dados que você coleta, mas, todo o processo</b> de como você fez, de marcar, por exemplo, o horário que você tá gravando. [...] Acho que tudo isso são dados da pesquisa. Tudo que é importante pra você, inclusive valorizar o seu trabalho. "Ah, eu gravei, eu transcrevi", você não sabe o tempo que você perdeu, ficou esperando o aluno sair aqui pra falar comigo, então, é assim, né? Isso tudo eu acho que é importante. Valoriza seu trabalho.”
E6	Ciências da Saúde (Odontologia)	“Os dados de pesquisa <b>são aqueles dados que eu vou coletar, seja na fonte primária, pesquisando diretamente onde o tema da pesquisa acontece, seja num lugar, seja com as pessoas</b> , certo? Seja num laboratório ou numa clínica. Então, são aqueles dados em que o pesquisador vai visualizar, medir, mensurar, anotar. <b>Seriam os nossos dados primários. E os secundários são os dados que a gente utiliza a partir de outras fontes de pesquisa</b> . Mas, normalmente a gente trabalha com dados primários.”
E7	Engenharias (Engenharia de Produção)	“Dado de pesquisa? Assim uma definição técnica ou o que eu na minha área entendo como dado de pesquisa? Na minha área, o que eu entendo como dado de pesquisa, <b>é um conjunto de dados brutos ou de informações que vão ser utilizados em alguma investigação</b> . A minha área é quantitativa, são matrizes numéricas.”

Quadro 16 - Percepção sobre os Dados de Pesquisa

(conclusão)

O QUE VOCÊ ENTENDE POR DADOS DE PESQUISA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E8	Ciências Humanas (Psicologia)	“Para mim <b>são dados que podem ser construídos de várias formas</b> [...], vamos dizer assim, mais padronizada [...]. Aplicação de um questionário, o dado que é gerado ali por essa análise, [...] <b>por exemplo, de grupos focais, de entrevistas [...]</b> <b>E também dados vindo de Diário de Campo, de observações, então, tanto dados mais estruturados, padronizados, como dados mais [...]</b> livres e que retratam a realidade, naquele momento, porque também a gente tem uma realidade muito dinâmica, e, principalmente a temática que eu trabalho que é as implicações psicossociais da pobreza. Então, ela é algo que tá mudando muito, principalmente com a mudança do governo. Então, assim, são dados também que eles estão muito relacionados a um momento, porque às vezes, depois de X anos aqueles dados já não estão mais contextualizados.”
E9	Ciências Agrárias (Ciência do Solo)	“Dados de pesquisa é <b>todo aquele resultado que eu obtenho quando tô fazendo um teste de alguma coisa</b> . Isso eu tô alinhando com uma visão minha. Se eu tô testando adubação, por exemplo, <b>todas aquelas informações que sejam qualitativas ou quantitativas, que eu observo, que eu anoto, que eu colete, a partir das minhas observações são dados de pesquisa. Mesmo que eu utilize ou não</b> . Às vezes, eu observo coisas que servem, às vezes, pro meu entendimento, não serve para outra coisa. Não preciso publicar. Mas, aquilo foi dado de pesquisa.”
E10	Ciências Biológicas (Ciências Marinhas e Tropicais)	“Eu acredito que sejam <b>todas as informações a respeito das variáveis que eu vou estudar, que eu estabeleci para estudar durante a pesquisa</b> . E essas, <b>não só as variáveis em si, mas, todo o desdobramento que pode vir delas</b> . Tipo, os gráficos que vão ser gerados dentro do processo de interpretação. As tabelas que são construídas. Então, essas informações são extraídas dos dados originais, acredito que sejam esses os dados da pesquisa.”
E11	Engenharias (Engenharia Civil / Sanitária / Recursos Hídricos)	“São os resultados, [...] só que não são só os resultados, é o que eu observo, certo? <b>Os dados de pesquisa é o que observo e resultados obtidos através de experimentos que eu fiz</b> . [...]”
E12	Ciências Exatas e da Terra (Ciência da Computação)	“ <b>São as informações coletadas e que serão processadas durante a análise</b> . Os dados serão transformados em informação ao longo da pesquisa.”

Fonte: Dados da Pesquisa (2019).

Definir dado por si só já é uma tarefa bastante complexa, especialmente diante do novo panorama que se apresenta em relação aos dados de pesquisa. Outrossim, a discussão sobre os dados foi amplamente debatida por autores como Blair (2010), Borgman (2015), Brown e Duguid (2017), Brookes (1980), Buckland (1991), Davenport (1998), Day (2001), Ingwersen e Järvelin (2005), Liu (2004), Nielsen; Hjørland, 2013, Oliveira (2002), Padoveze (2000), Russo (2010), Setzer (1999) e Zins (2007).

Diante do exposto, a percepção dos pesquisadores entrevistados apresenta semelhanças e ao mesmo tempo algumas diferenças, confirmando o quanto a variedade de dados produzida e os diferentes tipos de áreas e de abordagem dos estudos podem influenciar diretamente na forma como são vistos os dados de pesquisa. Obviamente, o entendimento geral é de que esses dados são aqueles oriundos de pesquisa científica, entretanto, destacam-se os que são vistos pelos entrevistados como unidade que não informa, mas, quantifica um fato ou uma observação; ferramentas utilizadas na pesquisa; elementos que subsidiam a investigação; dados que são construídos e retratam uma realidade; até todo o processo de estudos e metodologia são considerados como dados de pesquisa.

Por exemplo, a percepção de que os dados são ferramentas utilizadas na pesquisa não pode ser considerada incorreta do ponto de vista lógico, se considerarmos as características da área do sujeito entrevistado e as formas utilizadas para obter os dados, que são exatamente com o uso de ferramentas de extração de palavras para a montagem do *corpus* textual ou até a consulta aos repositórios e sites institucionais produtores de *corpora* textuais, que costumam disponibilizar gratuitamente esses materiais. Igualmente acontece com a definição de que o dado pode ser construído e retratar uma realidade, nada mais justo e tão verdadeiro quando se trata do trabalho desenvolvido, por exemplo, por psicólogos e cientistas sociais que investigam a pobreza ou a realidade em que se encontram os moradores de rua. Ou ainda, que dado de pesquisa é tudo relacionado à investigação, incluindo a metodologia e todo o processo que fez com que se chegasse a determinadas descobertas e resultados. Aliás, considera-se que qualquer área valoriza essencialmente o processo e a forma como a pesquisa é conduzida e feita, principalmente na realidade atual em que estão em voga a reprodutibilidade e a replicabilidade na ciência.

Em relação a estar ciente da existência de repositórios de dados de pesquisa generalistas ou de acordo com a área de conhecimento do pesquisador, boa parte dos entrevistados (dos doze sujeitos, oito conhecem) respondeu conhecer apenas o repositório institucional da UFC (que não é um repositório de dados). Um dos entrevistados afirmou não conhecer o repositório institucional da UFC, e, apesar de saber da existência de repositórios e sites onde se podem obter dados para subsidiar a pesquisa, este não recorda os nomes e afirma nunca ter precisado, sendo esta a justificativa para não utilizar. E2 afirmou estar ciente da existência do Repositório Institucional da UFC e de outros repositórios de dados e sites onde se podem obter dados para subsidiar a pesquisa. Apesar de saber que existem, não recorda os nomes, embora tenha dito que sabe onde pesquisar e obter seus insumos de pesquisa. Inclusive, E2 enviou o *corpus* textual utilizado em sua dissertação para um repositório de dados indicado pela bibliotecária (a mesma pessoa responsável por esta pesquisa) em ocasião de um atendimento personalizado na Biblioteca de Ciências Humanas da UFC (BCH-UFC).

Alguns entrevistados relataram conhecer e citaram os nomes desses repositórios de dados, embora nem todos costumem utilizar, ao passo que outros consideram o Portal de Periódicos da Capes e as bases de dados como repositórios de dados. É mister discutir o quanto a universidade necessita se inteirar das questões relacionadas à gestão de dados científicos e da Ciência Aberta, se compararmos com indícios simples, como o fato de ter ou não conhecimento sobre a infraestrutura de dados. Estar ciente da existência dos repositórios institucionais é o mínimo e que qualquer pesquisador da instituição deveria saber, entretanto, a realidade e os dados demonstram que ainda há uma parcela considerável na UFC que não só desconhece a existência dos repositórios de dados de pesquisa generalistas ou de acordo com a área de conhecimento, mas também desconhecem o próprio repositório institucional.

Olha, é porque tem uma hora [...] que às vezes eu fico com uma dúvida, certo? Se, às vezes, eu falo da minha experiência, porque eu acho uma experiência muito peculiar, [...] ou de como acontece no que eu vi. Não só na minha turma de mestrado, como no programa de pesquisa [...], e conversando com pessoas de outros programas [...]. [...] saber o que é um repositório institucional, [...] lá no Centro de Tecnologia da UFC, poucas pessoas sabem o que é [...], mesmo professores. [...] a minha observação foi essa. Eu, numa coisa muito solitária, muito individual, certo, por querer aprender, [...] é que com o tempo fui aprendendo o que são os repositórios institucionais e aí conheci vários [...]. Mas, não é uma coisa comum de ficar se conversando lá no Centro de Tecnologia. São poucos. São raros, é isso que eu quero dizer. Agora assim, eu não conhecia, **não conheço repositório na minha área, isso eu nunca vi.** Um **repositório na área ambiental**, da



engenharia ambiental, de saneamento, nunca nem ouvi falar que existia. Os que eu conheço é Repositório Institucional da UFC [...], o da Capes, Banco nacional de teses e dissertações, se eu não me engano é esse o nome. (E11).

Além do exposto acima, é preciso esclarecer que há uma grande confusão por parte de algumas áreas ou profissionais sobre o que são e quais as diferenças entre repositórios, bibliotecas virtuais, bibliotecas eletrônicas, portais de pesquisa e bases de dados. O relato de E11 reflete sua visão e percepção a respeito do desconhecimento sobre repositórios institucionais e de dados científicos, mas, também ele confunde a tipologia, e, inclusive afirma desconhecer repositórios de dados em sua área, todavia, a área ambiental possui vários repositórios e uma infinidade de projetos relacionados ao meio ambiente em que são coletados e disponibilizados dados de pesquisa, a exemplo dos que estão cadastrados no portfólio do Sistema de Informação sobre Biodiversidade Brasileira (SIBBR).

Investigando sobre as formas de documentação da pesquisa por meio da elaboração de um Plano de Gestão de Dados (PGD) ou instrumentos semelhantes para esta finalidade, nenhum dos 12 entrevistados nunca elaborou um PGD ou sequer souberam responder o que era. Em vista disso, a pesquisadora solicitou que informassem então qual seria o tipo de estratégia utilizada para documentar as pesquisas desenvolvidas (Quadro 17).

Quadro 17 - Tipos de instrumento ou estratégias de documentação das pesquisas

(continua)

<b>VOCÊ COSTUMA DOCUMENTAR SUAS PESQUISAS ELABORANDO UM PGD OU UTILIZA OUTRO INSTRUMENTO PARA ESTE FIM?</b>		
<b>ENTREVISTADO</b>	<b>ÁREA</b>	<b>INSTRUMENTO / ESTRATÉGIA</b>
E1	Ciências Sociais Aplicadas (Finanças)	Projeto e cronograma.
E2	Letras, Linguística e Artes (Letras Espanhol)	Planejamento mental e anotações.

Quadro 17 - Tipos de instrumento ou estratégias de documentação das pesquisas

(conclusão)

<b>VOCÊ COSTUMA DOCUMENTAR SUAS PESQUISAS ELABORANDO UM PGD OU UTILIZA OUTRO INSTRUMENTO PARA ESTE FIM?</b>		
<b>ENTREVISTADO</b>	<b>ÁREA</b>	<b>INSTRUMENTO / ESTRATÉGIA</b>
E3	Ciências Humanas (Geografia)	Anotações, clipping e esquemas mentais
E4	Letras, Linguística e Artes (Letras Inglês)	Anotações e documentação dos passos metodológicos.
E5	Ciências Sociais Aplicadas (Administração)	Anotações em arquivos eletrônicos criados no computador e usa o gerenciador de referências Mendeley como auxiliar no processo de documentação da pesquisa.
E6	Ciências da Saúde (Odontologia)	Cronograma.
E7	Engenharias (Engenharia de Produção)	Arquivo Read me.
E8	Ciências Humanas (Psicologia)	Cronograma e Relatórios.
E9	Ciências Agrárias (Ciência do Solo)	Caderno / Diário de campo e as anotações são posteriormente transpostas para planilhas.
E10	Ciências Biológicas (Ciências Marinhas e Tropicais)	Planilhas em que constam variáveis, metadados, dados observacionais, testes estatísticos etc.
E11	Engenharias (Engenharia Civil / Sanitária / Recursos Hídricos)	Planilhas em Excel ou tabelas no Word.
E12	Ciências Exatas e da Terra (Ciência da Computação)	Anotações e Arquivo Read me.

Fonte: Dados da Pesquisa (2019).

Conforme os dados do quadro acima, constata-se que os pesquisadores usam meios próprios e de acordo com a sua área para a documentação da pesquisa, como a criação de um projeto, cronograma de execução, planilhas, anotações, read me etc. Embora seja um documento desconhecido para eles, o modelo poderia ser adotado e todos os detalhes e informações a respeito dos dados de pesquisa desses pesquisadores poderiam estar expresso no PGD, de acordo com a opinião de E9, que coaduna com essa discussão: “[...] Eu sinceramente [...] desconhecia que tinha uma metodologia para isso ou um passo a passo [...] desse documento. Eu acho

interessante porque, assim, eu só deveria oficializar o que eu já faço.” De maneira complementar, o entrevistado 7 afirmou não usar o PGD e justificou suas razões:

Não, porque é muito simples, né? Nós, por exemplo, pegamos uma linguagem de programação, geramos um arquivo com esses números, são aleatórios, mas, tem significado para eles, e assim, não é algo que a gente precise de um sistema de gerência de banco de dados, nem de ferramentas mais robustas, nossos dados são muito simples. São matrizes com números, só. Somente isso. Nós fazemos isso empiricamente, em geral. (E7).

Apesar de dizer que não documenta sua pesquisa e justificar que são dados muito simples, apenas números e que não existe uma documentação para isso, a pesquisadora ficou intrigada com a resposta, então, tentou se aproximar da linguagem e jargão utilizado pelas Engenharias, e, dessa vez, perguntou se o pesquisador costumava utilizar, por exemplo, arquivos Read me para esse fim. Imediatamente, a resposta do pesquisador foi positiva e demonstrou através de sua atitude ter despertado para o fato que ele realiza sim a documentação de sua pesquisa, embora não seja por meio de um PGD, mas, por um documento bastante difundido em seu campo de pesquisa.

#### **4.3.2 Ciência Aberta**

Precipuamente, com o surgimento da nova tendência de disponibilização dos dados associados aos artigos, seja como material suplementar ou depositado em repositórios de dados, estudos demonstram como o aumento no compartilhamento de dados de pesquisa, o surgimento de publicações e entidades que passaram a exigir a disponibilização dos dados afetou a comunicação científica. Cerca de 500 artigos publicados em 2009 tinham apenas 9% dos artigos revisados e que depositaram os dados brutos de forma completa. (ASSANTE *et al.*, 2016; NICHOLAS *et al.* 2015; SQUAZZONI *et al.*, 2019). Perante esse cenário, a revisão dos dados de pesquisa por pares passou a ser adotada por diversas revistas, a exemplo da PLoS e Springer Nature.

Assim, investigou-se se algum dos pesquisadores, no exercício do papel de avaliador, já teria vivenciado alguma experiência em que além de avaliar o artigo foi necessária a solicitação para avaliar os conjuntos de dados, ou caso não tivessem passado por essa situação, perguntou-se se estariam preparados para avaliar tanto o artigo como os *datasets*. Os entrevistados E1, E3, E5, E6, E7, E10 e E12 responderam que depende da área de conhecimento e que poderiam fazer isso se fosse de seu

domínio ou especialidade. A entrevistada E8 respondeu que estaria preparada, dependendo do tipo de pesquisa e se fosse de sua área, contudo, faz várias ressalvas, levanta questões e aspectos relevantes que devem ser levados em conta:

No geral, é muito difícil enquanto avaliadora a gente analisar os dados da pesquisa num banco de dados. Por que? Porque a gente recebe muitos artigos para dar parecer. Analisar, pra mim, dados de pesquisa são dos meus alunos de mestrado e de doutorado, que eu acompanho. Então, [...] se uma revista chegar e disser assim: "Para você dar o parecer [...] vai ter que analisar os dados". Eu não vou dar o parecer. Por que? [...] Eu não tenho tempo, e às vezes pode ser que eu não tenha propriedade daquela técnica estatística ou de análise de conteúdo, ou do que for. O software que ele utilizou. Então, eu acho que para uma revista exigir esse nível, eu teria que ter algo a mais para dar aquele parecer. Eu não sei se um suporte... Na verdade, eu não teria interesse. Sendo bem sincera. Eu colocando na mesa quais são as quantidades de atividades que eu desempenho, essa não seria uma atividade onde eu pegaria o meu tempo para fazer isso. [...] Porque pra mim, [...] analisar os dados, que não for da minha pesquisa, onde é que eu vou tá fazendo isso? Com os meus alunos da pós. Ou alunos da graduação que eu esteja orientando. Então, eu analisar isso para um artigo, para mim, não é uma atividade que eu colocaria como prioridade nem como atividade rotineira, tá? Eu fazer uma análise de um artigo a partir dos dados que ele colocou lá como verdade, né? Como dados, eu consigo fazer, mas não chegar no nível de pegar um banco de dados para eu poder analisar e saber até que ponto aquela análise [...] tá correta ou [se] ela acaba em outras coisas. Eu acho que isso é exigir demais do parecerista. Eu acho que isso pode ser tarefa da revista. Se a revista acha que é importante, então, ela faça esse trabalho pela equipe dela, e passa depois para o parecerista [...] dar um parecer de uma forma geral. Mas, essa especificidade, eu acho que não cabe aos pareceristas, cabe à edição da revista. (E8).

O depoimento de E8 apresenta importantes preocupações em relação à prática da revisão de conjuntos de dados, pois, como a maioria das revistas se concentra no processo de revisão e publicação tradicional, além de ter políticas voltadas apenas para isso, não é de se estranhar que pensar na concentração de mais uma atividade que requer muito mais do que um parecer apropriado, também exija o domínio específico e habilidades para garantir que aqueles dados são condizentes com o que foi apresentado no artigo e que a estrutura dos dados brutos está completa. Similarmente, E11 acrescenta pontos que complementam os aspectos expostos por E8 em relação ao papel do avaliador e a tarefa de verificar e averiguar os *datasets*: "Isso talvez crie um problema para quem é avaliador e para as revistas. [...] São professores universitários com mil tarefas e tal, [...] ainda tem que analisar os dados? E tem que fazer tudo lá no deadline, [...] isso pode fazer com que as revistas atrasem."

Para a tarefa de avaliar um artigo e os *datasets*, E4 não se considera preparada porque não recebeu nenhuma formação para tal, limitando-se a cumprir o papel de avaliadora somente do manuscrito.

[...] Não recebi nenhum preparo para isso, nunca soube nem que se fazia esse tipo de avaliação, não teria critério, não sei nem que critérios poderiam ser, então, eu realmente não estaria em condição de avaliar uma coisa que foi produzida por outra pessoa, segundo critérios que só ela detém e que se ela divulgar, como eu só estou acostumada a analisar os critérios que só eu detenho, eu não ia saber comparar para saber se foi bem feito, mal feito, bem coletado, mal coletado, bem armazenado, mal armazenado. Vamos dizer assim, nem bem ou mal, mas, adequadamente ou não adequadamente. Eu não saberia dizer. A gente só recebe preparo para analisar os textos. Tá bem escrito, tá mal escrito, a metodologia faz sentido, não faz sentido, as conclusões fazem sentido, não fazem sentido, a fundamentação teórica faz sentido ou não faz sentido... Mas, os dados? A gente nunca vê também. Pelo menos nas Letras, ninguém vê os dados de ninguém, não. Acesso apenas ao manuscrito. [...] Os únicos dados que eu sei que são disponíveis, são os *corpora* linguísticos das grandes instituições que estão online ou que são disponibilizados via solicitação, mas fora isso... (E4).

É interessante notar que E4 é da área de Letras e trabalha com *corpus* textual, por isso faz todo o sentido a insegurança e a declaração de que não está preparada para avaliar os conjuntos de dados associados aos artigos, afinal, os *corpora* textuais são fundamentalmente montados de acordo com a preferência e conveniência do pesquisador, logo, mesmo que descreva o processo de montagem de seu *corpus*, ainda assim, existiriam elementos que não são facilmente detectáveis ou reproduzíveis até mesmo para um avaliador. Ademais, E4 afirmou que nenhum pesquisador recebe formação ou sequer uma preparação para a execução desse tipo de atividade, limitando-se, portanto, somente à avaliação do artigo.

Outro ponto relacionado às questões que se cruzam com os dados de pesquisa, a avaliação por pares e a Ciência Aberta é a revisão por pares aberta. Visionária, a pesquisadora Julie Hurd (2000) apresenta em seu artigo seminal intitulado “*The Transformation of Scientific Communication: A Model for 2020*”, apresenta a sua visão de futuro a respeito da revisão por pares. Segundo a autora, a revisão por pares será característica de qualquer sistema de comunicação, mesmo que os mecanismos dirigidos à qualidade possam diferir em um processo de submissão e revisão digital. Complementa ainda que os colégios invisíveis continuarão a existir apesar das estações de trabalho em rede possam servir de base da comunicação entre os membros e ampliar a sua afiliação, sendo assim um colégio invisível baseado na Internet. A descrição de Hurd (2000) é um retrato fiel do cenário atual da comunicação científica, especialmente nos aspectos relacionados à revisão por pares, dos periódicos científicos e dos diversos tipos de repositórios, incluindo-se

ainda as redes formais e informais de divulgação e publicação dos artigos, e também dos dados.

Nesse sentido, investigou-se qual a opinião dos pesquisadores sobre isso. O Entrevistado 1 revelou ser uma visão estranha para ele, pois, de acordo com a sua opinião, seria muito mais prático fazer o upload dos dados e do artigo em um repositório de sua área e acompanhar os comentários, críticas e sugestões para o trabalho, como já acontece no repositório NBER, ArXiv, entre outros. Ainda segundo o pesquisador, faria mais sentido, a partir dessas discussões, aplicar as melhorias ao artigo e aos dados, e posteriormente submeter para um periódico científico. A prática descrita por E1 assemelha-se a da disponibilização dos pré-prints em repositórios.

Declararam desconhecer esse tipo de prática E2, E3 e E11, e os demais informaram que já ouviram falar ou que conhecem a avaliação por pares aberta, embora nunca tenham passado pela experiência quer seja como avaliador ou avaliado, exceto E9 que já passou por avaliação aberta em um de seus artigos submetidos para um periódico.

Eu **acho que é positivo**, inclusive eu **já publiquei trabalhos que o nome dos avaliadores apareceu lá, [...] trabalhei com revistas que mostravam isso, [...] parecer aberto**. E assim, **eu me senti bem mais à vontade sabendo que o meu trabalho foi avaliado por uma pessoa que é bem conhecedora da área**. Então, se uma pessoa que é uma sumidade naquele ponto, se ele avaliar meu trabalho e ele for publicado [...] ganha mais credibilidade. Não tive problema. Por outro lado, tem uma revista de fora, [...] eu não conhecia ninguém, eu não sei o que vai acontecer quando o meu trabalho sair. [...] Mas, **eu concordo**, eu acho positivo sim. **Tanto eu acho positivo isso, colocar o nome dos avaliadores e divulgar, como deixar claro o que cada autor fez no trabalho também**. Ah! Isso sim eu tenho certeza que eu já publiquei dessa forma. (E9).

Sobre o depoimento de E9 acima, o pesquisador destaca que considera igualmente importante a informação da função e do que cada autor fez em um manuscrito, e, nesse caso, muitas revistas já adotaram a Taxonomia CRediT adotada por diversos periódicos e serve para discriminar os papéis desenvolvidos na criação de um trabalho.

A maioria dos pesquisadores entrevistados considera positiva a abertura no processo de avaliação por pares, embora algumas ressalvas a esse respeito tenham sido feitas.

Olha... Relativizando a coisa, eu acho que assim, **tem um aspecto positivo, tem um aspecto negativo**. Eu acredito que assim, estando aberta, você abre para mais críticas construtivas, [...] mais melhorias, [...] ideias novas, coisas que você pode não perceber antes, e aí chega alguém e ele diz o que você

pode fazer e ajuda, mas, eu acho um pouco delicado na **configuração brasileira**, né? [...] **Eu acho delicado** esse ponto, de você saber quem vai ser analisado e quem vai analisar, abrir assim também, [...] muitas vezes **falta um certo profissionalismo**. Então, por um lado, você pode abrir para pessoas que vão ajudar, por outro, você pode abrir para pessoas que vão atrapalhar. Eu não tenho um posicionamento claro [...], como eu não conhecia essa ideia, eu não consigo te dizer agora o que pensar sobre isso. [...] Tem o aspecto positivo que eu acho que tem quer ser considerado, dessa **possibilidade de uma troca mais ampla**, por assim dizer. (E2).

Na mesma linha de pensamento, E7 acrescenta que a avaliação por pares deve ser aberta, por que isso dá transparência ao processo. Além disso, coíbe práticas como a autocitação, plágio, autoplágio e roubo de ideias. Contrariamente aos que são a favor da avaliação por pares aberta, os pesquisadores E5, E6 e E11 preferem o modelo de avaliação cega (*Blind-Review*) para as publicações científicas, por considerar que a avaliação aberta pode ser tendenciosa, causar algum desconforto ou disputa acadêmica. O entrevistado 5 acredita que “[...] o sistema de publicação brasileiro atual não tem essa maturidade para seguir a avaliação aberta”. Da mesma forma, E6 concorda que o *Blind-Review* é a melhor forma de avaliação dos artigos.

Um dos temas recorrentes no universo da Ciência Aberta é a reprodutibilidade da pesquisa. Com efeito, o aprofundando das questões relacionadas à Ciência Aberta que se interligam com os dados de pesquisa e os processos de abertura das pesquisas. Assim, indagou-se cada um dos entrevistados a respeito da crise de reprodutibilidade da ciência na atualidade.

A importância de se discutir a reprodutibilidade da pesquisa reside no fato de que são pilares da pesquisa científica, principalmente, em áreas que não conseguem avançar o conhecimento sem que se possa realizar experimentos, simulações, testá-los, replicá-los e reproduzi-los. Exige-se de qualquer pesquisador que siga um método científico, e, como parte do processo de pesquisa, a reprodutibilidade (a capacidade de se atingir os mesmos resultados, usando bases e informações armazenadas) e a replicabilidade (busca alcançar as mesmas conclusões ou resultados usando dados e processos distintos) são essenciais. (ASCHWANDEN, 2015; FIRMIN, 2019; MARQUES, 2019).

Nesse sentido, os tópicos levantados pelos pesquisadores dizem respeito a várias peculiaridades e variáveis relacionadas à reprodução de um experimento, perpassando até por problemas de compreensão dos conceitos causados por uma educação incipiente ou de estudos sem profundidade:

Eu estou ciente disso. Mas, isso daí surgiu com um problema na psicologia, né [sic]? **Na ciência tem várias besteirinhas que mudam, por exemplo, eu fazendo uma reprodução de algo aqui e algo lá no Equador, físico, pode ter uma interferênciazinha só da latitude e longitude que eu tenho que fazer considerações e tal, e às vezes, isso não está no artigo, eu tenho que saber o conceito. [...] Um físico quando veio aqui para o Brasil falou que um dos defeitos da gente nas pesquisas é que a gente não absorve os conceitos da forma certa. Aí a gente começa a fazer ciência com conceitos aproximados do que são os conceitos fidedignos ou preestabelecidos. A leitura é ruim, tá entendendo [sic]? E aí, eu acho que reprodutibilidade tem que tomar muito cuidado porque senão fica aquela coisa, é reproduzir, é refazer o experimento. Sim, mas, **tem peculiaridades locais**, [...] de laboratórios, [...] que para eu reproduzir aquele a coisa eu tenho que fazer essas considerações. Em ciências sociais, idem. Tem muito caso aí de teorias que o pessoal diz "ah, não é aplicável no Brasil.". Sim, ela tem uma peculiaridade, ela foi aplicada em um determinado canto, quando venho pro Brasil, eu tenho que ver quais são as peculiaridades. **Eu trabalhei com o tráfego, uma época. Um dos problemas aqui que a gente via quando ia trabalhar com formação de fila e congestionamento, é porque no Brasil a gente usa muito carro 1.0. Aí tu vai pra outro canto [sic], os carros são 2.0, a resposta do carro é totalmente diferente. Aí tinha gente querendo refazer usando a mesma [coisa], o usando o mesmo software que usava na Inglaterra. É totalmente diferente. Eu tenho que entender os conceitos e tal pra reorganizar o modelo pra replicar o que foi que aconteceu lá. Se eu não fizer isso, eu não tô replicando, eu tô copiando [sic].** Esse é um grande problema. **Eu tenho que entender bem o conceito pra fazer replicabilidade. Por isso que eu tenho um certo pé atrás, que eu não sei se a replicabilidade vai servir é para refutar teorias que são corretas. [...]** 'Essa auditoria' científica, tem que tomar cuidado [...]. **O que tá acontecendo na verdade não é replicabilidade, é auditoria.** (E1).**

O depoimento de E1 apresenta elementos que estão intrinsecamente ligados à questão da reprodutibilidade da pesquisa em dimensões mais amplas, a exemplo dos déficits de aprendizagem, variáveis que podem mudar os resultados e sobre a preocupação de refutar uma teoria válida. Inclusive, a palavra 'reprodutibilidade' tem ganhado maior destaque após o estopim da crise de reprodutibilidade na ciência e os escândalos envolvendo a prestigiosa revista Science que publicou um artigo sobre um experimento, que mostrava que os conservadores políticos tinham reações fisiológicas mais fortes às imagens ameaçadoras do que os liberais. (ARCENEUX *et al.*, 2019; BARBA, 2018; FANELLI, 2018; LIU; SALGANIK, 2019).

Adicionando mais elementos ao assunto 'reprodutibilidade', E7 assevera que esse é considerado um fator de qualidade em seu campo de pesquisa, inclusive sugere que isso deveria ser inserido como um indicador de qualidade da ciência no Brasil, mas, o pesquisador relata algumas dificuldades encontradas nesse percurso:

**Principalmente na minha área, um dos pilares da qualidade da pesquisa é a reprodutibilidade.** Nós trabalhamos com algoritmos, e, geralmente, nós temos que comparar o que nós fazemos com os outros, então, [...] vamos ter que pegar o que aquela pessoa fez, de certa forma, vamos ter que implementar o algoritmo da outra pessoa [...]. Eu, de minha parte, [...] não



vejo nenhum problema em compartilhar um código, né? Mas, as pessoas não respondem. Aí se nós tivermos dúvidas acerca do conteúdo, como é que a gente vai comparar de forma justa? Então, **essa questão da reprodutibilidade é muito importante. Há algum tempo atrás eu tive acesso a um dado, não sei se era correto, consistente, mas que, das pesquisas aqui no Brasil, apenas um sexto tinha algum retorno, cinco sextos era desperdício. Então, acho que você atrelar isso ao fomento público, melhorar os indicadores da qualidade da ciência brasileira é fundamental.** (E7).

No que diz respeito à percepção sobre a Ciência Aberta, todos os entrevistados afirmaram ser este um tema novo ou que conhecem muito pouco. Assim expressaram sua percepção, conforme mostra o quadro 18 a seguir:

Quadro 18 - Percepção sobre a Ciência Aberta

(continua)

O QUE VOCÊ ENTENDE POR CIÊNCIA ABERTA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E1	Ciências Sociais Aplicadas (Finanças)	“Para mim, a Ciência Aberta é uma <b>colaboração continuada entre diferentes pesquisadores [...]. A ciência ter a função inicial que ela tinha, que é auxiliar a sociedade, a humanidade.</b> Hoje em dia, a ciência, como está, que é dita <b>'fechada'</b> , toda e qualquer produção científica, <b>ela não é acessível.</b> Então, o conhecimento não é acessível. A ideia para mim da Ciência Aberta é <b>compartilhar conhecimento.</b> É deixar o conhecimento fluído na sociedade.”.
E2	Letras, Linguística e Artes (Letras Espanhol)	“[...] Conforme o que a gente tem conversado, seja essa <b>nova tendência de disponibilização de ideias, [...] de informações referentes às pesquisas que são realizadas, os dados, que é um ponto central, essa abertura nas discussões enquanto produção de material, [...] quanto material é produzido, está sendo formulado, há uma ampliação na possibilidade de conversas e trocas de ideias, essa própria abertura das identidades das pessoas, de você saber quem lhe avalia, de você saber quem é avaliado.</b> Apesar de problemático, em certos aspectos, para mim, eu acredito que seja essa <b>maior clareza no aspecto da pesquisa em si,</b> do texto, do material, de tudo que envolve o documento, também das pessoas envolvidas.”.
E3	Ciências Humanas (Geografia)	“ <b>Não tenho a mínima noção... Ciência Aberta...</b> Eu penso assim, <b>algumas áreas da ciência, o negócio é muito fechado, por exemplo, nas ciências da saúde, a farmácia, têm segredos que são inclusive privados, de laboratórios. Então, ali não tem Ciência Aberta [...].</b> A empresa júnior, é um absurdo fazer um mapa ali... Porque cobra. Eu nem gosto do nome empresa júnior [...]. [Se] aquele tipo de iniciativa fosse como uma espécie de extensão...”.

Quadro 18 - Percepção sobre a Ciência Aberta

(continuação)

O QUE VOCÊ ENTENDE POR CIÊNCIA ABERTA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E4	Letras, Linguística e Artes (Letras Inglês)	<p>“<b>Só noção. Noção muito vaga.</b> Quando falam 'Ciência Aberta', falam assim: "[...] procura tal revista que é aberta". Ou então: "procura esse corpus, que é aberto". Ou então: "procura essa ferramenta de análise que o código é aberto". <b>Significa que você usa sem pagar</b>, basicamente. Simplificando muito, a ideia que você tem é que [...] <b>vai ter acesso à essas ferramentas, a esses dados e a essas publicações sem ter que remunerar as instituições que disponibilizam esses dados, ferramentas e publicações.</b> Mas, <b>eu imagino que tenha muito mais do que isso, que seja muito mais complexo do que isso</b> [...]. Eu acho que na origem disso tudo tem uma <b>questão ética muito forte</b>, uma questão de <b>paradigma de ciência, porque não chamam de dado aberto, não chamam de ferramenta aberta, chamam de Ciência Aberta.</b> Então, <b>o paradigma de ciência, de fazer científico e dentro de todo paradigma tem um conjunto de valores: éticos, humanos e a maneira como até você pensa o mundo tem influência quando você diz que uma ciência é aberta.</b> Então, <b>eu acredito que seja muito mais do que simplesmente não pagar, são as implicações desse não pagar que são importantes</b> [...].”</p>
E5	Ciências Sociais Aplicadas (Administração)	<p>“Engraçado, [...] quando você falou, <b>é um termo que eu não conhecia.</b> Você falou de Ciência Aberta e eu <b>fiquei pensando em 'universidade aberta'</b> do Brasil, que é o programa que eu [conheço]. Mas, não é um termo que eu conhecia, não. [...]. Essa ideia de <b>compartilhar a pesquisa, compartilhar dados, né?</b> Ter uma <b>integração dos pesquisadores</b>, tentar uma coisa <b>mais transparente...</b> Para mim, isso é uma coisa nova.”</p>
E6	Ciências da Saúde (Odontologia)	<p>“Pelo pouco que eu sei, eu acho que é... Acho que foi você que me falou. (risos). [...] Mas, eu acho que é essa tendência aí de... [...] se a pesquisa for financiada pelo órgão público, então, se entende que <b>aquela pesquisa não é sua, não no sentido de você ser o dono daqueles dados</b>, inclusive você deve torná-la <b>pública.</b> [...] Então, eu acredito que seja a Ciência Aberta nesse sentido, de <b>permitir uma ligação, uma conexão entre membros de uma comunidade científica</b>, de modo que você possa <b>minimizar o tempo</b>, porque você não vai coletar de novo os dados, né? Você permite que esses dados sejam mais divulgados, [...] então, eu acredito que é mais nesse sentido, de uma <b>cooperação.</b>”</p>
E7	Engenharias (Engenharia de Produção)	<p>“Esse termo '<b>Ciência Aberta</b>', eu nunca tinha ouvido falar... Assim formalmente, <b>eu acho que tá relacionado pelo sentido aí dos seus questionamentos, é para códigos abertos, repositórios abertos, revisões por pares abertas...</b> [...] Eu, inclusive, [...] só utilizo software aberto. Infelizmente, nesse computador não, porque eu sou obrigado a usar o sistema operacional fechadíssimo dele, mas, <b>todo o nosso laboratório</b>, tudo é Linux, tudo é Libre Office, tudo é Julia, <b>tudo é aberto.</b>”</p>

Quadro 18 - Percepção sobre a Ciência Aberta

(conclusão)

O QUE VOCÊ ENTENDE POR CIÊNCIA ABERTA?		
ENTREVISTADO	ÁREA	PERCEPÇÃO
E8	Ciências Humanas (Psicologia)	<p>“<b>Eu não sei o que é a Ciência Aberta</b>, não. [...] A única coisa que eu já ouvi falar foi esses bancos de dados abertos. É... Eu acho assim, por exemplo, eu tenho alguns requisitos: Eu não publico em revista que eu tenho que pagar. <b>Revistas que são em acesso aberto</b>. Ou ter acesso a algumas [revistas], que a gente consegue pelos periódicos CAPES, porque o governo já pagou, né? Então, <b>para mim isso é um elemento importante, eu acho que qualquer pessoa pode ter acesso</b>. Eu não sei, assim, essa bandeira 'Ciência Aberta', para mim... Ela falta alguma coisa, sabe? [...] Eu acho que tem que ser ciência comprometida, [...] que dê conta da realidade, [...] que preste conta dos recursos. Ah! Então, eu acho que o tema vai mais além [...]. Ciência pra quem? Pra quem é essa ciência? Com quais problemas ela tá lidando? Para que público ela chega? Então, eu acho que tem outras coisas também, que eu não conheço muito esse movimento, mas que eu questiono, que pode ser que esteja dentro, né? [...] Eu acho que tem que ter outros elementos. Universidade para quem? Então, eu acho que tem várias coisas aí, mas, se for adequada, que <b>as pessoas possam ter acesso</b>, eu concordo. [...]”.</p>
E9	Ciências Agrárias (Ciência do Solo)	<p>“<b>O termo, na verdade, é novo para mim</b>. Mas, eu acho que, ao pé da letra, Ciência Aberta é uma forma de fazer ciência de uma maneira mais transparente. De deixar disponível quem avaliou, deixar disponível dados, o que for preciso para que seja reproduzido ou não, para que seja testado. <b>Eu acho que Ciência Aberta é transparência</b>.”.</p>
E10	Ciências Biológicas (Ciências Marinhas e Tropicais)	<p>“[...] Não tenho grande conhecimento [...]. Eu acho que é a noção, assim, de que existe como se fosse uma grande comunidade internacional, [...] e as informações, tanto a respeito da maneira como você coleta os dados, os dados que foram coletados, a interpretação desses dados, [...] os resultados obtidos e as conclusões geradas, elas são de <b>domínio público</b> [...]. Então, você tem uma grande <b>comunidade que compartilha</b> não só as informações de conhecimentos que são gerados, mas, também <b>o processo como um todo de execução</b>, de chegada até aquela conclusão.”.</p>
E11	Engenharias (Engenharia Civil / Sanitária / Recursos Hídricos)	<p>“Assim, teve uma frase, que eu não sei exatamente de quem é [...]: “Se eu cheguei onde eu cheguei é porque eu subi em ombro de gigantes”. Então, na verdade, <b>a ciência para mim ela sempre é aberta</b>. [...] <b>Ela só consegue crescer, porque alguém já produziu algum conhecimento</b>. [...] Então, [...] a ciência só consegue dar saltos se for assim. E, me parece que a ciência no início, era feita [...] assim, era [...] mais aberta, mais dialogada, [...] e aí depois que passou a surgir essa preocupação com produtivismo, certo, aí começou a ter esses problemas. <b>Hoje, alguém tem que pagar a conta. E eu tô me referindo à história do artigo. Ou você paga para publicar ou você paga para baixar</b>.”.</p>
E12	Ciências Exatas e da Terra (Ciência da Computação)	<p>“Ouvi pouco falar a respeito do termo. Porém concordo com a ideia de que <b>a ciência não deve estar restrita aos laboratórios</b>.”.</p>

Fonte: Dados da Pesquisa (2019).

Conforme o quadro acima, percebe-se que a ideia de Ciência Aberta está muito ligada ao acesso ao conhecimento, à transparência, ao movimento do Acesso Aberto, entre outros. Na verdade, o termo em si tem diversas acepções e perspectivas, como já foi explorado anteriormente, nessa dissertação. Assim, acrescenta-se que para alguns dos entrevistados a ciência é universal, então, o termo em si não deveria

ser empregado dessa forma. Em outras palavras: “[...] as gerações futuras irão considerar o termo “Ciência Aberta” como uma tautologia - um retrocesso de uma época antes do surgimento da ciência. A “Ciência Aberta” simplesmente se tornará conhecida como ciência, e as práticas secretas e fechadas que definem nossa cultura atual parecerão tão primitivas para eles quanto a alquimia é para nós.” (NOSEK, CHAMBERS, 2015, não paginado, tradução nossa).

E2 destaca a importância de que os estudantes que ingressam na universidade tenham contato com esse tipo de abordagem da Ciência Aberta por meio das informações repassadas pelos bibliotecários e defende uma integração entre esses profissionais e os professores.

Eu entendo que é um trabalho grande, que seria assim, bastante puxado, complicado para os bibliotecários, mas, eu acredito que seria muito interessante, até por uma questão de contato com novos pesquisadores. Os ingressos, aqueles que entram na universidade, que possam ter um contato maior com essa questão da Ciência Aberta, ao ter um contato com os bibliotecários, com os professores, que possam até mudar certas mentalidades que são bem mesquinhas, assim, para não usar uma palavra um pouco delicada. Então, eu acredito que essa mentalidade, essa troca de informações por parte do corpo da universidade, eu acho que é muito importante para quem entra, para que possa entender como esse tipo de pesquisa é interessante e é importante esse tipo de ciência, assim, mais claro, mais disponível. (E2).

E4 evoca pontos interessantes que são muito caros para a Ciência Aberta, o compartilhamento e a produção de conhecimento hodiernamente, especialmente diante de um cenário em que há barreiras de acesso ao conhecimento.

A gente aqui ainda é muito elitista [...] com a produção de conhecimento e a divulgação de conhecimento. E a divulgação de conhecimento não é uma questão de ego, não. É uma questão de compartilhamento. Se você produz conhecimento só pra si [...] Mas, se você quer compartilhar conhecimento, você tem que divulgar porque as pessoas só podem compartilhar contigo o que elas sabem que você fez. Se elas não sabem o que você fez, não tem divulgação. Eu sou politicamente posicionada de um jeito, em termos de conhecimento, eu sou um tanto quanto diferente, eu sou liberal, do ponto de vista de conhecimento. Eu acho que tem que ser o livre acesso, não o livre mercado, mas, o livre acesso. Reservas de mercado em termos de conhecimento, elas são elitistas, não inclusivas, e eu considero antiéticas. Ética entendida como aquilo que é bom pra ti quando é bom pro outro. (E4).

A partir da visão de E4, Suber (2007, p. 194, tradução nossa) reforça que “O conhecimento não é uma mercadoria”. O conhecimento deve ser livre e deve estar disponível para todos como um bem público. De acordo com o autor, o pesquisador participa da economia da dívida, não tem interesse econômico e tampouco poderia ser reduzido a um bem patrimonial ou mercantilizado. Sua recompensa é o

reconhecimento, a reciprocidade, a reputação e a responsabilidade, em outras palavras, é o capital do conhecimento e reconhecimento (BOURDIEU, 2001).

Uma das reflexões trazidas por E7 que envolvem diretamente a Ciência Aberta, o acesso livre ao conhecimento e às publicações científicas, é um relato contundente da atualidade e dos modelos de negócio vigentes dos periódicos científicos. Esse é um fator que deflagra outros problemas, principalmente os que estão imbrincados com a ética e integridade da pesquisa.

Na minha área, eu acho que é uma tendência internacional, existe uma pressão muito grande por publicações, e, muitas vezes, as pessoas acabam simulando resultados. [...] Esse é o grande desafio, porque essa pressão, principalmente em países do Oriente, [...] acaba cometendo muitos desvios de conduta. O desafio é nós mudarmos esse ritmo da pesquisa, primar pela qualidade, acabar com essa restrição de ter que publicar em quantidade muito grande num curto intervalo de tempo, isso tá acabando com a pesquisa, está trazendo problemas psicológicos para os pesquisadores, está deturpando o processo de investigação, então, o desafio é quebrar essa lógica, só que isso não é possível porque as revistas ganham dinheiro com isso. Essas revistas internacionais. Nós assinamos um termo [...]. É pior do que aquela mais valia que o Marx [...]. Bem pior. Você desenvolve, você tá usando dinheiro público, você tá doando seu sangue para chegar lá e dizer não é mais meu esse conhecimento, é da editora que tá ganhando \$30 com cada download. Então, é uma lógica perversa, né? Mas a Capes fica como um cordeirinho para essas editoras internacionais. E tem que criar nossos indicadores de qualidade, usar revistas abertas, e se não tiver daquela editorazinha bonitinha, assim... A gente tem que... Pra gente se libertar dessa exploração, mas, isso não vai acontecer, infelizmente. [...] Mas, é essa pressão e essa indústria que está fazendo com que a pesquisa seja desvirtuada. O Einstein não ia ser bolsista do CNPq, né? (risos). [...] Insanidade, isso aí é insanidade.

De acordo com o depoimento de E7, deveriam os pesquisadores escolher primar pela qualidade, buscar seguir métricas alternativas mais justas para medir o impacto de suas pesquisas. Adicionando a esse cenário os dados de pesquisa, gerenciar e compartilhar os dados em repositórios institucionais e em repositórios de dados de sua escolha, e, principalmente, publicar em revistas de acesso aberto são imperativos, ou seja, questionar e desafiar o status quo instaurado na comunicação científica atual e adotar os princípios da Ciência Aberta.

O relato de E7 se entrelaça com os dados de pesquisa, afinal, esses mesmos periódicos, como enfatiza Albagli (2015), agora, não apenas querem ser os detentores dos artigos ou livros escritos pelos cientistas: eles também desejam obter os dados de pesquisa. Atualmente, os pesquisadores se encontram em uma situação em que produzem e pagam para acessar o conhecimento produzido por eles mesmos, perseguem ideais de publicações em revistas prestigiosas com alto fator de impacto;

seguem valores que pregam a quantidade versus a qualidade (produtivismo acadêmico), porém, que muitas vezes não se consegue alcançar tais ideais porque são excludentes e favorecem a poucos, conforme comprova o estudo de Gibbs (1995), que aborda esses pontos e o quanto os países emergentes e periféricos têm suas produções invisibilizadas em detrimento da produção de países ricos. Embora E7 busque alternativas para se desvencilhar dessa situação, por exemplo, publicar em revistas de acesso aberto, eis o que retrata o pesquisador e em outros tipos de problema ele se depara:

[...] As revistas, hoje, da minha área [Engenharias], praticamente todas são de acesso aberto, só que para tanto, elas impõem uma taxa que varia de \$500 [...] a \$1000 dólares [...]. Como a relação real e dólar não tá muito interessante, assim, fica muito difícil. Então, assim, as revistas Open Access, todas são pagas. Algumas não são, mas essas, geralmente, têm o qualis muito baixo. Então, na minha área, nós somos obrigados a publicar. [...] Inclusive, eu estou com um esforço [...] de mandar para uma revista até da Universidade de Havana, em Cuba, de acesso aberto, uma revista que tem certa qualidade, mas, o site não está funcionando... Eu tô há oito dias tentando, falando... Assim, o que acontece: como eles não têm dinheiro e fazem aquilo de uma forma amadora, aí não tem como competir, né? Mas, é assim, [...] eu busco sempre que possível alternativas, mas, na minha área é muito pouco provável. As revistas extrato A [classificação Qualis] são todas essas internacionais. E [ainda têm os] copyrights. A gente perde o nosso próprio [artigo]. Não é mais meu aquilo que eu escrevi, não posso nem usar de novo. [Colocar os artigos em alguns repositórios], no ResearchGate, algumas revistas permitem. Outras, não. (E7).

Adicionando outros pontos complementares, a Ciência Aberta é essencialmente desenvolvida e disseminada de forma que todos possam colaborar, desde pesquisadores até o público leigo, permitindo-lhes contribuir e acrescentar conhecimento à pesquisa, tais como: dados, metodologias, resultados e protocolos, disponibilizando-os gratuitamente em diferentes estágios do processo de investigação. Além disso, antes de compartilhar os dados, estabelecer as condições em que esses dados serão distribuídos a fim de divulgar e ampliar o impacto da pesquisa são fundamentais, pois também devem assegurar as questões referentes aos direitos autorais e de propriedade intelectual.

### **4.3.3 Pesquisa**

Partindo do princípio de que alguns dados de pesquisa são disseminados de diferentes maneiras, não apenas em plataformas de infraestrutura de dados, repositórios, sites e afins, perguntou-se aos pesquisadores quais são as estratégias

utilizadas para a divulgação de suas pesquisas, isto é, em quais canais formais e informais da comunicação científica. Nesse sentido, para Fecher e Friesike (2014, p. 44, tradução nossa) afirmam que “[...] a ciência experimentou uma nova dinâmica à luz das modernas tecnologias de comunicação”, como a utilização de redes sociais, blogs e o uso da internet na divulgação científica dos trabalhos desenvolvidos pelos cientistas, influenciando diretamente os fluxos informacionais ao produzirem novas formas de divulgação, disseminação, avaliação e comunicação entre os pares.

A maioria dos entrevistados prefere divulgar suas pesquisas em artigos, eventos e livros, isto é, pelos meios tradicionais, como explicam os estudos em comunicação científica de Meadows (1999). Realizam divulgação em sites os indivíduos E6, E7 e E8, e, utilizam a rede social para isso apenas E6 e E8 para publicizar atividades dos grupos de pesquisa. Igualmente, E9 declarou estar iniciando o uso de uma página do Facebook para divulgar as atividades do Programa de Pós-Graduação.

Em relação ao uso das redes sociais acadêmicas, usam o ResearchGate os pesquisadores E1, E7, E9 e E10. De todos, apenas E4 usa a Academia.edu como rede social acadêmica. Possuem perfil no Mendeley os sujeitos E1, E5 e E10. A maioria dos entrevistados também possui perfil no Google Acadêmico como uma das formas de divulgação das pesquisas. Somente E7 marca presença na rede social para avaliadores e editores Publons. Sobretudo, “Para um pesquisador, a disseminação é um ponto-chave, pois dá sentido ao seu trabalho ao permitir o registro, avaliação, disseminação e acúmulo de conhecimento.” (SILVA, 2019, p. 5,). Nesse aspecto, apenas E8 utiliza adicionalmente a mídia institucional da universidade, vídeos e a elaboração de cartilhas em linguagem acessível ao público. Entre todos os entrevistados, apenas E7 afirmou disponibilizar alguns dados de pesquisa em seu site profissional e no ResearchGate.

Outro fator ponderado na pesquisa considerou a tendência atual para a disponibilização dos dados associados aos artigos científicos, o que também tem se caracterizado como uma das práticas incentivadas pela Ciência Aberta. Entre os 12 (doze) entrevistados, nenhum publicou artigos em que os conjuntos de dados estavam associados ou foram disponibilizados nos repositórios de dados da revista ou qualquer outro. Alguns dos pesquisadores confessaram não conhecer esse tipo de opção para a publicação de seus artigos ou porque não era uma exigência do periódico. Apenas

E10 declarou estar em processo para publicação de um artigo e de seus dados como material suplementar pela primeira vez:

[...] Das que eu publiquei, não houve essa solicitação, né? Mas, a gente tá num processo de publicação agora em que a revista solicitou para colocar como material suplementar os códigos do programa estatístico que a gente utiliza, [...] que é o R. Então, a revista pede os scripts que foram utilizados na análise para serem colocadas como o material suplementar. E aí sim, como a gente ainda tá num processo de submissão, eu ainda não coloquei, mas, certamente irei colocar brevemente.

O entrevistado 1 informou que nunca publicou artigos em que os dados estão associados e disponíveis em repositórios, mas, afirmou que o periódico já chegou a solicitar os dados para verificação de seu trabalho. De acordo com o depoimento de E7, as revistas em que costuma publicar, recomendam a disponibilização dos dados de forma facultativa, mas, geralmente o referido pesquisador disponibiliza no ResearchGate ou em repositórios específicos da sua área e informam o link para os dados nos artigos disponíveis gratuitamente.

Investigou-se ainda a respeito de aspectos relacionados às pesquisas científicas e que envolvem a necessidade de acompanhamento e transparência do dinheiro público aplicado nesses estudos. Sobre isso, E2 se manifestou favorável e justifica que há o risco de se manipular uma pesquisa, portanto, concorda que há uma necessidade de vistoria ao longo dos processos de pesquisa. Ademais, complementa que se os impostos são pagos, então, deve haver a cobrança de saber como o dinheiro foi aplicado.

Diante do exposto, questões relacionadas à manipulação e mau uso dos dados também foi mencionado por E2. Na perspectiva de E10, a transparência é a melhor forma de combater fraudes e problemas com autoria na pesquisa científica, contribuindo para um profundo senso de ética e justiça no campo da ciência:

[...] **Quando você publiciza tudo, você abre para comunidade, é uma forma sim de controle de fraudes.** A gente sabe que, como eu disse, existem pessoas parece que tem um comichão por fazer coisas erradas, né? Mas, a gente não pode também pautar as nossas ações pensando nessas pessoas que fazem, que deveriam ser a exceção ou que independentemente de ser regra ou não, estão fazendo é a coisa errada. Então, a gente tem que tentar seguir os bons exemplos, independentemente se eles são maioria ou não. Então, eu acredito que **a abertura das informações e do processo em si, é também uma estratégia para tentar evitar esse tipo de fraude.** Assim, o cara sabendo que o processo de revisão é aberto, o autor, por exemplo, vai pensar duas vezes antes de colocar uma coautoria que não seja adequada. Porque eu acho que falta assim, às vezes, é... É tentar mesmo, não sei se seria a palavra certa, constranger, mas, acho que falta mesmo [...] você ir a público e dizer: "Não. Essa atitude é errada. Não se faz esse tipo de prática".



E, às vezes, a gente entra meio que num senso de transformar as coisas em práticas normais e acaba que isso meio que acontece com um monte de coisa ruim no nosso dia a dia? Então, essa questão, por exemplo, das coautorias, que muitas vezes são completamente inadequadas. A própria questão da análise dos resultados, a forma como os dados foram coletados, acho que quanto mais aberto e publicizado isso for, as pessoas que tem uma tendência a fazer isso, naturalmente, [...] vão [se] sentir mais acanhadas, não vão se sentir tão estimuladas [...]. Acho que tornar aberta também é uma maneira de evitar esse tipo de prática inadequada.

Sobre a questão da autoria, no que se refere ao dado, há controvérsias. Borgman (2015) afirma que autoria e outras formas de responsabilidade são apenas convenções, mas, essa visão não se sustenta diante do crescente aumento do número de autores ao longo dos anos (HILÁRIO; GRÁCIO; GUIMARÃES, 2018; LIMA; FARIAS, 2019; MEADOWS, 1999; TARGINO, 2010). Mesmo diante desse fato, a autoria do dado é complexa porque envolve saber quem produziu o dado (proveniência), quem coletou, quem é o responsável, entre uma série de outras questões.

Um ponto crucial levantado por E8 diz respeito à formação do pesquisador e os desafios do compartilhamento que envolvem sua participação no processo de pesquisa. A pesquisa desenvolvida em algumas áreas demanda intensa participação do pesquisador, por exemplo, em uma pesquisa de campo. As experiências adquiridas por aquele que participa ativamente no campo demonstram que, mais do que coletar dados, é preciso *viver* a pesquisa, o que implica participar do processo. Nesse aspecto, a dúvida de E8 agrega valor à pesquisa científica desenvolvida nas pesquisas humanas e sociais, à medida que:

[...] o desafio é exatamente isso [..]: a pessoa não participar do processo, porque pra mim, o dado ele é construído, tá? Então, [...] isso é um elemento, é um desafio. É possível, é. Se eu conseguir entender melhor como é esse compartilhamento. Eu não tô [sic] dizendo que eu não possa mudar. Posso mudar de ideia. Para mim também, [...] esse compartilhar, às vezes eu tenho muito a preocupação: Qual é o nível de conhecimento que eu tenho? Quem vai pegar meus dados? Porque eu tenho muito medo de hoje as coisas... [...] Elas serem construídas de uma forma muito rápida, então, para eu conseguir terminar logo o meu mestrado, eu faço qualquer coisa. Eu corro contra o tempo. E, às vezes, eu não vivencio o processo do que é se formar um pesquisador. **Então, a disponibilidade dos dados ela também passa pelo desafio, até que ponto ela ajuda a formar um pesquisador, ou se ela corta etapas na formação do pesquisador.** Porque é uma das coisas mais difíceis, é você pensar um tema de pesquisa, [...] delinear uma pesquisa, [...] pensar nas etapas, [...] construir o instrumento, [...] coletar, [...] analisar, [...] escrever, [...] devolver. Então, é um processo longo. E eu tenho medo até que ponto essa disponibilidade possa tirar algumas dessas etapas. E você ter às vezes um pesquisador especialista em trabalhar com bancos de dados abertos. Ele vira especialista nisso, mas, ele não sabe chegar no campo. Como às vezes a gente tem professores que são somente pesquisadores.

Ele não sabe chegar numa sala de aula. Ou é só de laboratório. Então, um desafio para mim é até que ponto etapas podem ser deixadas de lado por eu ter já um banco. Quem pode pegar [...]? Qualquer pessoa? Tem alguns limites? Então, assim, eu acho que a minha preocupação muito é a formação do pesquisador. E para mim, ainda mais do pesquisador comprometido. Porque para mim não é um pesquisador que só coleta dados. Para mim, é um pesquisador implicado com [...] o que ele faz e possa transformar a realidade, [...] que ele [...] possa dar conta do recurso público que ele recebeu pelo financiamento, ou por ser um professor da instituição Federal. Então, pra mim, tem outros elementos que, por falta de conhecimento, eu ainda fico meio sem entender como é que seria essa disponibilidade.

A colaboração em pesquisas também é um aspecto a ser discutido entre os pesquisadores da UFC. Para McKiernan *et al.* (2016), a colaboração é essencial para o avanço do conhecimento, entretanto, identificar e conectar-se com os colaboradores não é tarefa fácil. Para os autores, as práticas abertas facilitam a interconexão entre os cientistas, aumentam a visibilidade do pesquisador e de seus trabalhos, agilizando o acesso rápido aos dados e softwares, além de criar oportunidades de interação e de contribuição em projetos. A respeito dos desafios em relação à colaboração e compartilhamento nas pesquisas científicas, E5 indica a falta de interlocução e comunicação institucional:

[...] Eu acho que o maior desafio é você quebrar mesmo essas barreiras de castelo, sabe [...]? [...] Vamos pensar aqui na UFC: você tem o departamento de engenharia, departamento de administração, departamento de sistemas e mídias, a biologia, medicina, estão em prédios separados, um não fala com o outro, entendeu? Eu sempre falo assim para os meus alunos, eles vêm aqui ter aula comigo, eu pergunto: "Vocês conhecem a UFC? Vocês já foram no Pici?" Ele nunca pisou no Pici, entendeu? E o pessoal de biologia, de outras áreas, de física, sei lá, sistemas e mídias digitais, como eu vou lá, eu fiz essa parceria. São raríssimos os professores que fazem isso. Então, acho que esse é o maior desafio. [...] Ainda tem uma coisa individualista [...]. (E5).

Segundo a perspectiva de E5, há uma necessidade de maior integração na UFC para que todos os *campi* e cursos possam formar parcerias de trabalho e em pesquisas. Diversos pontos entre a pesquisa científica se cruzam ou estão imbrincados com a Ciência Aberta e o compartilhamento de dados, inclusive, alguns códigos foram recorrentes entre uma categoria e outra, evidenciando um elo e ao mesmo tempo uma linha tênue entre os temas, conforme o exposto ao longo dessa dissertação.

#### **4.3.4 Biblioteca**

Relacionando o papel dos bibliotecários aos dados de pesquisa, investigou-se junto aos entrevistados se eles consideram importante e se reconhecem a

biblioteca como um dos setores que pode desenvolver e ofertar serviços de dados e de apoio aos pesquisadores. Naturalmente, nas palavras de Sayão e Sales (2015, p. 6), a biblioteca é vista como “[...] a custodiante [...] dos registros acadêmicos, e este papel se estende agora para incluir os dados de pesquisa.”

Reforçando esse pensamento, para Packer (2005, p. 252), “A biblioteca é a instância que dá sustentação e democratiza o fluxo de informação [...]”, portanto, no que corresponde à questão dos dados de pesquisa não seria diferente. Nesse sentido, a maioria dos entrevistados concordou que a biblioteca tem relevante função a desempenhar no que se refere a oferta de tais serviços, como o auxílio na organização dos dados, inclusive em capacitações, cursos e treinamentos presenciais ou EaD. Na perspectiva dos entrevistados, a biblioteca pode ajudar na melhor organização das informações (E1, E4, E6, E9, E12), na capacitação, desenvolvimento de cursos e atendimento especializado em gestão de dados para a comunidade acadêmica (E1, E2, E3, E4, E5, E6, E8, E9, E10, E11, E12), na padronização e indicação dos metadados usados para cada área (E6). Conforme ilustra a fala de E8: “[...] quando eu digo que vocês podem ajudar, muitos nem sabem! Eu digo que vocês podem ser um suporte, que tem profissionais que podem ajudar. Os cursos que vocês divulgam [...]. É uma pós que está em sintonia com a biblioteca.”

Foi mencionado por E2, E3, E4, E6, E8 e E10 a necessidade de incorporar a gestão de dados de pesquisa aos cursos de graduação, pelo menos na iniciação ao preenchimento do PGD.

Talvez, se possível, cursos de capacitação, né? Um tipo de atendimento ao estudante que ainda não tem essas experiências porque muitas vezes a gente, principalmente, na graduação, aprende isso mais ao final, quando precisa lidar já com o trabalho em si, já com o documento, já com o *corpus*, já com a monografia, com o artigo, a gente acaba descobrindo isso bem ao final da graduação. [...] Acho que esse tipo de capacitação para o aluno, desde o começo, de entender como funciona, de saber fazer, de saber compartilhar. [...] Até para a pessoa entender a importância que isso traz para o trabalho dela. Isso traz assim, de certa forma, uma legitimidade que outros trabalhos não teriam. [...] Acho que esse tipo de mudança de mentalidade traz um benefício maior para o pesquisador [...], porque ele se torna uma pessoa mais disciplinada, mais atenta, mais cuidadosa. Então, traz um enriquecimento pessoal, de certa forma. (E2).

A Entrevistada 4 expressou que a demanda relacionada à gestão de dados na universidade é urgente. Sua percepção tem razão de ser, uma vez que a gestão de dados envolve uma série de novos processos que demandarão dos pesquisadores pelo menos o mínimo interesse para se apropriar de práticas como a documentação

da pesquisa, organização, preparação para tornar os dados FAIR, compartilhar, entre outras.

Por favor! Por favor! Isso é urgente! Isso é urgente! Isso é premente! Eu não sei que outra profissão poderia oferecer esse suporte já que vocês é que tem essa formação, então, não faz muito sentido formar outras pessoas para fazer o que vocês já fazem. Eu acho isso contraproducente. [...] **Então, até onde eu sei, são os profissionais de Ciência da Informação que tem que trabalhar com isso. E isso é urgente [...]. Antes que toda uma nova geração de pesquisadores se acabe sem que as pessoas mudem sua maneira de fazer ciência. E a gente mais uma vez perca o bonde da história porque os países desenvolvidos já estão lá, caminhando para isso, e a gente aqui fica nesse negócio fechado, estreito, restritivo, e essa restrição não beneficia o pesquisador.** Se beneficiasse, eu até entendia, não acho certo, não, mas, eu entendia. Mas, não beneficia, não. Só faz com que tudo caminhe mais devagar e o nosso trabalho seja menos valorizado, reconhecido. [...] Então, se a coisa abrir, [...] alcança o mundo. (E4).

E7 suscitou importantes reflexões a respeito do papel das bibliotecas, se seria ou não o setor competente para o oferecimento dos serviços de dados e de apoio ao pesquisador na UFC:

[...] A biblioteca, ela tá mudando muito o papel dela em comparação há alguns anos atrás, né? Eu acho que o papel dela, presencial, parece que tá ficando cada vez mais exíguo. E ela está migrando, talvez, daqui a alguns anos... A biblioteca como nós conhecemos, seja mais um repositório histórico das informações [...]. Nós estamos migrando para uma Era de Home Office, daqui a alguns anos as pessoas, talvez, não vão nem mais sair de casa, vai ser ensino a distância. Assim, eu acho que... Não sei se seria, no meu entendimento, papel da biblioteca, capacitar as pessoas. Não saberia dizer... [...] Não tô dizendo que seria contrário a isso, mas, não sei se faria parte do escopo da biblioteca fazer isso. [...] Teria que refletir mais, porque vocês já têm tantas atribuições. Será que esse seria papel? Assim, seria facultativo? Imperativo? Eu não sei. (E7).

Com base no questionamento, a literatura científica registra que os bibliotecários têm ocupado espaços e criado oportunidades na área da gestão de dados de pesquisa. (BURNETT, 2013; CORTI *et al.*, 2009; CUNHA, 2010; MORGAN-BRETT, 2011; PINFIELD, COX, SMITH, 2014; SAYÃO; SALES, 2014; SAYÃO; SALES, 2015; TENOPIR; BIRCH; ALLARD, 2012). O número de bibliotecas ao redor do mundo que oferecem serviços de dados e apoio ao pesquisador tem aumentado ao longo dos últimos anos. Diversos são os relatos de experiências que versam sobre as questões pragmáticas e operacionais ou ligados às competências que o bibliotecário deve ter para atuar nesse campo.

Reforçando a opinião dos pesquisadores sobre os benefícios de uma capacitação para ajudar na adoção das práticas da Ciência Aberta, no compartilhamento e no gerenciamento dos dados de pesquisa, E9 afirma que não há

nada melhor do que o interesse e a necessidade. A entrevistada 4 salienta que “As pessoas não adotam porque não conhecem. [...] Então, eu acho que não há a menor dúvida. **Muita gente não muda de atitude porque não sabe que existe outra atitude possível.**”

Após apresentar a perspectiva dos sujeitos de pesquisa nesta categoria e suas ligações ou semelhanças com a literatura científica, a próxima seção versará a respeito do compartilhamento.

#### **4.3.5 Compartilhamento**

O compartilhamento é um ponto em comum mencionado pela maioria dos entrevistados e corresponde inclusive à disponibilização e divulgação do trabalho em estágio de desenvolvimento. Houve unanimidade entre os pesquisadores em afirmarem que não teriam problema em compartilhar desde que entendessem de que forma isso seria feito, sem ferir princípios éticos e legais, e ainda em casos de pesquisas em desenvolvimento, sendo esse um ponto delicado porque fere a primazia e o ineditismo da pesquisa acadêmica. Inclusive, alguns sugeriram que compartilhariam, desde que antes explorassem os seus resultados de pesquisa, isto é, após um período de embargo.

Locupletando a discussão da problemática, E8 vai além e afirma que as regras relacionadas ao compartilhamento de dados deveria ser algo previsto pelas agências de fomento e que todas as disposições normativas devem constar nos editais de financiamento:

Eu acho que a primeira coisa é: [...] **teria que vir no edital de financiamento.** Certo? No edital de financiamento, teria que já ser dito, que aqueles dados, quando finalizar [a pesquisa], eles têm que estar abertos e que a agência que está financiando, ela respalda isso. E aí, eu, enquanto pesquisadora, eu tomo a decisão: se eu acho que é possível, eu me candidato ao edital, se não, então... Eu acho que essa questão, [...] o primeiro momento é de quem está financiando. Ter essa informação, se aquilo é um requisito para poder o projeto ser selecionado. [...] **Até porque aí entra também o sigilo da informação.** Eu tenho que passar pelo **comitê de ética**, quem tá respondendo tem que ser mantido o sigilo, né? **Eu não sei até que ponto, como é que esses dados são codificados.** Mas, assim... É uma coisa muito relativa, tá? Porque a pessoa, ela deu um dado dizendo que... Deu uma opinião sobre alguma coisa. **Eu trouxe uma informação e eu tô dizendo ali para ela que vai ser mantido o sigilo. Mas, se aquele dado, ele chega para qualquer pessoa, qualquer pessoa pode interpretar do jeito que quiser. Inclusive a pessoa pode até botar um nome para criatura. Porque se tá aberto, eu não tenho controle, né? Eu tenho um controle do que eu tô escrevendo, que não tem nenhuma identificação.** Mas, eu não sei quem pega e o que pode fazer com esse dado. E também assim, como é que ele

vai dizer que ele é... É obrigado a dizer que ele pegou de um banco aberto? O que que ele vai escrever? Por exemplo, de metodologia. Porque ele não participou. Ou ele vai pegar os artigos que eu escrevi dizendo como eu fiz a metodologia, então, como é que fica essa questão inclusive do ineditismo dos trabalhos? **Isso eu tô dizendo assim por falta de informação**, realmente, eu não tenho. **Para mim é um tema muito novo**, então, eu tô trazendo aqui vários questionamentos do que seria ter essa possibilidade desses dados abertos. Então, **o próprio comitê de ética, quando você passa um projeto de pesquisa lá eles não perguntam isso**. Pois é. Então assim, **eu acho que para isso, para essa possibilidade, eu acho que teria que ter elementos anteriores para se mudar uma cultura, para se amarrar, ter uma legislação mais amarrada em relação a esse acesso**. Se esse acesso obriga a pessoa mencionada aonde é, de qual é a fonte? (E8).

De acordo com um dos pontos abordados por E8 em relação às questões éticas da pesquisa, é mister salientar que até o próprio comitê de ética, inclusive via Plataforma Brasil, apresenta ao pesquisador, se assim desejar, manter o sigilo de seu projeto durante a etapa de execução do estudo até a sua finalização. Dando seguimento a linha de pensamento de E8, assim como E10, sobre o compartilhamento, foram mencionados receios e problemas relacionados ao plágio nas pesquisas se estenderem à esfera do uso e compartilhamento dos dados de pesquisa.

Argumentando o desenlace referente ao compartilhamento dos dados, E8 traz à tona importantes reflexões no campo da ética e integridade da pesquisa científica que estão intrinsecamente ligadas ao processo de pesquisa e as particularidades existentes nas pesquisas de campo em oposição às que são desenvolvidas de maneira distinta:

Então, acho que também tem essa cultura, do que vem se transformando. Muitas vezes, as pessoas, elas estão numa pressa de fazer as coisas com muita rapidez, então, muitas vezes elas **não passam por um processo de maturação, do que é construir um questionário, do que é fazer uma pesquisa, então, de repente, ela já pega um banco formado**. O que eu disponibilizo, meu banco, é para assim, os parceiros da pesquisa. Então, se eu faço uma pesquisa no Paraná e tem pessoas lá, esse banco, a gente disponibiliza. Por que? Porque é um parceiro daquela pesquisa. Então, para mim, essa questão ela ainda é uma questão que eu acho que precisa ter muito esclarecimento. Para que ela possa dar uma segurança e você se sentir, vamos dizer assim, **sem ferir os princípios éticos**. Para mim, essa questão mexe muito com a questão da ética. A ética da pesquisa. E não só da ética do ponto de vista formal, para mim, entra com processos. **Como minhas pesquisas são muito implicadas, elas são muito envolvidas, então, não é só um dado de pesquisa coletado, é um processo que foi construído. Então, disponibilizar isso, para mim complica, porque não vai ter esse processo de pesquisa**. Vai ter um dado ali nu e cru onde eu vou pegar. **Em outras áreas pode ser mais fácil porque é uma amostra, porque é um componente que eu fiz...** No caso da Psicologia, especificamente da Psicologia Comunitária, para mim, ainda é algo muito difícil de se apropriar. **Mas, assim, acho que é algo importante de se entender, até para poder ter um posicionamento. Eu tô tendo hoje um**

**posicionamento muito mais como leiga no tema. Mais uma sensibilidade do que uma conhecedora na temática. (E8).**

Além da preocupação ética com a pesquisa, os aspectos legais como a Lei Geral de Proteção de Dados, deve ser considerada no âmbito do compartilhamento dos dados científicos. Assim, E12 pontua que é favor do compartilhamento dos dados, desde que estes não contenham informações sensíveis, e, caso existam, que sejam adotados os contratos de confidencialidade ou mecanismos semelhantes.

Com o intento de investigar se os pesquisadores da UFC compartilham ou não seus dados de pesquisa e para tentar compreender como funciona esse processo, inquiriu-se cada um dos entrevistados sobre isso. A queixa de alguns pesquisadores apontou que no Brasil não há esse costume, pouquíssimas pessoas respondem a um pedido dessa natureza, ou em algumas ocasiões em que chegam a responder o indivíduo não possui mais os dados. Alguns pesquisadores relataram que no exterior isso é visto com maior naturalidade.

Não há problema nenhum [o compartilhamento entre membros do mesmo grupo]. **Geralmente, ou é porque às vezes perdem esses dados.** Uma vez eu entrei em contato com uma professora lá no Estados Unidos, o artigo era de 2009, não, 2013, ela disse: "Não, não tenho mais, não. Perdi". Perdeu tudo, né? Às vezes não é por má índole. A pessoa perde. Eu acho que muitos casos, o que a pessoa na minha área vai querer, às vezes, não quer passar o algoritmo, tem medo, mas, os dados ela quer passar porque **vai ganhar uma citação** com isso. E é o que na nossa área hoje, mais do que a publicação, **a citação é mais importante do que a própria publicação.** (E7).

A partir do exemplo mencionado por E7 acima, percebe-se que a gestão dos dados científicos tem sua relevância na salvaguarda dos arquivos, preservação a curto e em longo prazo, e, poderiam os dados terem sido compartilhados se a professora tivesse executado essa atividade. Afinal, uma vez que os arquivos foram perdidos, é praticamente impossível recuperá-los. A citação dos dados também tem sido uma das vantagens bastante destacadas na literatura científica sobre a gestão de dados de pesquisa (FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER, 2019; SALES, 2014; SALES; SAYÃO, 2018; SAYÃO; SILVA, 2019).

Entre os membros da comunidade científica da UFC, o compartilhamento entre grupos e parceiros de pesquisa parece ser o mais comum. Sobre a definição de 'comunidade científica', Le Coadic (2004, p. 28) afirma que "[...] é o grupo social formado por indivíduos cuja profissão é a pesquisa científica e tecnológica." Para o

referido autor, essa é uma noção ambígua, revestida do mito da ‘república das ideias’ (Cidade do Saber) surgida no século XIX. De acordo com esse mito, os cientistas se encontravam para discutir ideias e teorias, e tinham como preocupação em comum o compromisso com a verdade. O autor faz uma analogia ao mito e explica:

[...] existem as **comunidades científicas** reais, **segmentadas** em função de disciplinas, línguas, nações e mesmo ideologias políticas; comunidades de trabalhadores científicos motivados por forte espírito de competição, onde o cientista que **vence é quem primeiro publica a informação**. (LE COADIC, 2004, p. 29).

Alguns pesquisadores informaram que costumam compartilhar seus dados entre os grupos de estudo que lideram, entre alunos e colegas ou parceiros de pesquisa, mas, segundo os depoimentos, é cada um dentro de seu próprio espaço. O entrevistado 10 informou que em sua área há mais compartilhamento de dados entre os grupos, mas, alerta que ele conhece em maior quantidade pessoas em sua área (Ciências Biológicas) receosas em relação ao compartilhamento. Ainda sobre o impacto causado no processo de pesquisa por conta do compartilhamento dos dados científicos, E8 expõe suas preocupações:

Para mim, isso impacta muito, é o processo da pesquisa. Para mim, o dado, ele tá ali filtrado. [...] **Como é que eu posso falar de uma pesquisa que eu não participei?** Até porque todas as pesquisas que eu faço, eu vou pra campo. Então, eu tô lá pesquisando, [...] envolvida, estou lá no dia a dia. [...] Então, para mim é muito estranho eu receber um dado sem eu saber como ele foi construído. [...] O que me impacta muito é o processo de construção, que não tá naquele dado que está num banco de dados. Aquele processo, é só quem fez. E aí aquele dado, para mim, ele não tem vida. **Porque se ele não tem o processo, como ele foi construído, ele não tem vida.** Então, para mim é muito estranho esse compartilhar, porque para mim tem que ter essa vida, que foi como é que ele [...] foi gerado, como é que ele chegou ali, [...]? (E8).

Ao contrário de algumas áreas que declaram ter dificuldades em conseguir acessar ou disponibilizar dados científicos, áreas como a linguística de *corpus* tem o compartilhamento e a disponibilização como sua razão de ser, ou seja, se essas práticas não existissem, seria impossível que houvesse produção de conhecimento nesse referido campo do saber.

**Nesse sentido, a linguística de *corpus* é vanguardista**, dentro da área linguística. **Porque tem uma transparência muito grande, inclusive as ferramentas de análise, a maioria são gratuitas, porque são elaboradas pelos próprios pesquisadores, que vão estudar computação pra se tornar capaz de desenvolver essas ferramentas porquê de outra forma eles não conseguem. E aí disponibilizam também as ferramentas.** Então, quanto a isso, eu acho bem tranquilo. Eu, talvez esteja falando desse meu



lugar de conforto. Se eu gerasse dados que tivessem tido um custo alto pra ser gerados, eu talvez não me sentisse assim tão tranquila de dizer "disponibilizo". Mas, eu nunca passei por isso, então, eu não sei como é que eu ia me comportar, se eu tivesse um apego ao meu dado, que eu não tenho porque eu não tenho porque ter o que nem meu é. Tá disponível. [...] **Mas, eu continuo achando que os dados não importam, o que importa é a sua metodologia.** Se você conduz um experimento rigoroso, correto, você vai chegar em conclusões que são válidas, replicáveis, em qualquer circunstância, e, as conclusões vão ser sempre as mesmas se for replicado corretamente. E se for usado os mesmos dados com outra metodologia, vai se chegar a conclusões que serão complementares e não competitivas. (E4).

Apesar da farta disponibilidade de *corpora* textuais na Linguística, nem sempre esse fator representa colaboração. E2 e E4 informaram que os pesquisadores trabalham individualmente, tendo pouca participação em projetos de cooperação, o que é um contraponto interessante, uma vez que se tem compartilhamento e disponibilização de materiais (*corpus*), mas, em termos de colaboração, cooperação e trabalho em rede são incipientes, conforme o depoimento de E4:

Não. Todo mundo individual. Nem dentro do próprio estado, nem a UECE dialoga com a UFC, instituições da mesma cidade. Então, os projetos de cooperação são muito poucos. Existe diálogo, mas, é tudo muito incipiente [...]. Não existe uma [...] cultura forte de cooperação. Não existe, não tem, nem dentro da mesma cidade quanto mais internacionalmente. [...] Não existe uma atitude institucional de manutenção desses laços. Não existe uma política. É cultura mesmo. Não existe uma cultura de valorizar essas relações. São iniciativas individuais, que só beneficiam os currículos individuais das pessoas, por mais que elas queiram. Elas têm toda uma boa intenção, elas socializam, publicam, fazem eventos, divulgam. Dali não sai. Acaba. Parcerias assim internacionais fortes, duradouras, são muito poucas. Na minha área, não tem não.

A respeito da falta de uma cultura colaborativa mais sólida, conforme o relato de E4, destaca-se a necessidade de se estabelecer parcerias institucionais internacionais. Na ótica de Russel (2000, p. 37), "Uma das formas de compensar a pequena estatura científica das comunidades da periferia é colaborar com instituições dos países situados no centro científico. Gibbs (1995) e Meneghini (1998) já alertaram em seus estudos como o conhecimento produzido nos países que não fazem parte do *mainstream* da ciência são relegados a parte submersa e invisível do saber produzido e publicado nos periódicos científicos mundialmente.

Alguns dos entrevistados informaram que compartilhariam seus dados apenas se isso fosse solicitado, e, há pesquisadores que aceitam fazê-lo desde que saibam como serão usados ou mediante a imposição de algumas condições. Outro depoimento informa que alguns dados não são compartilhados porque são dados de empresas, e, muitas vezes, quando essas repassam os dados aos pesquisadores,

eles vêm embaralhados, para que não possam ser identificados. O entrevistado E3 relatou sua insatisfação e experiência vivenciada em relação à retenção de dados necessários para o desenvolvimento de sua pesquisa.

Eu tive problemas [...]. Por exemplo, a história das poligonais. Na geografia tem muita gente que trabalha com as poligonais, e que tinha as poligonais. Não compartilharam. Tem uma lógica perversa, muito ruim no caso da geografia. [...] A minha geração não teve muito acesso a fazer o recurso do mapa, porque é uma técnica, na realidade, essa geração tem, e aí quando eu voltei pro mestrado eu vi uma coisa que eu sempre combati que era o negócio de empresa júnior. Então, agora montou uma empresa júnior na geografia, então, ele até faz o mapa, se você pagar. Dentro da lógica da universidade, eu acho um absurdo. E... Mas, o que mais me impressionou foi essa história de compartilhar os dados. Dados simples, uma poligonal é uma coisa simples, e eu não tive isso lá dos colegas [...]. O que eu senti lá, especificamente, na geografia, nesse caso, é assim: tem compartilhamento se você é do laboratório [...]. Se você não for, não compartilham. (E3).

Posteriormente à coleta dessa informação, a pesquisadora conversou com E3 sobre este assunto, e, de acordo com o relato do sujeito pesquisado, ele conseguiu a colaboração de um colega da Geologia que não apenas montou as poligonais, mas, o ensinou como fazer.

Resumindo os desafios de compartilhamento elencados pelos pesquisadores, destacam-se os seguintes:

- a) Segurança e garantia da integridade dos dados a fim de prevenir mau uso, manipulação ou fabricação dos dados;
- b) Confiança de ter realizado um trabalho bem feito, bom o suficiente para ser avaliado abertamente;
- c) Cooperação e solidariedade entre os pesquisadores;
- d) Questões éticas e legais relacionadas à pesquisa;
- e) Reconhecimento;
- f) Monopólio das grandes editoras;
- g) Diálogo e interlocução institucional e entre os cursos;
- h) Primazia e ineditismo da pesquisa;
- i) Processos de pesquisa;
- j) Padronização e organização;
- k) Mecanismos de avaliação da qualidade da pesquisa;
- l) Produtivismo.

Os desafios supramencionados também correspondem a muitos que estão presentes na literatura científica e em pesquisas realizadas sobre a Ciência Aberta e

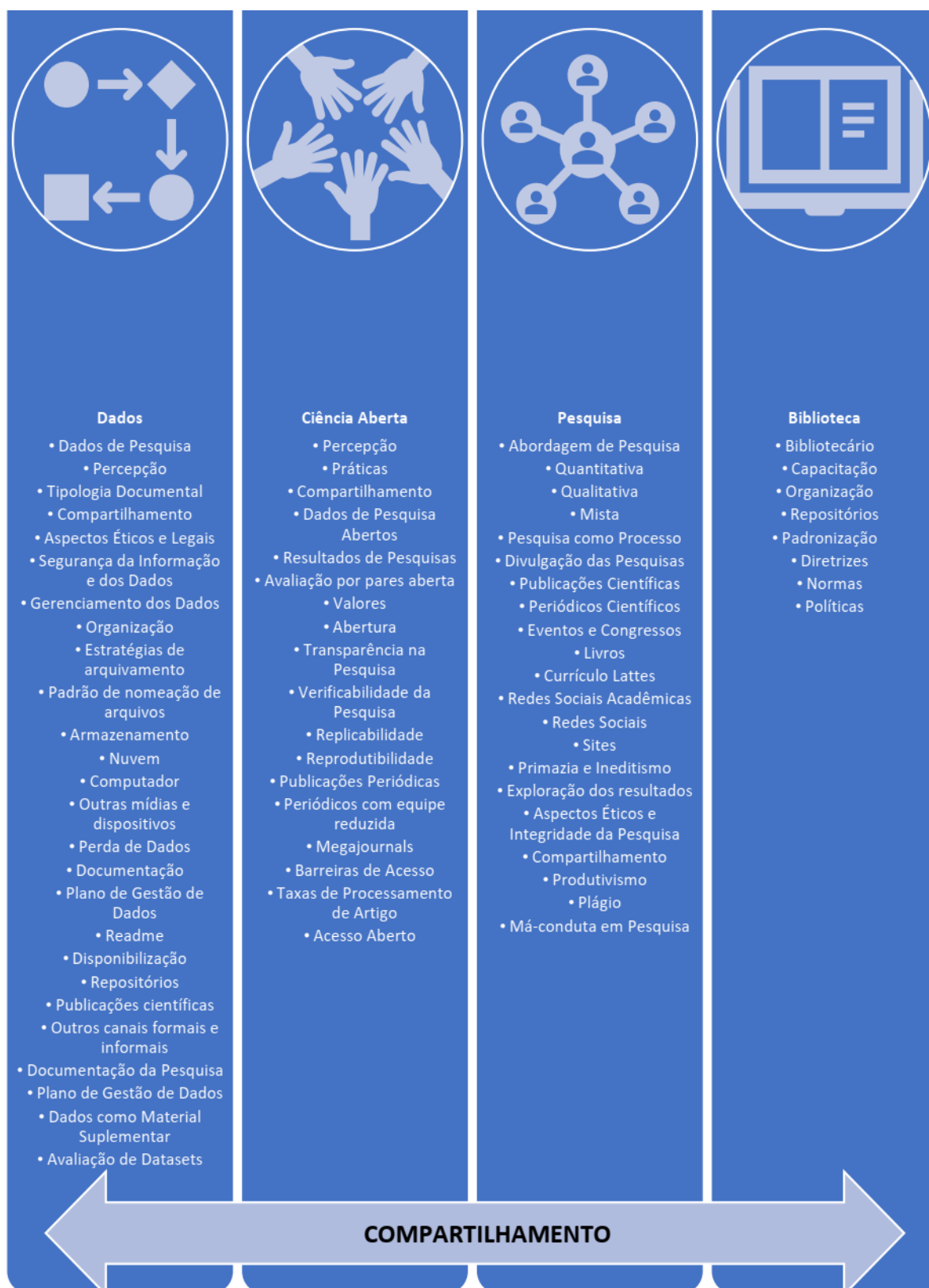
a Gestão de Dados de Pesquisa. (DIGITAL SCIENCE, 2017; REDE DE DADOS DE PESQUISA BRASILEIRA, 2019; SAYÃO; SALES, 2015; SILVA, 2019). Apesar disso, o compartilhamento tem conseguido maior destaque a cada dia, e o incentivo a essa prática vem crescendo, especialmente no atual cenário da pandemia do Coronavírus e síndromes respiratórias agudas (Covid-19, MERS e SARS).

Mediante o exposto, conclui-se que a investigação empreendida e o contato com os pesquisadores propiciaram uma ampla visão do panorama atual da UFC em relação ao tema. É evidente, fundamentado nos dados, que os pesquisadores da universidade já realizam em certa medida algum tipo de compartilhamento, seja de informações ou dados de pesquisa. Mas, o mais importante e perceptível ao longo desse processo de pesquisa e de diagnóstico da instituição é que eles desejam compartilhar, e, quando não o fazem, é porque desconhecem, ou não sabem como fazer isso, ou ainda por esbarrarem em questões éticas e de integridade da pesquisa. Nesse sentido, com o intento de contribuir e modificar a realidade, considera-se que a informação é conhecimento em ação (WERSIG, 1993), portanto, é o que nos conduz ao próximo passo executado durante a pesquisa: o minicurso online.

#### **4.4 Ação de Informação Educacional: Minicurso online**

A partir das coletas anteriores, a fim de se obter maior compreensão sobre o fenômeno em estudo, adotou-se como uma das propostas principais a elaboração e criação de um curso online para os pesquisadores da UFC com dois objetivos distintos: o primeiro, diz respeito à exigência de um aprofundamento e observação do comportamento, envolvimento e engajamento dos pesquisadores no ambiente virtual, convertendo-se em uma estratégia mais branda para observar e se aproximar da comunidade investigada. Assim, o emprego desse procedimento buscou comprovar os resultados obtidos durante a investigação, na primeira etapa referente ao questionário, a qual revelou algumas das práticas e necessidades informacionais dos pesquisadores. A análise da coleta desses dados foi útil para elencar os assuntos que iriam compor os conteúdos abordados no módulo introdutório a ser executado na terceira etapa da pesquisa. Os assuntos listados têm relação direta com a formação das categorias de análise desse estudo (Figura 30).

Figura 30 – Categorias e temas levantados para compor o módulo online do curso



Fonte: Dados da Pesquisa (2019).

O segundo objetivo relacionado à oferta do minicurso online na última fase da pesquisa visou a aplicação de uma ação de educação e informação para demonstrar a importância da gestão de dados de pesquisa para os pesquisadores, além de ser uma das maneiras empregadas para fazer o teste piloto de um dos cursos que poderá ser parte integrante de um programa de gestão de dados. Entre outros motivos para a realização do minicurso, destacam-se:

- a) Oferecimento de uma ação de informação / educacional sobre o tema para a comunidade, pois, evidenciou-se a necessidade pelo alto índice de pessoas que demonstraram desconhecer o tema;
- b) Observação participante e a aplicação do método netnográfico, uma vez que a presença física da investigadora poderia ser interpretada como invasiva e seria complicado intervir diretamente na rotina dos sujeitos investigados, dessa forma, a virtualidade da presença e nas comunicações suavizou o processo de entrada, deixou-os mais à vontade e estabeleceu uma relação horizontal entre todos os participantes;
- c) Teste dos conteúdos introdutórios do curso, identificação das falhas e ajustes a serem realizados antes da sugestão de implementação de qualquer proposta de ação de educação e informação por meio de cursos e treinamentos, ou em serviço de dados, suporte e apoio ao pesquisador.

Conforme já explicitado na seção de metodologia, o curso online usou a abordagem do design instrucional para o seu delineamento. Tal atividade, é definida por Filatro e Piconez (2004) como a atividade que se dedica ao planejamento, preparação, elaboração do projeto, produção e publicação de textos, imagens, gráficos, sons e movimentos, simulações, atividades e tarefas em suportes virtuais.

Desse modo, o modelo esboçado na etapa de planejamento considerou o *gap* existente a respeito do pouco conhecimento da temática por parte dos pesquisadores, o que influenciou também a escolha dos tópicos presentes no módulo introdutório. Apesar de o tema central ser a gestão de dados de pesquisa, outros assuntos que tem relação direta com os dados, como os temas da comunicação científica e a Ciência Aberta, também foram inseridos no minicurso online, por exemplo, o pré-print, as avaliações por pares aberta e as plataformas colaborativas online, a exemplo da Open Science Framework.

Além disso, por se tratar de um teste piloto do minicurso online, os sujeitos foram estimulados a dar *feedbacks*, fazer sugestões ou até contribuir com algum conteúdo se assim desejarem. O intuito dessa ação de informação funcionou muito mais como uma troca, em que tanto a pesquisadora quanto os sujeitos aprenderam e contribuíram uns com os outros. Ademais, a apropriação da informação ocorre no processo de mediação, quando se materializa em uma ação que permite produzir e não apenas consumir, isto é, pressupõe uma transformação do conhecimento do usuário onde este determina se a informação foi ou não realmente efetiva (ALMEIDA JÚNIOR, 2004, 2007; FARIAS, 2014).

Para a aplicação do minicurso online, conforme já explicitado, foi utilizada como Ambiente Virtual de Aprendizagem (AVA) a plataforma Google Classroom. Junqueira (2018), define um Ambiente Virtual de Aprendizagem como uma sala de aula online, um lugar na internet onde o aluno executa suas atividades e estudos. Geralmente, possui acesso via login e senha, agrega funcionalidades diversas de um site, inclusive conta com recursos de envio de mensagens, postagens em mural ou *timeline*, inspirado na dinâmica das redes sociais, chats, conteúdo das aulas (slides, vídeos e outros materiais), testes, agenda, calendário de atividades, fórum etc. Em suma, busca propiciar uma experiência do mundo real, isto é, presencialmente.

Inicialmente, a mensagem de boas-vindas foi elaborada e programada para publicação no Google Classroom (Figura 31), assim como os slides de orientação e ambientação na plataforma que constam as informações referentes ao uso e as ferramentas disponíveis, inclusive algumas regras de boa convivência e netiqueta (etiqueta em ambiente virtual). A proposta de atividades formativas do minicurso é opcional, mas, ainda assim solicitou-se a cada participante que se apresentasse, sendo essa uma das formas iniciais de interação entre a tutora e os cursistas.

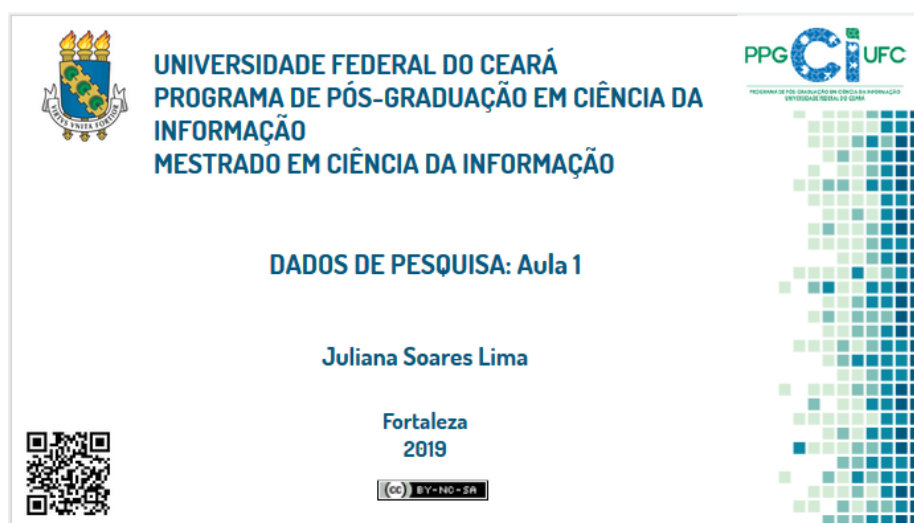
Figura 31 - Interface do Google Classroom



Fonte: Google Classroom (2019).

Conforme relatado na seção de metodologia, referente ao planejamento da ação de informação educacional, os materiais textuais criados para o minicurso (slides e notas de aula), receberam um QR Code (Figura 32) para acessar os materiais e ampliar a experiência imersiva na plataforma.

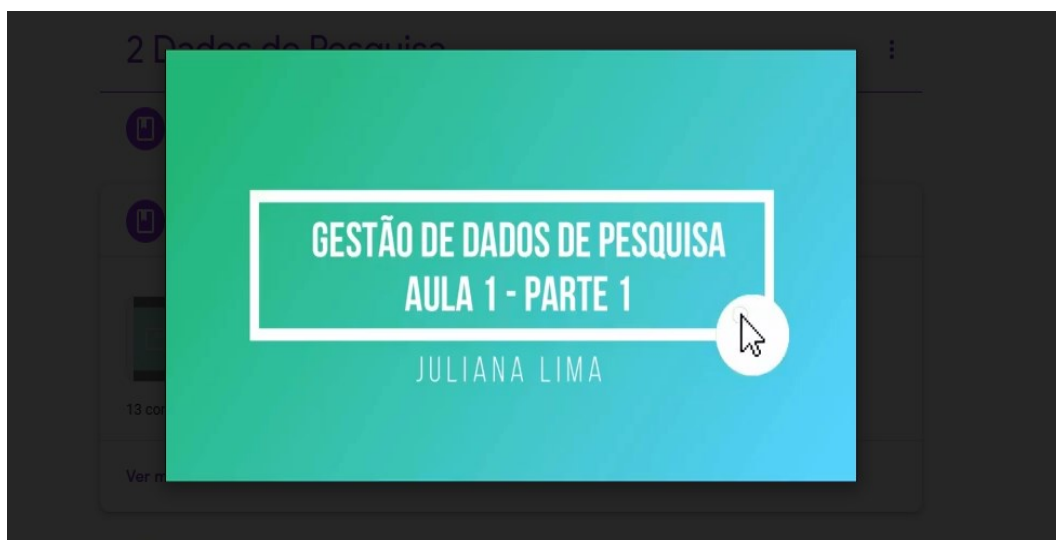
Figura 32 - Exemplo de slides com QR Code



Fonte: Lima (2019).

Foram preparadas quatro videoaulas com os conteúdos do minicurso, e para a última aula, foi gravado um podcast. Todos os vídeos eram acompanhados dos slides apresentados nas videoaulas (Figura 33). Outros conteúdos adicionais também foram disponibilizados a fim de complementar os materiais elaborados.

Figura 33 - Exemplo de videoaula

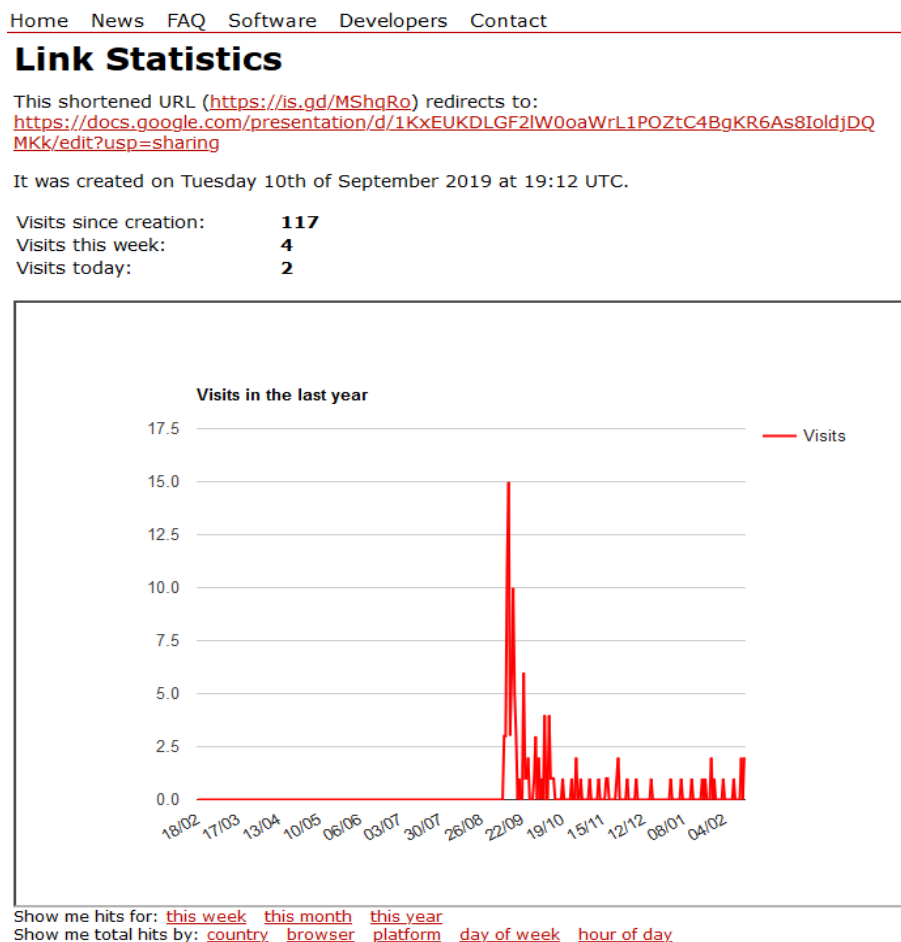


Fonte: Lima (2019).

Para acompanhar as estatísticas de acesso aos materiais criados para o minicurso, utilizou-se o encurtador de URL is.gd, pelo fato de ser uma ferramenta útil que disponibiliza gratuitamente a opção de personalizar o endereço encurtado, QR Code e as estatísticas básicas são visíveis para todos publicamente, conferindo assim maior transparência e rigor (Figura 34).



Figura 34 - Exemplo dos dados sobre estatísticas de acesso aos materiais elaborados



Fonte: is.gd (2019).

O número de downloads e acessos aos conteúdos, dias, meses e horários podem ser vistos por meio das estatísticas geradas no Is.gd. Geralmente, o número intenso de acessos ocorre bem próximo do período de disponibilização dos conteúdos no Google Classroom.

Em relação ao número de participantes, o curso teve 178 cursistas acompanhando o minicurso online. Apesar de ter tido uma boa interação entre os cursistas e entre a tutora, a maioria dos participantes permaneceu apenas assistindo as videoaulas, fazendo o download dos conteúdos, e, em algumas ocasiões, enviaram mensagens privadas para tirar dúvidas sobre o conteúdo. Mesmo assim, no mural de postagens, a interação foi boa e constante durante o minicurso, conforme ilustra o quadro 19 a seguir.

Quadro 19 - Número de interações em comentários nas postagens e mural de cada aula

<b>NÚMERO DE COMENTÁRIOS NAS POSTAGENS DE CADA AULA</b>	
<b>AULA</b>	<b>Nº DE COMENTÁRIOS</b>
Ambientação	3
Apresentação	101
Aula 1	53
Aula 2	69
Aula 3	42
Aula 4	14
Aula 5	31
Avaliação	4
<b>TOTAL</b>	<b>317</b>









Fonte: Dados da pesquisa (2019).

Conforme o quadro acima, a principal preferência da maioria dos cursistas foi interagir por meio do mural do curso para deixar seus comentários e dúvidas a respeito de cada aula. Segundo Junqueira (2018), um ambiente virtual de aprendizagem bem organizado facilita as interações entre as pessoas e objetos de conhecimento. Infere-se assim, que isso se deve pela facilidade, acessibilidade e disposição do recurso na interface da plataforma.

Após a postagem de cada conteúdo, os cursistas assistiam as videoaulas e acessavam o conteúdo no tempo livre e no seu próprio ritmo, e, depois registravam seus comentários ou dúvidas a respeito do conteúdo. No minicurso, essa facilidade de acesso em qualquer lugar ou hora, e em qualquer dispositivo, é descrita por Filatro e Piconez (2004) como uma das características da educação online, por meio da mediação tecnológica através da conexão em rede, sem limitação de tempo ou lugar.

Assim, por meio da mediação tecnológica, o processo interativo ao longo do curso ocorreu de forma satisfatória, como mostram as interações assíncronas via mural no quadro 20. Alguns dos participantes do minicurso online também foram entrevistados durante a fase de coleta de dados, como é o caso de E7 (Engenharia de Produção) e E6 (Odontologia).

Quadro 20 - Exemplo de interação entre os participantes - Aula 1

INTERAÇÃO ENTRE OS PARTICIPANTES – AULA 1: DADOS DE PESQUISA (MURAL)		
CURSISTA	DATA	COMENTÁRIO
 E7	12/09/2019	A questão do planejamento da pesquisa para evitar a obsolescência dos dados é fundamental em áreas relacionadas com tecnologia. A reprodutibilidade de um experimento, com base nos dados de pesquisa, é de suma relevância e tem de ser garantido. Creio que disponibilizar [sic] material suplementar, seja com os dados primários, bem como com uma análise estendida[sic] dos resultados, pode ser interessante nesse contexto.
 Tutora	12/09/2019	+E7, comentário perfeito! Há muitas análises assim em sua área, Professor?
 E7	16/09/2019	Olá Juliana, sobre seu comentário do dia 12 de setembro (desculpe a demora em responder), em minha área a necessidade de comparar experimentos computacionais realizados em hardwares distintos (e em linguagens de programação distintas) é uma realidade. Assim, a disponibilização de material suplementar poderia auxiliar a facilitar/viabilizar tais comparações. Dentre esses tipos de materiais, podemos citar código fonte, tabelas com os resultados detalhados (nos artigos as tabelas e gráficos são resumidos), dentre outros.
 E7	16/09/2019	A obrigatoriedade/tendência de publicação em revistas de acesso aberto (geralmente com taxas entre 700 dólares e 2000 dólares) será um fator de impedância para os pesquisadores de países ditos "emergentes" conseguirem publicar nesses periódicos, bem como conseguir fomento de agências nacionais e internacionais. Por trás de um argumento nobre, vem uma política para dificultar o desenvolvimento da nossa ciência.
 Tutora	17/09/2019	+E7, sem problema, não se preocupe com a demora na resposta. O mais importante é a participação de vocês. Seu relato sobre os tipos de dados e materiais produzidos em sua área mostram a variedade e formatos dos dados produzidos, muito bom! Além disso, concordo com você sobre a possibilidade de disponibilizar materiais suplementares, pois, isso ajuda o leitor na compreensão do que foi apresentado em um manuscrito, amplia as possibilidades. Abraço!
 Cursista 1	20/09/2019	O material foi muito bem elaborado. E esta é uma grande problemática, trabalhar com banco de dados, gerir, analisar, todos nós nas diversas áreas temos muitos dados e não sabemos como interpretá-los com eficiência. Onde fica a Datamining [sic]...
 E6	23/09/2019	Concordo com a Cursista 1 sobre o material didático que é muito bom! De fato, fica muitos dados na pesquisa que não são analisados até porque temos que selecionar alguns dados mais impactantes visando à publicação, restando quantitativos de dados que se perdem e não publicados. Disponibilizar esse banco de dados num repositório permitirá que outros pesquisadores tenham interesse de aproveitá-los e publicá-los.
 Tutora	24/09/2019	+Cursista 1, de fato é uma grande problemática. [...]. +E6, muito bem observado. De fato, nem todos os dados das pesquisas são utilizados, e, o compartilhamento propicia que novas pesquisas possam ser desenvolvidas. Abraços!



Fonte: Dados da pesquisa (2019).

Conforme o exposto no quadro anterior, E7 discutiu a respeito da reprodutibilidade de um experimento, característica de sua área de atuação (Engenharia de Produção). Adicionou ainda a essa questão que a disponibilização dos dados como material suplementar pode ajudar a outros pesquisadores em futuras comparações e desenvolvimento de novos modelos, pois, assim como E7 relatou em ocasião da entrevista, sua área se baseia em constantes comparações e busca desenvolver novos modelos superiores aos antecessores com bastante frequência.


Por sua vez, E6 comentou sobre o fato de muitos dados que não são analisados serem perdidos e não publicados, enquanto os dados que foram selecionados e utilizados nas pesquisas estão materializados apenas no formato do artigo e que nem sempre estão acessíveis para o público e disponíveis para uso e reuso. O comentário de E7 já demonstra uma mudança de pensamento, e reconhece que disponibilizar os dados em repositórios permitirá que outros possam aproveitá-los no desenvolvimento de novas pesquisas, conforme amplamente defendido na literatura científica (BÉCARD et al., 2016; BORGMAN, 2007, 2010; DIGITAL SCIENCE; FIGSHARE, 2017; EUROPEAN COMMISSION, 2016; FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER; LEHMKUHL; GAVRON; ALVES; MORAES DE BEM, 2018; LIMA; FARIAS, 2019; ROYAL SOCIETY OF LONDON, 2012; SALES; SAYÃO, 2019; SAYÃO; SALES, 2015).

Outra ocasião em que o despertar para a questão dos dados de pesquisa na universidade é premente surgiu em diversas situações no minicurso. Entre os cursistas que se manifestaram, todos destacaram passaram a perceber a relevância em aderir a tal prática e as benesses trazidas para as suas pesquisas, conferindo inclusive maior rigor metodológico. Semelhante caso é registrado nas conclusões do trabalho de Sales (2014) em sua pesquisa desenvolvida no Instituto de Engenharia Nuclear (IEN). A pesquisadora destaca a mudança de comportamento ocorrida entre os pesquisadores e os seus cuidados que passaram a empregar em seus dados de pesquisa. A figura 35 apresenta algumas dessas mudanças de comportamento registradas via Google Classroom já no início do minicurso online.


Figura 35 – Exemplo de interação entre os participantes no mural


 **Juliana Lima** postou um novo material: 2.1.2 Gestão de Dados de Pesquisa 



Item postado em 9 de set. de 2019 Editado às 17 de nov. de 2019  
 Vídeo de 1 minuto sobre a Gestão de Dados de Pesquisa.



 **Gestão de Dados de Pes...**  
<https://is.gd/gdados1minpro...>



**9 comentários para a turma**



 **E7** 16 de set. de 2019  
 Estou mais do que convencido de que vinha dando menos importância do que deveria aos dados de pesquisa. Vou mudar de postura!!


 **C2** 16 de set. de 2019  
 Tenho a mesma sensação! Curso muito necessário!

 **C3** 17 de set. de 2019  
 Estou com o mesmo sentimento dos colegas  A luz acendeu para a questão dos dados! Maravilha, Juliana!!!! 🍌🍌🍌

 **Juliana Lima** 17 de set. de 2019  
 + , que legal! Ótimo!

 **Juliana Lima** 17 de set. de 2019  
 + , gratidão! Quero ver você e todos os pesquisadores da UFC no topo, com muito sucesso! Abraços! 😊

 **Juliana Lima** 17 de set. de 2019  
 + , que maravilha! Estou à disposição para ajudar todos vocês no que eu puder e for preciso em relação à gestão dos dados de pesquisa. Abraços!

 **C1** 20 de set. de 2019  
 Estou bastante entusiasmada com o conteúdo relevante do curso até na elaboração das minhas hipóteses.

Fonte: Dados da pesquisa (2019).

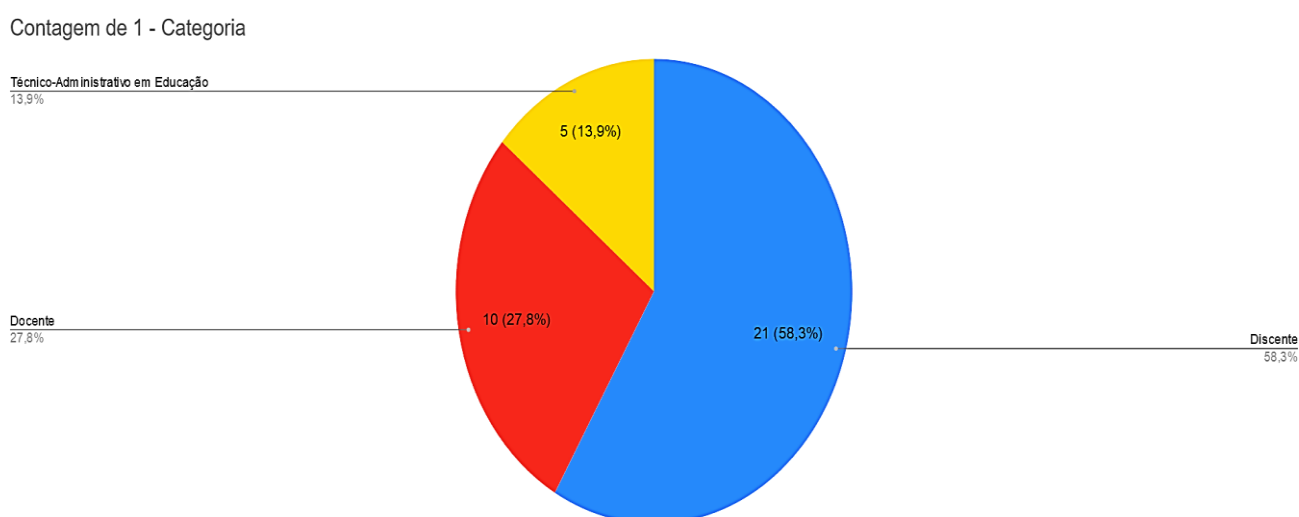
Perante os resultados iniciais demonstrados pelos comentários, as percepções dos pesquisadores assinalam que, a partir do momento em que se percebe a necessidade de informação e se identificam as lacunas de conhecimento, os sujeitos são capazes de identificar as suas próprias necessidades informacionais e buscam maneiras de preenchê-las. (KUHLETHAU, 1993; MIRANDA, 2006). Logo, os cursistas compreenderam que a sua necessidade de informação em relação aos dados poderia ser satisfeita por meio das informações e conhecimentos repassados durante o minicurso e contribuíram para melhor aprendizagem.

#### 4.4.1 Avaliação do Minicurso Online

Após o período de vigência estipulado para o curso online (30 dias), o momento final correspondente à avaliação por parte dos cursistas foi iniciado. Nessa etapa, alguns pedidos foram feitos para que o curso não fosse retirado do ar logo após o seu encerramento, assim como algumas perguntas relacionadas ao material também foram dirigidas para a tutora. Diante dos pedidos, o curso ficou disponível até 01 de novembro de 2019. Sobre os materiais, foi esclarecido que todo o conteúdo escrito (slides e notas de aula) estaria disponível para download e as videoaulas disponíveis no canal do YouTube da Biblioteca de Ciências Humanas da UFC.

Apesar de ter atingido o número de 242 inscritos, 178 cursistas estavam ativos na plataforma do curso. Desses 178, apenas 37 responderam a avaliação do minicurso online. A maioria dos inscritos foram discentes (170 inscrições), seguidos dos docentes (57 inscrições) e técnico-administrativos em educação (14 inscrições). Entre aqueles que responderam a avaliação do curso (Apêndice D), temos os seguintes percentuais: 58,3% são discentes; 27,8% docentes e 13,9% Técnicos-Administrativos em Educação. (Gráfico 21).

Gráfico 21 - Categoria dos cursistas que avaliaram o minicurso



Fonte: Dados da pesquisa (2019).

Provavelmente, o baixo número de respondentes tenha sido devido ao fato de não ser uma exigência para a obtenção de uma declaração de participação no

minicurso e por ter sido mencionado no início do curso que as atividades seriam opcionais. De qualquer maneira, apesar do baixo número de respondentes, é possível ter uma noção das impressões em relação ao minicurso.

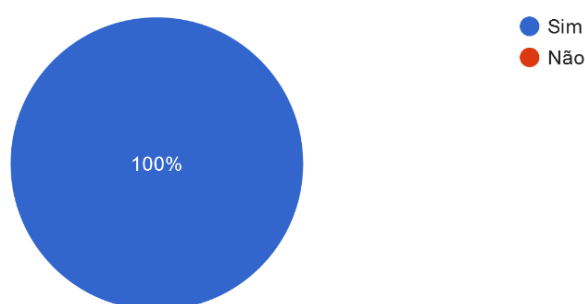
O maior número de respondentes da avaliação do curso foi dos cursos de Pós-Graduação em Ciências Biológicas (8,3%, entre docentes e discentes), do Doutorado em Ciências Marinhas (5,6%, docentes e discentes) e Pós-Graduação em Direito (5,6%, discentes). Os demais cursos de Pós-Graduação tiveram percentuais semelhantes (2,8%).

Em relação aos conteúdos ministrados e o conhecimento repassados, 100% dos participantes informaram que o curso contribuiu para a sua aprendizagem. Infere-se que a aprendizagem foi significativa, tendo em vista que fez parte do contexto, da situação e do ambiente (LE COADIC, 1996) dos cursistas, portanto, considera-se que o primeiro passo para o desenvolvimento das competências necessárias para lidar com os dados advém desse reconhecimento por parte dos sujeitos que passaram a entender as suas necessidades de informação e em como supri-la.

Gráfico 22 - Contribuição do curso para a aprendizagem dos cursistas

3 - Você considera que os conteúdos ministrados e os conhecimentos adquiridos contribuíram para a sua aprendizagem?

37 respostas



Fonte: Dados da pesquisa (2019).

Para investigar se os cursistas se sentem preparados para aplicar os conhecimentos adquiridos em sua prática cotidiana, 10,8% se sentem pouco preparados; 45,9% se consideram regularmente preparados; 29,7% preparados e 13,5% muito preparados. Em relação ao nível geral de satisfação sobre o curso, em uma escala de 0 a 10, a média de satisfação dos respondentes é 9 (nove). Esses dados estão sintetizados na tabela 9.

Tabela 5 - Média de satisfação e nota do curso

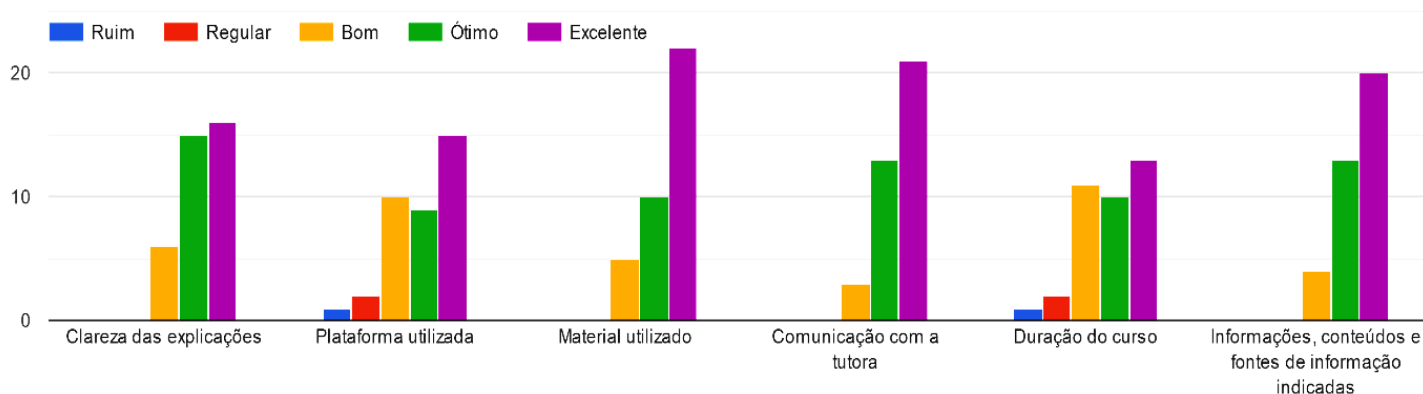
<b>Categoria</b>	MÉDIA de 4 - Atribua uma nota de acordo com a escala linear sobre o quanto você se sente preparado(a) para aplicar os conhecimentos adquiridos:	MÉDIA de 6 - Atribua uma nota para o curso de acordo com o seu nível geral de satisfação.
Discente	5	10
<b>Discente Total</b>	<b>4</b>	<b>9</b>
Docente	3	7
<b>Docente Total</b>	<b>3</b>	<b>8,5</b>
Técnico-Administrativo em Educação	4	9
<b>Técnico-Administrativo em Educação Total</b>	<b>3</b>	<b>9</b>
<b>Total geral</b>	<b>3</b>	<b>9</b>

Fonte: Dados da pesquisa (2019).

A satisfação em relação a itens como clareza das explicações, plataforma utilizada, material disponibilizado, comunicação com a tutora, duração do curso e informações, conteúdos e fontes de informação extras também demonstrou bom índice de aceitação.

Gráfico 23 - Grau de satisfação com itens gerais do curso

5 - Informe seu grau de satisfação com os itens abaixo:



Fonte: Dados da pesquisa (2019).

Além da avaliação do minicurso ter sido solicitada aos participantes, disponibilizou-se uma lousa virtual (Jamboard) para que os alunos pudessem se expressar e deixar uma mensagem a respeito do curso (Figura 36).



Figura 36 - Jamboard (Lousa Virtual)



Fonte: Dados da pesquisa (2019).

Por esse prisma, é possível inferir que a capacitação dos pesquisadores foi relevante e significativa para eles à medida que a realização do curso e até de determinadas atividades profissionais os auxilia na capacidade de compreensão do mundo, de lidar com os seus problemas e adquirir novas competências (BELUZZO,

2002). Não é por acaso que o reconhecimento das necessidades de informação perpassa pela identificação de lacunas no conhecimento, pois estas são geradas a partir de fatores emocionais e situacionais que reverberam em nosso cotidiano. Aliás, quando se trata de competência, “[...] já se reconhece a necessidade de aprender ou valorizar diversas áreas do saber.” (SILVA *et al.*, 2019, não paginado, locations 75%).

Terminado o período de avaliação do curso pelos participantes, passou-se para a avaliação dos aspectos técnicos do minicurso concernentes aos dados coletados no questionário de avaliação, e, procedendo assim a uma avaliação enquanto tutora do minicurso proposto. Para tanto, inspirada em Filatro *et al.* (2019), usou-se o modelo da matriz de Feedback (Figura 37) a fim de identificar o que pode ser melhorado, aprimorado e excluído em relação a estrutura do curso, modelo e estratégias.

Figura 37 - Matriz de Feedback

<p><b>+</b> o que funcionou?</p> <ul style="list-style-type: none"> <li>- Interações via mural do Google Classroom;</li> <li>- Videoaulas;</li> <li>- Podcast;</li> <li>- Slides;</li> <li>- Notas de aula.</li> </ul>	<p><b>▲</b> o que pode ser aperfeiçoado</p> <ul style="list-style-type: none"> <li>- Exploração de mais recursos e integração dentro do Google Classroom;</li> <li>- Tamanho das Videoaulas (Diminuir);</li> <li>- Tempo de duração e período de disponibilização do curso;</li> <li>- Explorar mais a mídia Podcast;</li> <li>- Plataforma Google Classroom.</li> </ul>
<p><b>?</b> questionamentos</p> <ul style="list-style-type: none"> <li>- Quais são os motivos para que as interações via chat/hangout com a turma não tenham funcionado?</li> <li>- Quais são os motivos de os participantes terem preferência pela mensagem privada para o tutor via plataforma (Google Classroom) ou por e-mail?</li> </ul>	<p><b>✓</b> novas ideias</p> <ul style="list-style-type: none"> <li>- Inserção de quizzes, jogos e exercícios;</li> <li>- Sugestões de atividades formativas (pelo menos uma obrigatória);</li> <li>- Sugestão de uma atividade colaborativa (em grupo);</li> <li>- Incentivar a construção de conteúdo colaborativa (prossumidores);</li> <li>- Criação de uma Persona.</li> </ul>

Fonte: Adaptado de Filatro *et al.* (2019).

Como as estratégias que funcionaram bem durante o minicurso, destacam-se as interações via mural do Google Classroom, as videoaulas, o podcast e os materiais textuais (slides e notas). Os feedbacks dos participantes em relação às notas de aula e os slides foi bastante positivo e foram elogiados pelos cursistas. As videoaulas

também tiveram retornos positivos, entretanto, o tamanho das aulas ainda precisa ser diminuído, considerando o comentário de um dos participantes no questionário de avaliação do curso: “Gostei muito do curso. As videoaulas foram ótimas: como sugestão poderia diminuir o tamanho das videoaulas e dividir o conteúdo em maior quantidade de vídeos”. Outro ponto a ser repensado é o tempo de duração do curso e adotar a sugestão de alguns cursistas sobre deixar a plataforma sempre aberta, sem limitação de período para que as pessoas possam acompanhar de acordo com o seu próprio ritmo. Em cursos online sobre Ciência Aberta e Dados de Pesquisa, por exemplo, essa foi uma estratégia usada pelo projeto FOSTER para que os pesquisadores interessados pudessem se capacitar.

Durante o curso, a exploração no uso de mídia como o podcast foi reduzida: somente um podcast foi gravado em caráter experimental e disponibilizado. Portanto, há a pretensão de explorar mais esse tipo de mídia em cursos futuros produzidos por esta pesquisadora e pela equipe da Biblioteca de Ciências Humanas, seu local de trabalho atualmente.

Sobre a plataforma Google Classroom, a interface é bastante intuitiva e constantemente passa por melhorias implantadas pelo Google. Entretanto, nem sempre as mudanças aplicadas em um site ou em ambientes virtuais de aprendizagem funcionam como o esperado, e, durante a realização do curso, o Google implantou uma mudança de layout semelhante a linha do tempo de redes sociais como o Facebook, onde as publicações recentes vão ficando acima e as mais antigas abaixo, mesmo com a hierarquização dos tópicos e conteúdos feitos pela tutora. Em certos momentos, isso causou uma impressão visual ruim, de um ambiente fora de ordem, e demandou que a tutora arrastasse e reordenasse os conteúdos constantemente para não comprometer a experiência de aprendizagem dos cursistas, mas, apesar de todos os cuidados, ainda foi perceptível para um dos participantes:

Ótimo curso, com diversos exemplos e diversas fontes de dados, mas creio que é muita informação para apenas um curso, além do fato de ser informação nova, o que torna o aprendizado uma tarefa talvez não tão fácil a todos. **A plataforma também não é tão organizada** no sentido de nem sempre ser possível saber onde a pessoa parou da última vez. Porém, tirando esses detalhes que são mínimos em comparação com a qualidade do curso, achei tudo perfeito e o aprendizado daqui espero conseguir levar adiante. (CURSISTA 4).

Dando prosseguimento a matriz de feedback, alguns questionamentos foram levantados sobre a participação dos cursistas em chats e Hangouts durante o

minicurso. Praticamente quase nenhum demonstrou interesse em se comunicar por esses canais, mas, pelo que foi possível captar instintivamente, isso se deve ao fato de que a maioria das pessoas estavam bastante assoberbadas com as tarefas do semestre e usavam o pouco tempo que tinham para acessar o curso. Não é à toa que a interação assíncrona foi a mais utilizada por meio do mural da plataforma, pois é o recurso de mensagem que está posicionado assim que se entra na plataforma e localizada logo abaixo das postagens dos conteúdos.

No último ponto elencado na matriz de feedback, foram registradas as ideias que se deseja aplicar futuramente na elaboração de novos cursos, como jogos, quiz, exercícios, atividades colaborativas em grupo, incentivar a criação de conteúdo por parte dos cursistas (prossumidores<sup>152</sup>) e a criação de uma persona. Filatro *et al.* (2019) define como persona um personagem fictício criado para representar diferentes tipos de usuário, inclusive usando as mesmas características, atitudes e comportamento do aprendiz que participa do curso.

Além disso, vale ressaltar que já se sabe que alguns pesquisadores podem ter receio ou medo de compartilhar seus dados, mas, o curso foi totalmente dirigido para desmistificar isso. Assim, foi preciso mostrar aos pesquisadores que o uso da informação e do conhecimento não faz com que se acabem ou sejam consumidos (BORGES, 2000). O uso e reuso dos dados é possível, além de ser capaz de gerar novos conhecimentos a partir de seus antecessores.

Mais uma das preocupações iniciais foi exatamente mostrar que nem todo tipo de dado de fato pode ser compartilhado, considerando-se as questões éticas e legais. Percebeu-se que após o acesso às informações disponibilizadas no conteúdo do minicurso, os pesquisadores passaram a olhar a questão do compartilhamento de forma mais apurada, crítica, inclusive se autoavaliando e identificando-se com o tema.

Alguns deles declararam na plataforma do curso que após a experiência e o conhecimento obtido no minicurso, passariam a documentar suas pesquisas e a compartilhar os seus dados de pesquisa que não infringissem nenhuma norma, legislação ou ética da pesquisa. Essa mudança de comportamento demonstra que se antes os pesquisadores não documentavam ou compartilhavam dados, isso se deve ao fato de que não tinham conhecimento do que poderiam fazer, por que e de que forma isso poderia ser benéfico para a sua pesquisa. Portanto, a partir da

---

<sup>152</sup> Termo cunhado por Alvin Toffler, autor da obra intitulada "A Terceira Onda". Significa o papel do consumidor na Sociedade da Informação, que consome e também produz conteúdo.

sensibilização e do conhecimento adquiridos, os pesquisadores se tornaram conscientes de suas atitudes e responsabilidades, passando a enxergar os benefícios da adesão à essas práticas.

Nesse sentido, uma das estratégias utilizadas para demonstrar aos pesquisadores a importância da documentação da pesquisa e do compartilhamento dos dados de pesquisa foi a apresentação de um tipo de publicação científica bastante recente: o *data paper* (artigo de dados). Foi interessante notar que alguns pesquisadores que se demonstraram mais céticos em relação ao compartilhamento logo se interessaram em compartilhar seus dados por meio de um *data paper* em um *data journal*.

Apesar de ser um formato totalmente diferente de um artigo de periódico, pode-se inferir que o *data paper* demonstrou ser bastante atrativo para os pesquisadores. Em certa medida, eles ressignificaram a prática da escrita e publicação de um artigo, reconhecendo-se em práticas do que eles já fazem, contudo, com a adição de mais alguns passos, exigências, documentação da pesquisa e compartilhamento dos dados que sustentam os argumentos do artigo escrito. Portanto, essa predileção coaduna com as pesquisas que apontam o periódico como a fonte preferida pelos pesquisadores, utilizada tanto como insumo para o processo de produção do conhecimento quanto como um meio de publicação para veicular seus resultados de pesquisas (ALVARENGA, 2000).

Mesmo que a elaboração de um *data paper* seja distinta de um artigo tradicional, torna-se mais um incentivo e estratégia para que todos os pesquisadores façam bem o que já estão acostumados: compartilhar conhecimento. Escrever e publicar um artigo ou outro material é literalmente uma forma de compartilhamento.

Em menor ou maior grau, a sutil diferença é que além do artigo científico, agora os pesquisadores precisam também documentar suas pesquisas, gerenciar e compartilhar os seus dados, a fim de contribuir para o avanço da ciência e com a criação de novos estudos que podem ser desenvolvidos por meio do acesso aos dados. Acredita-se que explorar esse e outros formatos de publicação das pesquisas será um forte atrativo para aqueles que desejam iniciar no universo do compartilhamento de dados. Mais do que isso, publicar dados e *data papers* gera citação para os pesquisadores que participam da economia da dívida, na qual o conhecimento científico é um presente; portanto, deve ser compartilhado e estar

disponível para as pessoas como um bem público, isto é, sua recompensa é o reconhecimento (MAUSS, 2003).

Finalmente, só resta apresentar, entre tantos comentários deixados pelos cursistas, aquele que melhor representou uma aprendizagem profunda (*deep learning*) do conteúdo do curso online:

Empregando uma analogia apresentada no início do curso, a publicação é apenas o estágio final de uma investigação (a ponta do iceberg), enquanto os dados de pesquisa são o alicerce do iceberg, estando submersos e não vistos pelo público geral. Com o curso despertei a atenção para a importância de gerenciar os dados de pesquisa, com vistas a melhorar o processo de investigação como um todo, garantir a reprodutibilidade dos experimentos conduzidos, bem como melhorar a repercussão das publicações porventura produzidos. Confesso que não tinha essa sensibilidade no início do curso e hoje com certeza tenho ciência dessas questões. Muito obrigado! (E7).

O comentário acima expressa sobremaneira uma modificação nas estruturas de conhecimento de E7, materializando-se em uma aprendizagem não de superfície, mas sim profunda: aquela que ocorre quando são usadas as capacidades cognitivas de alta ordem (análise e síntese, julgamento e interpretação), integram a aprendizagem a conhecimentos e experiências anteriores, possuem uma visão holística do contexto, conseguem dar sentido e atribuir significado ao que está sendo aprendido, enxergam a aprendizagem sob a ótica de sua relevância para a vida e a carreira, demonstram retenção em longo prazo e são capazes de transferir aquilo que foi aprendido (FILATRO *et al.* 2019; LITTO; FORMIGA 2009).

Finalizando essa seção, reitera-se que a aplicação do curso comprovou que por meio da sensibilização e da promoção de ações educacionais e informacionais para os usuários contribuiu-se para a mudança no estado de conhecimento dos indivíduos, além da mudança comportamental e a adesão às práticas da gestão de dados de pesquisa. Essa experiência de aplicação do minicurso para os pesquisadores da UFC forneceu elementos para a criação de uma proposta de programa para gestão de dados de pesquisa, a qual será apresentada na próxima seção.

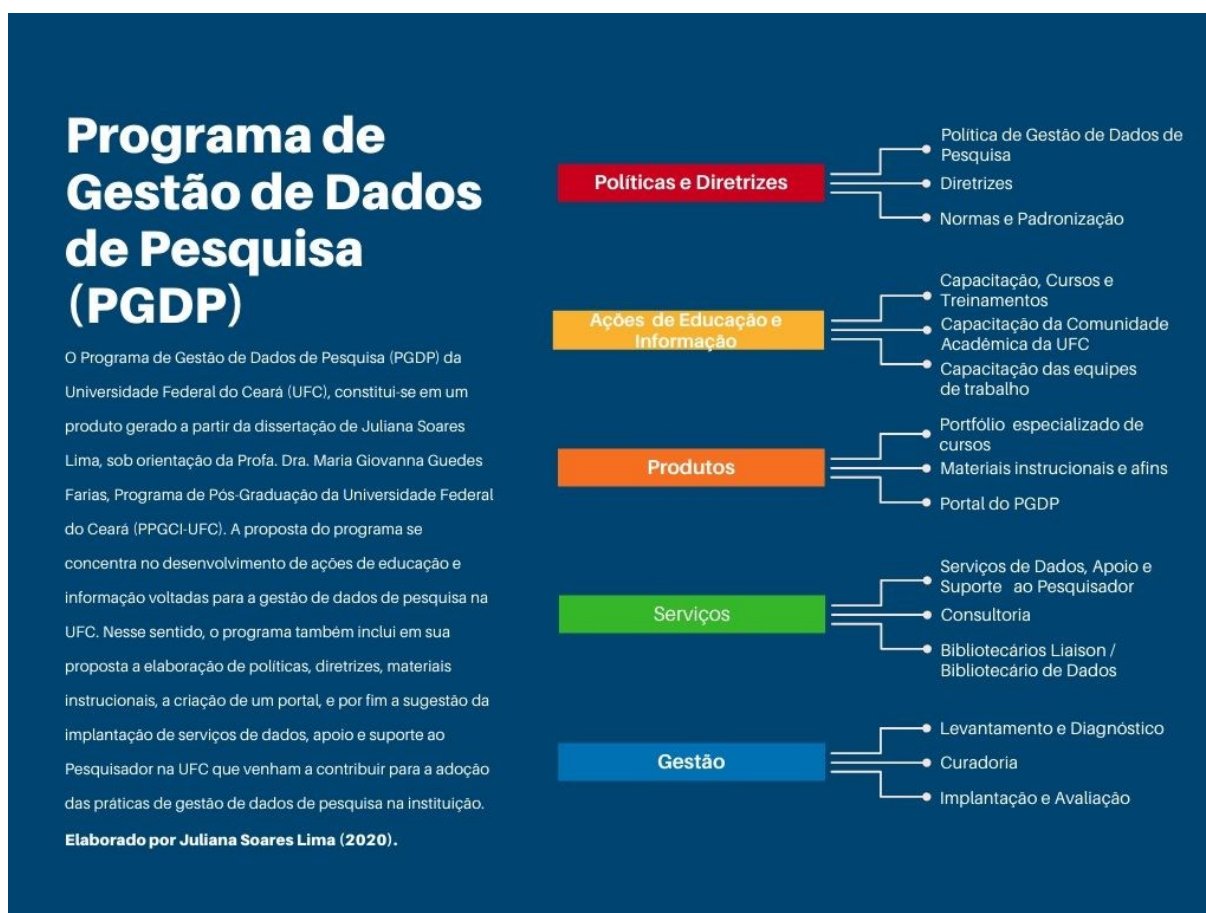
#### **4.4.2 Proposta de Programa Para a Gestão de Dados de Pesquisa na UFC**

Com base na revisão conceitual construída neste estudo, bem como nas reflexões suscitadas pelos resultados de pesquisa, na vivência de realização do minicurso e nos pressupostos metodológicos da *Grounded Theory* (Teoria

Fundamentada em Dados), a qual prevê a concepção de uma teoria assentada nos dados, visando acrescentar ou trazer novos conhecimentos à área do fenômeno (CASSIANI; CALIRI; PELÁ, 1996) ou ainda gerar uma explicação (teoria) de um processo, ação ou interação construída pelas visões dos sujeitos de pesquisa. Exatamente por isso a Grounded Theory é considerada substantiva, pois deriva do resultado representativo da realidade dos sujeitos e situações estudadas. (CRESWELL, 2014; MARTINS; THEÓPHILO, 2018).

Nesse sentido, com base nos dados coletados e a partir da investigação realizada, esta pesquisa sugere uma proposta de **Programa de Gestão de Dados de Pesquisa (PGDP) para a UFC** (Apêndice H), que se concentra no desenvolvimento de políticas, diretrizes, ações de educação e informação, produtos, serviços e gestão dos dados de pesquisa na universidade (Figura 38).

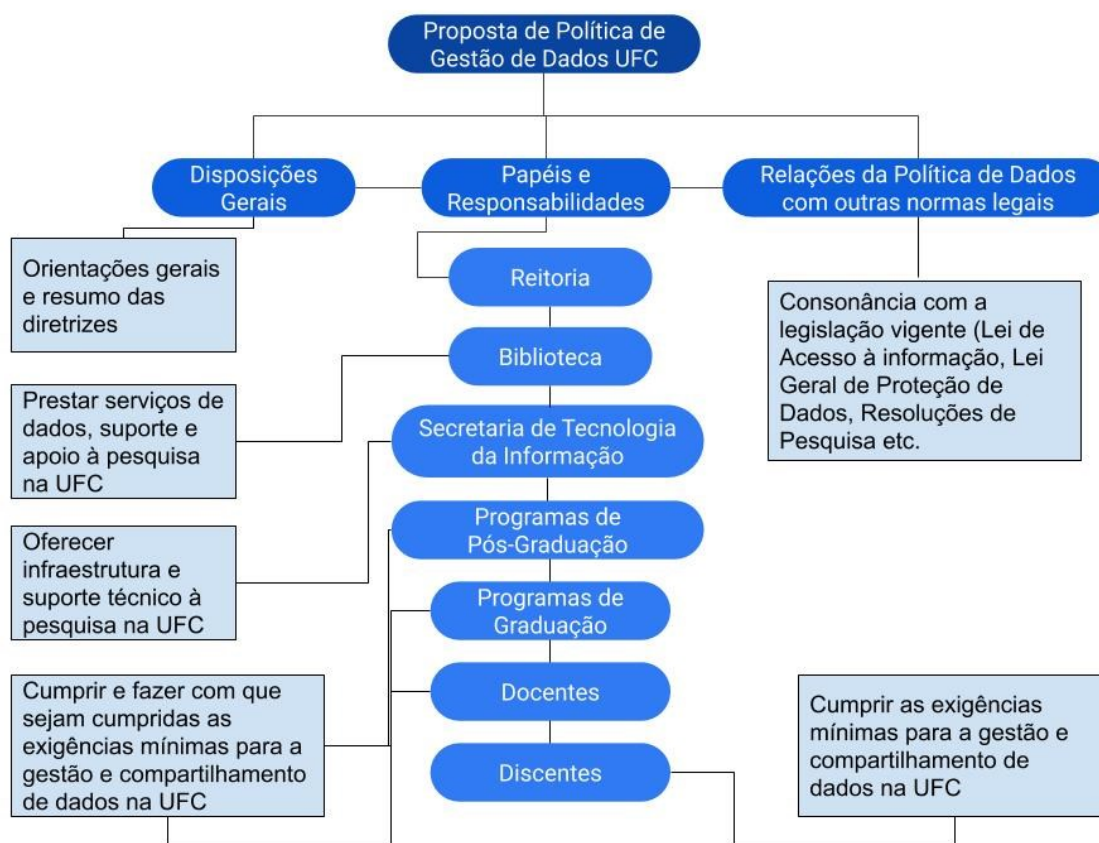
Figura 38 - Estrutura do Programa de Gestão de Dados de Pesquisa



Fonte: Elaborado pela autora. (2020).

A estrutura do PGDP representada na figura 38 ilustra a elaboração de uma proposta de Política de Gestão de Dados de Pesquisa, Diretrizes para a criação e aplicação dessas políticas, assim como o estabelecimento de normas e padrões para facilitar o processo de gestão dos dados de pesquisa na instituição. A Política visa ser um dispositivo norteador e quadro geral abrangente para a Universidade Federal do Ceará (UFC), na orientação da aplicação de procedimentos específicos voltados para a pesquisa aberta e a gestão dos dados oriundos de pesquisa científica na instituição. Apresentará as condições, os papéis e as responsabilidades da universidade e da comunidade acadêmica interna referente à gestão e preservação dos dados científicos atuais ou futuros (Figura 39). Evidentemente, reconhece-se que a implementação da política depende do apoio institucional e poderá levar alguns anos até a sua efetiva implantação.

Figura 39 - Estrutura da Proposta da Política de Dados



Fonte: Elaborado pela autora. (2020).

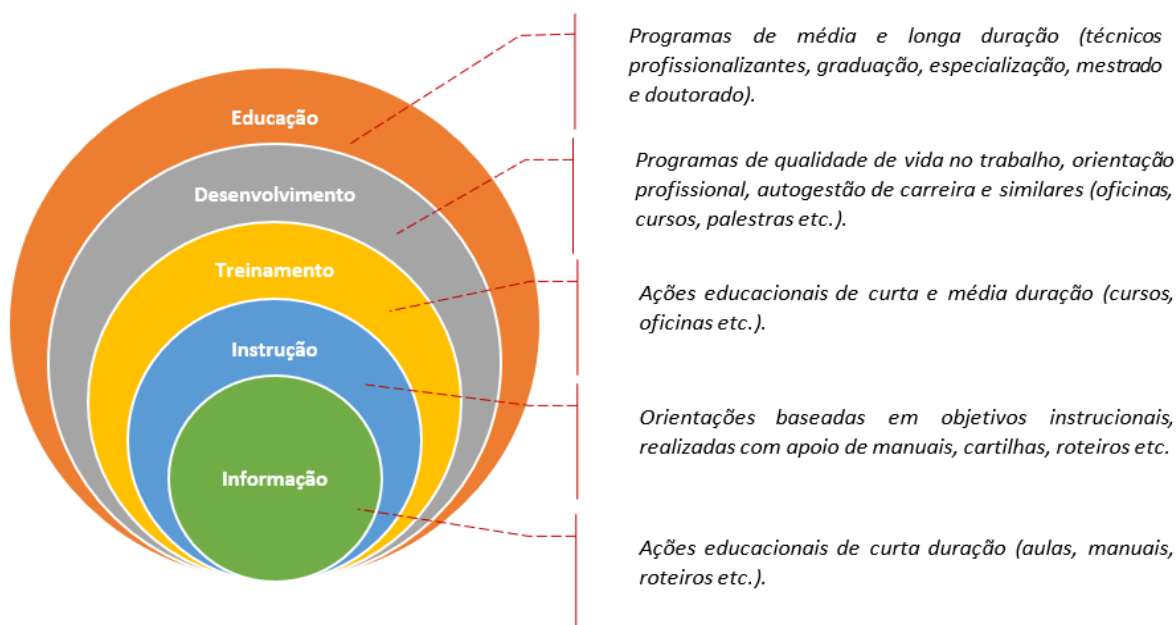


Diante de um cenário de explosão de dados e informações hodiernamente, estabelecer políticas, diretrizes e padrões torna-se fator estratégico para a administração e governança dos dados de pesquisa científica (pré-registro / registro, coleta, documentação, arquivamento, preservação, compartilhamento e publicação), tão necessário para acompanhar as tendências no campo científico nacional e internacional, além de advogar cada vez mais pela adoção de processos de pesquisa sustentáveis, abertos (abertura quando possível, fechado quando necessário), verificáveis, reproduzíveis e dinâmicos. Como parte do **Programa de Gestão de Dados de Pesquisa**, esta pesquisa também desenvolveu uma minuta que apresenta uma Política de Gestão de Dados para a UFC (Apêndice I), a qual dispõe sobre as responsabilidades da universidade e da comunidade acadêmica interna referente à gestão e preservação dos dados científicos atuais ou futuros, além de aplicar-se a todas as pesquisas conduzidas por servidores docentes, servidores técnico-administrativos em educação e discentes da UFC, independentemente de ser a pesquisa financiada ou não.

Entende-se que a idealização de um PGDP que contempla a criação de políticas, também precisa incorporar outras dimensões que dependem da ação humana, do contexto e da infraestrutura. Além disso, a gestão dos dados de pesquisa representa muito mais do que o apoio à excelência em pesquisa, pois os dados são a matéria-prima do conhecimento científico, por isso, estão se tornando tão importantes para todas as áreas da sociedade. Para isso, serão necessários serviços, políticas e infraestrutura. (COSTA, 2017, SHEARER, 2015).

Com efeito, o PGDP prevê ainda a execução de ações de educação e informação com vistas a sensibilizar a comunidade acadêmica para a necessidade do gerenciamento dos dados de pesquisa, além de ser uma outra maneira de incentivar a adoção dessas práticas na instituição. A figura 40 a seguir ilustra os tipos de ações de educação e informação que serão envidadas no programa.

Figura 40 - Tipos de ações de educação e informação



Fonte: Baseado em Filatro *et al.* (2019).

Embasado na literatura científica, esse eixo da proposta do programa inspirou-se em uma série de referências internacionais que recomendam a criação e realização de treinamentos nesse âmbito, a exemplo do documento elaborado pela European Commission e o Working Group on Education and Skills under Open Science (2017), intitulado “Providing researchers with the skills and competencies they need to practice Open Science”; e o documento “Implementing Research Data Management Policies Across Europe: Experiences from Science Europe Member Organisations” publicado pela Science Europe (2020); além das políticas e orientações de universidades internacionais (University of Edinburgh - Escócia; University of Queensland - Austrália; Delft University of Technology - TUDELFT - Holanda; Universidade do Minho - Portugal) e treinamentos consagrados como o MANTRA, Open Science MOOC, FOSTER, Formação Modular em Ciência Aberta da Fiocruz, entre outras.

As universidades supracitadas implantaram seus serviços de gestão de dados iniciando pela conscientização de seus usuários, ou seja, começar por uma campanha de sensibilização junto aos pesquisadores sobre a importância de os dados científicos estarem abertos em prol do avanço da ciência é uma prática consolidada comprovada e registrada na literatura científica. Ademais, alguns autores reforçam que essas

ações sejam acompanhadas da criação de uma política institucional que contemple toda a rede de pesquisadores. (COSTA, 2017, SAYÃO; SALES, 2015). Inclusive, vale ressaltar que o PGDP está em consonância com o Plano de Desenvolvimento Institucional da UFC (PDI 2018-2022).

Este programa baseia-se ainda na abordagem da mediação da informação na Ciência da Informação, na qual “[...] o bibliotecário atua como um “fornecedor” de informações que atendem as demandas e necessidades dos usuários.” (FARIAS; FARIAS, 2017, p. 333). Proposta por Silva (2015), Silva e Farias (2017) e também aplicada em um programa de mediação técnica da informação por Araújo (2019), a mediação da informação é caracterizada como um conjunto de práticas construtivas, um programa de ação, intervenção e interferência do bibliotecário, por exemplo, empregada nos serviços de informação, constituindo-se como uma das principais atividades técnica, pedagógica, institucional e pragmática da mediação.

Ademais, a mediação é voltada para os estudos de usuários e para a busca de uma aproximação da comunidade atendida, inclusive no incentivo à autonomia dos usuários, a exemplo da realização de capacitações e treinamentos como uma das formas de os usuários apreenderem e apropriarem-se da informação, além de participarem da construção coletiva de conhecimento na universidade. No caso específico deste programa, a mediação da informação contribuirá sobremaneira para a organização, gestão da informação e gerenciamento dos dados de pesquisa na instituição.

Constitui-se como outra interface do PGDP a criação de produtos, tais como um portfólio de cursos especializados, totalmente focado nas necessidades de treinamento para a gestão dos dados de pesquisas científicas desenvolvidas na universidade, além da elaboração de materiais instrucionais e afins para as capacitações a serem desenvolvidas. Apesar de algumas instituições ofertarem excelentes capacitações modulares e sem tutoria, a exemplo da Fiocruz e da iniciativa FOSTER, o desenho e modelo de capacitação no PGDP poderá utilizar abordagens com ou sem tutoria (em capacitações à distância) e treinamentos presenciais (com a mediação / intervenção de um bibliotecário). Para complementar a disponibilização de conteúdos, sugere-se dentro da proposta do PGDP a criação de um portal institucional em que constem informações sobre as capacitações oferecidas, dicas de gerenciamento dos dados de pesquisa, entre outras informações relacionadas.

O PGDP também sugere a criação de Serviços de Dados, Apoio e Suporte ao Pesquisador. De acordo com essa proposta, a biblioteca é a responsável pela prestação desses serviços, assim como ocorre em outras bibliotecas universitárias do mundo. Será dada assistência ao pesquisador em assessoria, consultoria, aconselhamento e treinamentos baseados nas ações de educação e informação como parte do programa.

O último componente do PGDP está relacionado ao eixo da gestão, parte que será dedicada à realização do levantamento, mapeamento e diagnóstico institucional prévio sobre a produção dos dados de pesquisas desenvolvidas na UFC. A etapa de diagnóstico do programa também irá verificar quais as possíveis limitações, as condições de infraestrutura e recursos humanos para as atividades de gestão dos dados de pesquisa. Posteriormente, será realizada uma avaliação comparativa a cada ano para acompanhar o desempenho do programa durante o seu período de execução. Durante esses processos, outras formas de avaliação poderão ser empregadas para avaliar o impacto, desempenho e resultados obtidos no decurso do programa.

A gestão dos dados exige apoio institucional, além de uma parceria entre a biblioteca e os pesquisadores, portanto, um dos objetivos do programa é estabelecer uma relação profissional de confiança entre os pesquisadores e os bibliotecários. Assim, entende-se que a atuação de um bibliotecário *liaison* que possa estabelecer uma relação de proximidade com os cursos atendidos possa colaborar para construir essa ponte, de forma que também seja possível auxiliar e desenvolver a curadoria dos dados de pesquisa junto aos laboratórios, grupos de pesquisas e pesquisadores. Um dos últimos pontos do PGDP, porém, não menos importante, é a sua implantação e avaliação após o seu início.

A proposta de programa prevê ainda uma interlocução com a comunidade acadêmica a fim de engajar e envolver todos os *stakeholders*, angariar voluntários e apoiadores da causa, promovendo a prática do *advocacy* em torno da gestão dos dados de pesquisa e fortalecendo os laços entre todos os atores envolvidos. Dessa forma, reconhece-se o valor da ‘ecologia de saberes’, preconizada por Boaventura de Sousa Santos, e o quanto tem extrema relevância no cerne desta proposta, tendo em vista que “[...] promover o diálogo entre vários saberes que podem ser considerados úteis para o avanço das lutas sociais [e] pelos que nelas intervêm” (SANTOS, 2014, p. 332). Em outras palavras, trata-se de um processo coletivo de construção e

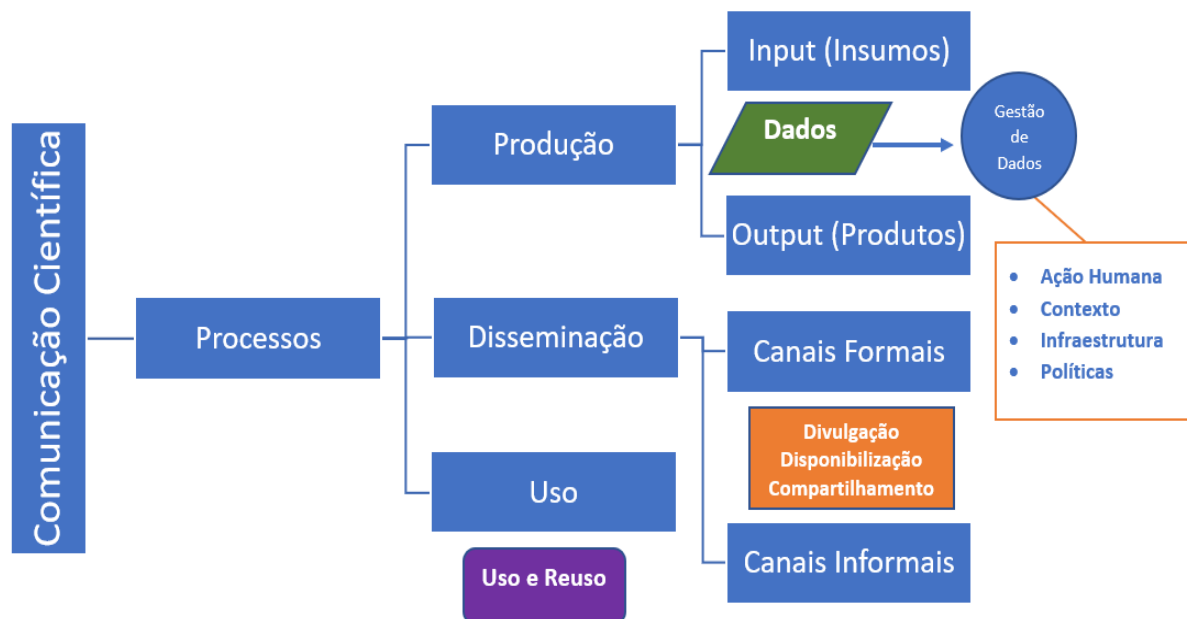
produção do conhecimento. Portanto, a integração, colaboração e a pluralidade de saberes existentes na universidade e no mundo corroboram com o esforço de superação das desigualdades sociais através do compartilhamento do conhecimento científico produzido na UFC, sendo exatamente este o papel da universidade, especialmente no cumprimento de sua responsabilidade social e de seus três pilares: ensino, pesquisa e extensão.

## 5 CONSIDERAÇÕES FINAIS

Indubitavelmente, os dados têm transformado o fluxo da comunicação científica, especialmente em relação à coleta, armazenamento, recuperação, acesso e compartilhamento. Exatamente por isso, a preocupação com o aumento do volume, a variedade e a velocidade têm sido uma constante em um cenário em que são cada vez mais exigidos de todos os atores envolvidos (pesquisadores, bibliotecários e instituições) a capacidade de documentar, armazenar, salvaguardar, preservar e curar dados a curto e em longo prazo. Por isso, a gestão dos dados se torna uma premissa fundamental, porque irá assegurar que as etapas supracitadas sejam compreendidas e aplicadas em prol do conhecimento científico.

Isto posto, contextualizando a relação da comunicação científica com os dados e sua relevância para a produção de conhecimento científico, percebe-se que são ao mesmo tempo insumo (*input*) e produto (*output*) de pesquisa, e sua disseminação pode se dar pelos canais formais e informais da comunicação científica. Um exemplo clássico disso é o próprio *corpus* textual da pesquisadora da área da Linguística: o *corpus* é insumo de entrada (*input*) e após processado pode gerar novos conjuntos e formar os *corpora* textuais, ou seja, torna-se um produto (*output*). Nos processos de comunicação científica presentes na literatura, apenas o **uso** era previsto, logo, incorporou-se a dinâmica do **uso e reuso**, considerando as possibilidades de compartilhamento (prática) e sustentabilidade (valor) da pesquisa científica hodierna (Ciência Aberta / *e-Science*). Assim sendo, a figura 41 ilustra a **gestão** no fluxo dos processos da comunicação científica, uma vez que este passou a ser incorporado às práticas científicas vigentes.

Figura 41 - Constituintes da comunicação científica agregada à gestão de dados de pesquisa



Fonte: Elaborado pela autora. (2019).

Esta pesquisa iniciou com indagações elementares, como a percepção dos pesquisadores sobre os dados de pesquisa e acerca dos princípios da Ciência Aberta. Para complementar a questão inicial, buscou-se investigar quais os desafios e dificuldades encontradas pelos pesquisadores no gerenciamento dos dados, no 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.

Com base nos resultados desta investigação, compreende-se que os objetivos delineados para a condução teórica e metodológica desta investigação foram alcançados. Foi realizada a análise da percepção dos pesquisadores sobre os princípios da Ciência Aberta e o gerenciamento de dados de pesquisa, a partir das coletas realizadas via questionário e entrevista, ao investigar e analisar a compreensão dos pesquisadores a respeito do que é um dado de pesquisa e o que é Ciência Aberta. Essa estratégia foi fundamental para que fossem extraídos os conceitos e o olhar dos pesquisadores e dos estudos desenvolvidos em cada área. Pesquisas anteriores documentadas na literatura científica também foram essenciais para afunilar as ações desenvolvidas na pesquisa.

Averiguou-se ainda as práticas e as necessidades informacionais dos sujeitos da pesquisa por meio da coleta de dados, visitas realizadas nos *campi* e conversas com os usuários, as quais foram fundamentais para descobrir as particularidades de cada área, os tipos de dados produzidos, os métodos utilizados, entre outros aspectos. Elaborou-se uma Política de Gestão de Dados com a finalidade de incentivar a adoção de práticas em gerenciamento de dados de pesquisa, a qual é parte integrante do Programa de Gestão de Dados (PGDP). Um programa proposto com o objetivo de sugerir a criação de serviços de dados na UFC, visando o desenvolvimento de ações formais, de propostas educacionais e de informação voltadas para a gestão dos dados de pesquisa, o qual inclui a elaboração de políticas, diretrizes e materiais instrucionais. Por último, foram desenvolvidas ações de informação com a realização do minicurso online como parte integrante do programa.

Todas essas ações realizadas durante a pesquisa ainda evidenciaram o papel da biblioteca e dos bibliotecários na gestão de dados de pesquisa, especialmente pelos seus conhecimentos e competências que são primordiais para a condução dos serviços de dados e de apoio ao pesquisador; demonstram a importância da adoção da gestão de dados de pesquisa tanto pelos pesquisadores bem como pelas instituições; mostram que as ressalvas e receios em relação à gestão de dados, assim como o compartilhamento, existe muito mais pelo desconhecimento ou pela preocupação em não ferir aspectos éticos e legais da pesquisa científica, pois, conforme foi visto nesse estudo, os pesquisadores da UFC demonstraram interesse ou abertura para adotar as práticas de gestão de dados de pesquisa e o compartilhamento, evidentemente, desde que tenham a vantagem de explorar seus dados antes de compartilhar; e indicam que por meio da sensibilização, conscientização, capacitação, ações de educação e informação para os usuários (pesquisadores) da comunidade acadêmica é possível contribuir para a mudança de comportamento e adoção das práticas voltadas para a gestão de dados de pesquisa.

Mais uma das evidências de que as ações de educação e informação são efetivas nesse cenário da gestão de dados é o fato de diversas bibliotecas em todo o mundo, terem iniciado a sua oferta de serviços de dados exatamente por meio da conscientização e ações educativas, antes mesmo de oferecer uma robusta infraestrutura tecnológica, a exemplo da University of Edimburgh, University of Queensland, University TUDELFT, University of Leeds, University of Bristol, entre tantas outras. Obviamente, isso exigiu antes que os bibliotecários dessas instituições



estudassem, planejassem, se capacitassem e que se mantivessem em educação permanente para atender a sua comunidade. Aliás, em várias instituições de ensino a biblioteca é o coração da universidade.

Partindo para a realidade dessa pesquisa, durante as visitas ao campo, as conversas com os pesquisadores e as coletas de dados, percebeu-se que trazer à tona esse tema para discussão serviu para evidenciar que a demanda latente já existe na universidade, especialmente pelo fato de que desde o início do estudo, foram recebidos alguns e-mails solicitando participação no minicurso online, outros pediram o agendamento de atendimento personalizado para orientações sobre o gerenciamento de seus dados, e ainda houve um pedido da ementa do curso online como modelo por parte de um professor a fim de propor uma disciplina para a sua turma. Antes de defender essa dissertação, a Pós-Graduação em Direito da UFC convidou a pesquisadora responsável por esse estudo, para uma aula sobre dicas de como gerenciar os dados de pesquisa para os alunos do mestrado e doutorado. O LOCUS, um dos laboratórios do curso de Psicologia, demonstrou interesse em colaborar com essa pesquisa, e ainda espera que a parceria continue para que os mapas afetivos e os dados produzidos por eles possam seguir os moldes da gestão de dados de pesquisa.

Após esses eventos, algumas mudanças de comportamento por parte dos pesquisadores já são perceptíveis, algumas práticas já estão sendo incorporadas, a exemplo do preenchimento do PGD por parte de alguns professores, e estes, estão sendo multiplicadores, pois também afirmaram que irão solicitar de seus alunos o preenchimento e entrega de um PGD. Docentes também estão solicitando auxílio da BCH para a elaboração de um PGD em projetos coordenados por eles.

Diante do exposto, é necessário frisar que, se os pesquisadores não estiverem cientes do que eles têm à sua disposição, dos recursos, das ferramentas e das possibilidades existentes para o gerenciamento dos seus dados de pesquisa, mesmo antes de surgir a oferta da infraestrutura, de nada adiantará. Portanto, a sensibilização e a educação, atreladas ao **fator humano**, é que tornam possível a concretização da gestão dos dados de pesquisa em uma instituição e por seus pesquisadores, aliás, não é segredo lembrar que o bibliotecário está no cerne dessas ações.

Um dos pontos fundamentais discutido nessa dissertação trata da questão do compartilhamento, o que tem ganhado cada vez mais destaque, estando em evidência

por conta da pressão de uma parcela considerável da comunidade científica, que vem questionando o *status quo* das editoras que restringem o acesso ao conhecimento científico publicado em periódicos por meio da cobrança de taxas e paywalls, das grandes corporações e até mesmo do governo. O exemplo mais recente desse fato é o surto global da pandemia do Covid-19 em que os pesquisadores do mundo inteiro se depararam com publicações sobre o assunto em acesso restrito, dependendo assim da solidariedade de outros cientistas ou instituições, que passaram a compartilhar e prover as informações iniciais necessárias para o desenvolvimento de pesquisas em busca da cura para o vírus ou mesmo a criação de uma vacina.

Rapidamente, cientistas de dados e de diversas áreas do conhecimento têm trabalhado também na agregação e disponibilização de dados, para que pesquisadores possam acessar as informações essenciais para suas pesquisas. Esses conjuntos de dados incluem ainda os dados governamentais sobre as taxas de internação, mortes e outras informações relevantes para ajudar no monitoramento da doença e previsão de picos do Covid-19.

Nesse panorama, os periódicos passaram a reforçar a publicação e uso de pré-prints, assim como a criação de números especiais temáticos a respeito do Covid-19. Diante dessa situação, até as editoras e bases de dados têm disponibilizado gratuitamente todos os conteúdos relacionados à pandemia. De forma semelhante, os repositórios de dados têm trabalhado na manutenção de suas coleções e ampliado o espaço de armazenamento para os datasets de pesquisas relacionadas ao vírus que são depositadas, a exemplo do Figshare e Zenodo.

Esses fatos evidenciam a relevância do compartilhamento, do acesso livre ao conhecimento, da necessidade de maior abertura e transparência na pesquisa, da importância da gestão dos dados de pesquisa e de que maneira o conhecimento pode estar disponível para todos.

Como agenda de sugestões para pesquisas futuras, recomenda-se fortemente o estudo da competência em dados (*data literacy*), inclusive em diversos níveis (escolar e universitário). Outra recomendação é o estudo dos termos 'disponibilização' e 'compartilhamento': são usados como sinônimos para se referir ao compartilhamento de dados, entretanto, há que se considerar que existem nuances práticas e terminológicas entre eles, a exemplo de disponibilizar dados em um repositório ou compartilhar os dados com um de seus pares por solicitação. Também ensejam estudos terminológicos as palavras ciclo de vida dos dados, curadoria de

dados, gestão de dados de pesquisa e gestão de dados científicos, pois muitas vezes esses termos são empregados como se tivessem o mesmo significado.

Um dos assuntos que merece estudo aprofundado é a investigação sobre padronização dos metadados ideais para a descrição de cada tipo de dado de pesquisa. Os pesquisadores precisam de orientações a esse respeito, capazes de organizar e auxiliar na recuperação e descoberta dos dados futuramente.

Apesar de existirem modelos consagrados de Planos de Gestão de Dados, este é um tema que merece ser discutido amplamente, pois, há modelos distintos, necessidades e áreas com particularidades que talvez mereçam a elaboração de um novo modelo ou mesmo da proposta de um PGD mais dinâmico, prático e interligado com os processos de pesquisa.

Estudos sobre os processos de comunicação científica que envolvem os dados de pesquisa, assim como a Ciência Aberta, devem ser incentivados e aprofundados, considerando que ainda há pontos de inflexão que precisam ser melhor compreendidos em profundidade. Pesquisas sobre os **princípios FAIR e sua aplicação na prática** são importantes, pois representam benefícios para todos os envolvidos e contribuem para que os dados possam ser usados e reutilizados. Mais do que necessários, são os estudos voltados para as questões éticas e legais dos dados de pesquisa, os processos de anonimização dos dados na teoria e na prática, incluindo a Lei Geral de Proteção de Dados, a fim de trabalhar a temática dos dados sensíveis. Sugere-se ainda o estudo sobre a definição do que é um dado, até se chegar aos dados de pesquisa na corrente do neodocumentalismo.

Finalmente, ressalta-se que essa dissertação delineou um panorama institucional sobre a gestão de dados de pesquisa no contexto da Ciência Aberta e apresentou uma visão geral sobre o tema, portanto, não é intenção desse estudo esgotar a temática discutida, mas sim mostrar que é só o começo do quanto ainda há para se explorar nesse universo.

## REFERÊNCIAS

- ACOSTA-HOYOS, Luís Eduardo. **Colégios invisíveis**: uma nova alternativa para o problema de informação técnico-científica, Brasília: Embrapa, Departamento de Informação e Documentação, 1980.
- 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. p. 9-25. Disponível em: <https://bit.ly/2o2b6c4>. Acesso em: 10 maio 2018.
- ALBAGLI, Sarita; APPEL, André Luiz; MACIEL, Maria Lucia. e-Science e Ciência Aberta: questões em debate. *In*: Encontro Nacional de Pesquisa em Ciência da Informação - ENANCIB, 14., 2013, Florianópolis. **Anais [...]**. Florianópolis: ANCIB, 2013. Disponível em: <http://ridi.ibict.br/handle/123456789/465>. Acesso em: 7 set. 2018.
- ALBAGLI, Sarita; CLINIO, Anne; RAYCHTOCK, Sabryna. Ciência Aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, nov. 2014. Disponível em: <http://revista.ibict.br/liinc/article/view/3593/3072>. Acesso em 10 junho 2015.
- ALMEIDA JÚNIOR, Oswaldo Francisco de. Leitura, mediação e apropriação da informação. *In*: SANTOS, J. P. (org.). **A leitura como prática pedagógica na formação do profissional da informação**. Rio de Janeiro: Fundação Biblioteca Nacional, 2007. p. 33-45.
- ALMEIDA JÚNIOR, Oswaldo Francisco de. Mediação da informação e múltiplas linguagens. **Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 2, n. 1, p. 89-103, jan./dez. 2009. Disponível em: <https://bit.ly/2X9sIRr>. Acesso em: 8 abr. 2019.
- ALMEIDA JÚNIOR, Oswaldo Francisco de. Profissional bibliotecário: um pacto com o excludente. *In*: BAPTISTA, S. G.; MUELLER, S. P. M. (org.) **Profissional da informação**: o espaço de trabalho. Brasília: Thesaurus, 2004. p. 70-86.
- ALVARENGA, Lídia. Alguns enunciados sobre a comunicação e o uso de fontes de informação entre pesquisadores brasileiros da área da Educação. *In*: MUELLER, Suzana Pinheiro Machado; PASSOS, Edilenice Jovelina Lima (org.). **Comunicação Científica**. Brasília: Departamento de Ciência da Informação Universidade de Brasília. p. 123-138.
- ALVES, Virgínia. **Informação científica em Biblioteconomia**: o livre acesso nas universidades federais. Curitiba: Appris, 2017.
- ALVES-MAZZOTTI, Alda Judith. O método nas ciências sociais. *In*: ALVES-MAZZOTTI, Alda Judith. **O método nas ciências naturais e sociais**: pesquisa quantitativa e qualitativa. São Paulo: Pioneira Thomson Learning, 2002. p. 109-187.
- AMARAL, Adriana; NATAL, Geórgia; VIANA, Lucina. Netnografia como aporte metodológico da pesquisa em comunicação digital. **Sessões do Imaginário**, Porto

Alegre, ano 13, n. 20, p. 34-40, 2008. Disponível em:  
<http://revistaseletronicas.pucrs.br/ojs/index.php/famecos/article/view/4829/3687>.  
 Acesso em: 10 maio 2019.

AMERICAN PRODUCTIVITY & QUALITY CENTER. **Knowledge Mapping Concepts and Tools (Collection)**. 2019. Disponível em: <https://bit.ly/2Kbhmsw>.  
 Acesso em: 10 maio 2019.

ANDERSON, Chris. **A cauda longa**: do mercado de massa para o mercado de nicho. Rio de Janeiro: Elsevier; Alta Books, 2006.

ANDERSON, Chris. The Long Tail. **Wired**, [S.l.], 2004. Disponível em:  
<https://www.wired.com/2004/10/tail/>. Acesso em: 10 abr. 2019.

ANGROSINO, Michael. **Etnografia e observação participante**. Porto Alegre: Artmed, 2009.

ANTONELLI-PONTI, Mayra *et al.* Divulgação científica em psicobiologia: educação a distância como estratégia para a promoção da extensão universitária. **Expressa Extensão**, Pelotas, v. 23, n. 3, p. 183-197, 2018.

ARAÚJO, Ana Carolina. **Universidade públicas realizam mais de 95% da Ciência no Brasil**. Alfenas: Universidade Federal de Alfenas, 2019. Disponível em:  
<https://bit.ly/2wspyMK>. Acesso em: 18 abr. 2019.

ARAÚJO, Ana Rafaela Sales de. **Mediação no âmbito da organização da informação**: proposta de um programa de atuação para o Sistema de Bibliotecas da Universidade Federal do Ceará. 2018. 208 f. Dissertação (Mestrado) – Universidade Federal do Cariri, Centro de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Biblioteconomia, Mestrado Profissional em Biblioteconomia, Juazeiro do Norte, 2018.

ARROW, Kenneth N. Economic welfare and the allocation of resources for invention. *In*: UNIVERSITIES NATIONAL BUREAU COMMITTEE FOR ECONOMIC RESEARCH. Committee on Economic Growth of the Social Science Research Council. **The Rate and Direction of Inventive Activity**. Princeton: Princeton University Press, 1962. p. 609-625.

ARCENEUX, Kevin *et al.* We Tried to Publish a Replication of a Science Paper in Science. The Journal Refused. **Slate**, 2019. Disponível em:  
<https://slate.com/technology/2019/06/science-replication-conservatives-liberals-reacting-to-threats.html>. Acesso: 11 nov. 2019.

ASCHWANDEN, Christie. Science Isn't Broken: It's just a hell of a lot harder than we give it credit for. **FiveThirtyEight**, 2015. Disponível em:  
<https://fivethirtyeight.com/features/science-isnt-broken/#part1>. Acesso em: 10 mar. 2019.

ASSANTE, Massimiliano *et al.* Are Scientific Data Repositories Coping with Research Data Publishing? **Data Science Journal**, [S.l.], v. 15, n. 6, p. 1-24, 2016. DOI: <http://dx.doi.org/10.5334/dsj-2016-006>.

AUER, Sören; BRYL, Volha; TRAMP, Sebastian (Ed.). **Linked Open Data – Creating Knowledge Out of Interlinked Data**: Results of the LOD2 Project. New York: Springer, 2014. Disponível em: <https://link.springer.com/content/pdf/10.1007%2F978-3-319-09846-3.pdf>. Acesso em: 7 set. 2017.

AUGUSTO, Diego Menezes; FEITOSA, Maria Zelfa de Souza; BOMFIM, Zulmira Áurea Cruz. A utilização dos mapas afetivos como possibilidade de leitura do território no CRAS. **Estudos Interdisciplinares em Psicologia**, Londrina, v. 7, n. 1, p. 145-158, jun. 2016. Disponível em: <http://twixar.me/XHST>. Acesso em: 10 maio 2019.

AVENTURIER, Pascal; ALENCAR, Maria de Cléofas Faggion. Os desafios dos dados de pesquisas abertos. **RECIIS**, Rio de Janeiro, v. 10, n. 3, jul./set. 2016. Disponível em: <https://bit.ly/2WaOAiV>. Acesso em: 10 abr. 2019.

ÁVILA, Thiago José Tavares. **Uma proposta de modelo de processo para publicação de Dados Abertos Conectados Governamentais**. 2015. Dissertação (Mestrado em Modelagem Computacional do Conhecimento) – Instituto de Computação, Universidade Federal de Alagoas, Maceió, 2015. Disponível em: <https://www.researchgate.net/publication/317087042>. Acesso em: 7 set. 2017.

BALL, Alexander. **Review of Data Management Lifecycle Models**. Bath, UK: University of Bath, 2012. Disponível em: <https://is.gd/gFZzzk>. Acesso em: 7 set. 2018.

BARBA, Lorena A. **Terminologies for Reproducible Research**. arXiv:1802.03311 [cs.DL], 2018.

BAUER, Florian; KALTENBÖCK, Martin. **Linked open data: the essentials – a quick start guide for decision makers**. [S.l.]: Semantic Web Company, 2012. 59 p. Disponível em: <http://www.reeep.org/LOD-the-Essentials.pdf>. Acesso em: 7 set. 2017.

BÉCARD, N.; CASTETS-RENARD, C.; CHASSANG, G.; COURTOIS, M. A.; DANTANT, M.; GANDON, N.; MARTIN, C.; MARTELLETTI, A.; MENDOZA-CAMINADE, A.; MORCLETTE, N.; NEIRAC, C. **Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France**. 2016. Disponível em: <https://is.gd/6UMY6g>. Acesso em: 08 fev. 2019.

BECKER, Howard S. **Métodos de pesquisa em Ciências Sociais**. 3. ed. São Paulo: Hucitec, 1994. 178 p.

BELL, Gordon. Prefácio. *In*: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (org.). **O quarto paradigma**: descobertas científicas na era da e-Science. São Paulo: Oficina de Textos, 2011. p. 11-15.

BELLONI, Maria Luiza. **Educação a distância**. Campinas: Autores Associados, 2011.

BELUZZO, Regina C. B. **Liderança & formação e desenvolvimento de equipes**. São Paulo: USP, 2002.

BENTES PINTO, Virgínia *et al.* Netnografia: uma abordagem para estudos de usuários no ciberespaço. *In*: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 9., 2007, Açores. **Anais [...]**. Lisboa: APBAD, 2007. Disponível em: <https://bit.ly/2VZiAcz>. Acesso em: 5 ago. 2015.

BENTES PINTO, Virgínia; CAVALCANTE, Lídia Eugênia. Pesquisa bibliográfica e documental: o fazer científico em construção. *In*: BENTES PINTO, Virgínia; VIDOTTI, Silvana Aparecida Borsetti Gregório; CAVALCANTE, Lídia Eugênia (org.). **Aplicabilidades metodológicas em Ciência da Informação**. Fortaleza: Edições UFC, 2015. p. 15-34.

BERNAL, John Desmond. Science and Industry. *In*: HALL, Daniel *et al.* **The Frustration of Science**. New York: Arno Press, c1935. p. 42-78.

BERNAL, John Desmond. **The Social function of Science**. London: George Routledge & Sons, 1939. Disponível em: <https://bit.ly/2W2OpkG>. Acesso em: 03 dez. 2018.

BERNERS-LEE, Tim. **Linked data**. 2006. Disponível em: <https://goo.gl/katw>. Acesso em: 7 set. 2017.

BERRÍO-ZAPATA, Cristian. **Tecnologia da informação, discurso e poder**: análise de domínio a partir do conceito de exclusão digital na perspectiva da teoria centro-periferia. 2015. 380 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, 2015. Disponível em: <https://bit.ly/2wtsCZ7>. Acesso em: 10 maio 2019.

BEZJAK, Sonja *et al.* **Open Science Training Handbook**. [S.l. : s.n.], 2018. Disponível em: <https://foster.gitbook.io/manual-de-formacao-em-ciencia-aberta/>. Acesso em: 20 abr. 2018 DOI: <https://doi.org/10.5281/zenodo.1212496>.

BITTENCOURT, Maíra. Grounded theory como metodologia para o estudo das mídias digitais. **C&S**, São Bernardo do Campo, v. 39, n. 1, p. 143-167, jan./abr. 2017. Disponível em: <https://bit.ly/2l5YcSd>. Acesso em: 10 maio 2019.

BLAIR, Ann M. **Too Much to Know**: Managing Scholarly Information before the Modern Age. [S.l.], Yale University Press, 2010.

BOMFIM, Zulmira Áurea Cruz. **Cidade e afetividade**: estima e construção dos mapas afetivos de Barcelona e de São Paulo. 2003. 237 f. Tese (Doutorado em Psicologia) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2003.

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

BORGMAN, Christine L. **Research Data: Who Will Share What, with Whom, When, and Why?** RatSWD Working Paper No. 161. 2010. Disponível em: <https://is.gd/T72Ucv>. Acesso em: 10 mar. 2019.

BORGMAN, Christine L. **Scholarship in the Digital Age: Information, Infrastructure, and the Internet**. Cambridge, MA: MIT Press, 2007.

BORKO, Harold. Information science: what is it? **American Documentation**, Santa Monica, p. 3-5, jan. 1968. DOI: <https://doi.org/10.1002/asi.5090190103>.

BOURDIEU, Pierre. O campo científico. *In*: ORTIZ, Renato. **A sociologia de Pierre Bourdieu**. São Paulo: Olho d'Água, 1976. p. 112-143.

BOURDIEU, Pierre. The specificity of the scientific field and the social conditions of the progress of reason. **Social Science Information**, [S./], v. 14, n. 6, p. 19-47, 1975.

BRAGA, Adriana. Netnografia: compreendendo o sujeito nas redes sociais. *In*: NICOLACI-DA-COSTA, Ana Maria; ROMÃO-DIAS, Daniela (org.). **Qualidade faz diferença: métodos qualitativos para a pesquisa em psicologia e áreas afins**. Rio de Janeiro: PUC-RIO: São Paulo: Loyola, 2013.

BRASIL. Secretaria de Logística e Tecnologia da Informação; Ministério do Planejamento, Orçamento e Gestão. **Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0**. Brasília: Ministério do Planejamento, Orçamento e Gestão, 2019. Disponível em: <https://bit.ly/30YAi3s>. Acesso em: 10 set. 2018.

BRESSER-PEREIRA, Luiz Carlos. Capitalismo financeiro-rentista. **Estudos avançados**, São Paulo, v. 32, n. 92, p. 17-29, abr. 2018. Disponível em: <https://bit.ly/2WutXxv>. Acesso em: 10 jan. 2019.

BRIET, S. **Qu'est-ce que la documentation?** Paris: Éditions Documentaires Industrielles et Técnicas, 1951.

BROOKES, Bertram C. The foundations of information science. Part I. **Journal of Information Science**, Cambridge, v. 2, n. 1, p. 125-133, 1980. Disponível em: <http://journals.sagepub.com/doi/abs/10.1177/016555158000200302>. Acesso em: 10 abr. 2018.

BROWN, John Seely; DUGUID, Paul. **The social life of information**. [S./]: Harvard Business School Publishing, 2017.

BRUYNE, Paul de; HERMAN, Jacques; SCHOUTHEETE, Marc de. **Dinâmica da pesquisa em Ciências Sociais**. 3. ed. Rio de Janeiro: Francisco Alves, 1991.

BUCKLAND, Michael K. Information as thing. **Journal of the American Society for Information Science (JASIS)**, v. 45, n. 5, p. 351-360, 1991. DOI: [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<351::AID-ASI5>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<351::AID-ASI5>3.0.CO;2-3).



BURNETT, Peter. What is the role of a librarian in Research Data Management? *In: INASP Blog*, 2013. Disponível em: <http://blog.inasp.info/research-data-management-role-librarians/>. Acesso em: 22 abr. 2019.

BUSH, Vannevar. As we may think. *Atlantic Monthly*, [S. l.], v. 176, n. 1, 1945. Disponível em: <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>. Acesso em: 22 abr. 2019.

CALARCO, Pascal; SHEARER, Kathleen; SCHMIDT, Birgit; TATE, Dominic. Librarians' Competencies Profile for Scholarly Communication and Open Access. 2016. Disponível em: <https://bit.ly/2xHTQgs>. Acesso em: 20 jul. 2017.

CAMPBELL, D. T.; FISKE, D. W. Convergent and discriminant validation by the multitrait - multimethod matrix. *Psychological Bulletin*, [S. l.], v. 56, n. 2, p. 81-105, 1959. DOI: <http://dx.doi.org/10.1037/h0046016>.

CAPURRO, R.; HJORLAND, B. O conceito de informação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 12, n. 1, jan./abr. 2007. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54/47>. Acesso em: 17 abr. 2018.

CAPURRO, Rafael. Epistemologia y ciência de la informacion. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 5., 2003, Belo Horizonte. *Anais [...]*. Belo Horizonte: UFMG, 2003.

CASSIANI, Sílvia Helena de Bortoli; CALIRI, Maria Helena Larcher; PELÁ, Nilza Teresa Rotter. A teoria fundamentada nos dados como abordagem da pesquisa interpretativa. *Rev. Latino-Am. Enfermagem*, Ribeirão Preto, v. 4, n. 3, p. 75-88, dez. 1996. Disponível em: <https://bit.ly/2We3fK1>. Acesso em: 10 maio 2019. DOI: <http://dx.doi.org/10.1590/S0104-11691996000300007>.

CASTELLS, Manuel. *A sociedade em rede*. São Paulo: Paz e Terra, 1999.

CERVO, Amado; BERVIAN, Pedro. *Metodologia científica*. 5. ed. São Paulo: Prentice Hall, 2002.

CHARMAZ, Kathy. *A construção da Teoria Fundamentada*. Porto Alegre: Artmed, 2009.

CHARMAZ, Kathy. *Constructing Grounded Theory: A Practical guide through qualitative analysis*. London: Sage, 2006.

CHARMAZ, Kathy. Constructionism and the Grounded Theory Method. *In: HOLSTEIN, James; JABBER, Gibrium. Handbook of Constructionist Research*. [S. l.:s.n.], 2008a.

CHARMAZ, Kathy. Shifting the Grounds: Constructivist Grounded Theory Methods. *In: MORSE, J.; STERN, P.; CORBIN, J.; BOWERS; CHARMAZ, K.; CLARKE, A. E. Developing Grounded Theory: The Second Generation*. Walnut Creek: Left Coast Press, 2008b.

CONSORTIUM OF EUROPEAN SOCIAL SCIENCE DATA ARCHIVES - CESSDA Training Working Group (2017 - 2018). **CESSDA Data Management Expert Guide**. Bergen, Norway: CESSDA ERIC. Disponível em: <https://www.cessda.eu/DMGuide>. Acesso em: 10 mar. 2019.

CORBIN, Juliet; STRAUSS, Anselm L. Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. **Qualitative Sociology**, [S.l.], v. 13, n. 1, 1990. Disponível em: <https://med-fom-familymed-research.sites.olt.ubc.ca/files/2012/03/W10-Corbin-and-Strauss-grounded-theory.pdf>. Acesso em: 10 maio 2019.

CORRÊA, Maurício de Vargas; ROZADOS, Helen Beatriz Frota. A netnografia como método de pesquisa em ciência da informação. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, João Pessoa, v. 12, n. 2, 2017. DOI: 10.22478/ufpb.1981-0695.2017v12n2.36531.

CORTI, Louise; VAN DEN EYNDEN, Veerle; BISHOP, Libby; MORGAN-BRETT, Bethany. **Managing and sharing data: best practice for researchers**. 3. ed. Colchester: UK Data Archive, 2011. Disponível em: <https://data-archive.ac.uk/media/2894/managingsharing.pdf>. Acesso em: 10 mar. 2019.

COSTA, Maíra Murrieta. **Diretrizes para uma política de gestão de dados científicos no Brasil**. 2017. 288 f. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília. Faculdade de Ciência da Informação, Brasília, 2017. Disponível em: <http://repositorio.unb.br/handle/10482/24895>. Acesso em: 10 maio 2018.

COSTA, Maíra Murrieta; CUNHA, Murilo Bastos da. O bibliotecário no tratamento de dados oriundos da e-science: considerações iniciais. **Perspectivas em Ciência da Informação**, v. 19, n. 3, p. 189–206, 2014. DOI:10.1590/1981-5344/1900.

COSTA, Sely Maria de Souza. 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, ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1139/1294>. Acesso em: 24 jun. 2018.

COX, A.M.; KENNAN, M.A.; LYON, L.; PINFIELD, S. Developments in research data management in academic libraries: Towards an understanding of research data service maturity. **Journal of the Association for Information Science and Technology**, [S.l.], 68 (9). pp. 2182-2200, 2017. Disponível em: <https://bit.ly/30PWsoE>. Acesso em: 10 fev. 2019.

CRANE, Diana. **Invisible colleges: diffusion of knowledge in scientific communities**, Chicago: The University of Chicago Press, 1988.

CRESWELL, John. W. **Investigação qualitativa e projeto de pesquisa: escolhendo entre abordagens**. 3. ed. Porto Alegre: Penso, 2014.

CRESWELL, John. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 2. ed. Porto Alegre: Artmed, 2010.

CRUZ, Carlos Henrique de Brito. **100 universidades e institutos no Brasil com mais artigos científicos**. 2019. Extraído da Base de dados InCites da Clarivates. Disponível em: <https://bit.ly/2VZ4IPu>. Acesso em: 18 abr. 2019.

CUNHA, Murilo Bastos da. A biblioteca universitária na encruzilhada. **DataGramZero**: Revista de Ciência da Informação, Rio de Janeiro, v. 11, n.6, dez. 2010. Disponível em: <https://tinyurl.com/yx9xuhuj>. Acesso em: 18 dez. 2013.

DASGUPTA, R.; DAVID, P. A. Information disclosure and the economics of science and technology. In: FEIWEL, G. R. **Arrow and the Ascent of Modern Economic Theory**. Houndmills: Macmillan Press, 1987. p. 519-542.

DASGUPTA, R.; DAVID, P. A. Towards a new economics of science. **Research Policy**, [S.l.], v. 23, p. 487-521, 1994.

DATAONE. **DataONE Education Module**: Data Entry and Manipulation. 2012. Disponível em: <https://tinyurl.com/y6d9muvq>. Acesso em: 10 mar. 2019.

DATAONE. DataONE Public Participation in Scientific Research Working Group. **Data Management Guide for Public Participation in Scientific Research**. 2013. Disponível em: <https://tinyurl.com/y3tyhbnl>. Acesso em: 10 mar. 2019.

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

DAVENPORT, T. H. **Big data no trabalho**: derrubando mitos e descobrindo oportunidades. Rio de Janeiro: Elsevier, 2014.

DAVID, Paul A. The historical origins of 'Open Science'. An essay on patronage, reputation and common agency contracting in the scientific revolution. **Capitalism and Society**, v. 3, n. 2, 2008. Disponível em: <https://is.gd/UFjzOh>. Acesso em: 7 set. 2018.

DAY, Ronald. **The modern invention of information**: discourse, history and power. Illinois: Southern Illinois University Press, 2001.

DENZIN, N. K.; LINCOLN, Y. S. **O planejamento da pesquisa qualitativa**. Porto Alegre: Penso, 2006.

DIGITAL CURATION CENTRE. **DCC Curation Lifecycle Model**. 2019. Disponível em: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>. Acesso em: 7 set. 2018.

DIGITAL SCIENCE. Figshare. **Blockchain for Research**: Perspectives on a New Paradigm for Scholarly Communication. London: Holtzbrinck Publishing Group: Springer Nature: Wiley, 2017. DOI: <https://doi.org/10.6084/m9.figshare.5607778>. Disponível em: [https://figshare.com/articles/\\_/5607778](https://figshare.com/articles/_/5607778). Acesso em: 8 abr. 2019.

DIGITAL SCIENCE. Figshare. **The State of Open Data 2017**: a selection of analyses and articles about open data, curated by Figshare. London: Holtzbrinck Publishing Group: Springer Nature: Wiley, 2017. DOI: [10.6084/m9.figshare.5481187](https://doi.org/10.6084/m9.figshare.5481187). Disponível em: <https://www.digital-science.com/resources/portfolio-reports/state-open-data-2017/>. Acesso em: 10 mar. 2019.

DRUCKER, P. F. **The age of discontinuity**. London: Heinemann, 1968.

DUARTE, T. A possibilidade da investigação a 3: reflexões sobre triangulação (metodológica). **CIES e-Working Papers**, Lisboa, n. 60, 2009. Disponível em: [https://repositorio.iscte-iul.pt/bitstream/10071/1319/3/CIES-WP60%20\\_Duarte.pdf](https://repositorio.iscte-iul.pt/bitstream/10071/1319/3/CIES-WP60%20_Duarte.pdf). Acesso em: 11 maio 2019.

EGBERT, Henrik. The Gift and the Open Science. **Bulgarian Economic Papers**, [S.l.], n. 4, p. 1-20, 2018. Disponível em: <https://bit.ly/2MhmM7O>. Acesso em: 17 jun. 2018.

ELIAS, Claudia de Souza Rodrigues *et al.* Quando chega o fim? Uma revisão narrativa sobre terminalidade do período escolar para alunos deficientes mentais. **SMAD: Revista Electrónica en Salud Mental, Alcohol y Drogas**, v. 8, n. 1, p. 48-53, 2012. Disponível em: <http://www.revistas.usp.br/smad/article/view/49594>. Acesso em: 10 dez. 2018.

ERWAY, Ricky *et al.* **Building Blocks**: Laying the Foundation for a Research Data Management Program. Dublin, Ohio: OCLC Research, 2015. <https://bit.ly/2KgXAMc>. Acesso em: 7 set. 2018.

ETZKOWITZ, Henry; LEYDESDORFF, Loet. The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. **Res. Policy**, [S.l.], 29, p. 109–123, 2000.

ETZKOWITZ, Henry; LEYDESDORFF, Loet. **Universities in the Global Knowledge Economy**: A triple helix of university-industry-government relations. London: Cassell, 1997.

EUROPEAN COMMISSION. **Guidelines on Data Management in Horizon 2020**. Version 2.1. 2016. Disponível em: <https://tinyurl.com/y3vlqymy>. Acesso em: 10 mar. 2019.

EUROPEAN COMMISSION. Working Group on Education and Skills under Open Science. **Providing researchers with the skills and competencies they need to practise Open Science**: Open Science Skills Working Group Report. Brussels: European Union, 2017. Disponível em: <https://tinyurl.com/y3byn4eh>. Acesso em: 7 set. 2018.

FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER. **Planos de Gestão de Dados no Horizonte 2020**. 2019. Cursos. Disponível em: <https://www.fosteropenscience.eu/node/2643>. Acesso em: 10 mar. 2019.

FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER. **What is Open Science**. 2015. Disponível em: <https://tinyurl.com/y4l3l44g>. Acesso em: 7 set. 2018.

FALKEMBACH, Elza Maria F. Diário de campo: um instrumento de reflexão. **Contexto e educação**, Ijuí, v. 2, n. 7, p. 19-24, jul./set. 1987.

FANELLI, Daniele. Is science really facing a reproducibility crisis, and do we need it to? **PNAS**, [S.l.], v. 115, n. 11, p. 2628-2631, mar., 2018. DOI: <https://doi.org/10.1073/pnas.1708272114>.

FARIAS, Maria Giovanna Guedes. **Análise da produção, implementação e avaliação de um modelo de mediação da informação no contexto de uma comunidade urbana**. Salvador, 2014. 283 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal da Bahia, Instituto de Ciência da Informação, 2014.

FARIAS, Maria Giovanna Guedes. Mediação e competência em informação: proposições para a construção de um perfil de bibliotecário protagonista. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 6, n. 2, p. 106-125, set. 2015. Disponível em: [www.revistas.usp.br/incid/article/view/101368](http://www.revistas.usp.br/incid/article/view/101368). Acesso em: 10 maio. 2018.

FARIAS, Maria Giovanna Guedes; FARIAS, Gabriela Belmont de. Mediação na Ciência da Informação: uma análise bibliométrica na coleção Benancib. **Revista Ibero-Americana de Ciência da Informação**, v. 10, n. 2, p. 332-349, 2017.

FARIAS, Maria Giovanna Guedes; VARELA, Aida Varela; FREIRE, Isa Maria. Competência em informação para comunidades: empoderamento e protagonismo social. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 24, n. 1, p. 4-24, abr. 2019. Disponível em: <https://bit.ly/2xnRZ2l>. Acesso em: 29 abr. 2019.

FECHER, Benedikt; FRIESIKE, Sascha. Open Science: One Term, Five Schools of Thought. In: BARTLING, Sönke; FRIESIKE, Sascha. **Opening Science: the evolving guide on how the internet is changing research, collaboration and scholarly publishing**. New York: Springer, 2014. p. 17-47. Disponível em: <https://goo.gl/yj2wYU>. Acesso em: 20 jul. 2017.

FERGUSON, Adam *et al.* Big data from small-data: data sharing in the 'long tail' of neuroscience. **Nature Neuroscience**, v. 17, n. 11, p. 1442-1447, nov. 2014. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4728080/>. Acesso em: 10 maio 2019.

FERREIRA, Valdinéia Barreto. **E-Science e Políticas Públicas: ciência, tecnologia e inovação no Brasil**. Salvador: EDUFBA, 2018.

FILATRO, Andréa; PICONEZ, Stela Bertholo. **Design instrucional contextualizado: planejamento, elaboração e avaliação de materiais didáticos para a educação a**

distância. [S.l. : s.n.], 2004. Disponível em: <https://bit.ly/2Sin2nW>. Acesso em: 10 abr. 2018.

FILATRO, Andréa *et al.* **DI 4.0: inovação em educação corporativa**. São Paulo: Saraiva, 2019.

FIRMIN, Sydney. **Reproducibility, Replicability, and Data Science**. Alteryx, 2019. Disponível em: <https://community.alteryx.com/t5/Data-Science-Blog/Reproducibility-Replicability-and-Data-Science/ba-p/471058>. Acesso em: 11 nov. 2019.

FLICK, Uwe. **Desenho da pesquisa qualitativa**. Porto Alegre: Bookman: Artmed, 2009.

FOX, Peter, HENDLER, James. E-science semântica: o significado codificado na próxima geração de ciência digitalmente aprimorada. *In*: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (org.). **O quarto paradigma: descobertas científicas na era da e-science**. São Paulo: Oficina de Textos, 2011.

FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. **Métodos de pesquisa para internet**. Porto Alegre: Sulina, 2011. Disponível em: <https://bit.ly/2WyXE0l>. Acesso em: 10 maio 2019.

FREIRE, Isa Maria. Dinâmica das ações de informação no regime de informação do Laboratório de Tecnologias Intelectuais. **Prisma.com**, Porto, n. 35, p. 3-21, 2017. DOI: 10.21747/16463153/35a1.

FREIRE, Paulo. **Pedagogia da Autonomia: saberes necessários à prática educativa**. São Paulo: Paz e Terra, 1996.

FROHMANN, B. Revisiting “what is a document?” **Journal of documentation**, v. 65, n. 2, p. 291- 303, 2009.

FULLER, Buckminster. **Critical Path**. New York: St. Martin’s Press, 1981.

FUNDAÇÃO OSWALDO CRUZ (Fiocruz). Ciência Aberta. *In*: **Glossário do Acesso Aberto**. 2019. Disponível em: <https://portal.fiocruz.br/en/glossario-acesso-aberto/C>. Acesso em: 10 maio 2019.

GALBRAITH, J. K. **The new industrial State**. New York: Mentor Books, 1979.

GAMA, Ana Claudia Soares Cavalcante. **Competência informacional: aprendizado individual ao longo da vida**. 2013. 509 f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2013. Disponível em: <https://bit.ly/2Mh6lly>. Acesso em: 10 maio 2019.

GARVEY, William D. **Communication: the essence of science: facilitating information among librarians scientists, engineers and students**, Oxford: Pergamon Press, 1979.

GIARETTA, David. **Insight into issues of Permanent Access to the Records of Science in Europe**. [S.l.]: European Commission, 2010. Project Final Report. Disponível em: <https://cordis.europa.eu/project/id/223758/reporting>. Acesso em: 10 abr. 2018.

GIBBS, Wayt. Lost Science in the Third World. **Scientific American**, [S.l.], p. 92-99, 1995. Disponível em: <https://www.researchgate.net/publication/259056309>. Acesso em: 10 ago. 2019.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GLASER, B. G.; STRAUSS, A. **Discovery of Grounded Theory**: Strategies for Qualitative Research. Sociology Press, 1967.

GOULDING, Christina. Grounded Theory: some reflections on paradigm, procedures and misconceptions. **Working Paper Series**, Telford, Number WP006/99, June, 1999. Disponível em: <https://bit.ly/2lbvMGv>. Acesso em: 11 maio 2019.

GRAHAM, Dan. Academic Publishing Survey of funders supports the benign Open Access outcome priced into shares. *In*: HSBC BANK. **HSBC Global Research**. Europe: HSBC Bank, 2013. Disponível em: <https://bit.ly/2VYZxyX>. Acesso em: 18 abr. 2019.

GRAY, Jim. Jim Gray on eScience: A Transformed Scientific Method. *In*: HEY; Tony; TANSLEY, Stewart; TOLLE, Kristin. **The fourth paradigm**: Data-Intensive Scientific Discovery. Redmond: Microsoft Research, 2009. p. XVII-XXXI.

GREEN, Ann; MACDONALD, Stuart; RICE, Robin. **Policy-making for Research Data in Repositories**: A Guide. Edinburgh: EDINA and University Data Library, 2009. Disponível em: <https://goo.gl/sdyd2p>. Acesso em: 20 jul. 2017.

GREEN, Daron. Infraestrutura científica: introdução. *In*: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (org.). **O quarto paradigma**: descobertas científicas na era da e-Science. São Paulo: Oficina de Textos, 2011. p. 129-130.

GRIFFITH UNIVERSITY. **Open Scholarship**. 2019. Disponível em: <https://bit.ly/2vKf4qp>. Acesso em: 10 abr. 2019.

GÜNTHER, H. Pesquisa qualitativa versus pesquisa quantitativa: Esta é a questão? **Revista Psicologia**: Teoria e Pesquisa, [S.l.], v. 22, n. 2, p. 201-210, 2006.

GÜNTHER, H.; ROZESTRATEN, R. J. A. Psicologia Ambiental: Algumas considerações sobre sua área de pesquisa e ensino. **Psicologia**: Teoria e Pesquisa, Brasília, v. 9, n. 1, p. 107-122, 1993. Disponível em: <https://psiambiental.files.wordpress.com/2016/03/1993-gunther-rozenstraten-psi-amb.pdf>. Acesso em: 10 maio 2019.

HAGSTROM, Warren O. Gift-giving as an organizing principle in science. *In*: BARNES, B. (Ed.). **Sociology of Science**. Harmondsworth: Penguin, 1972.

HAGSTROM, Warren O. **The Scientific Community**. New York: Basic Books, 1965. 304 p.

HAGUETTE, Teresa Maria Frota. **Metodologias qualitativas na sociologia**. 12. ed. Petrópolis, RJ: Vozes, 2010.

HAIG, B. Grounded Theory as Scientific Method. *In*: **PHILOSOPHY OF EDUCATION**, 1995. Disponível em: <https://bit.ly/2KiYlie>. Acesso em: 10 maio 2019.

HARDY, Georgina; CORRALL, Sheila. Revisiting the subject librarian: A study of English, Law and Chemistry. **Journal of Librarianship and Information Science**, [S.l.], v. 39, n. 2, p. 79-91, jun. 2007.

HARZING, Ann-Wil. **Publish or Perish**. 2007. Disponível em: <https://bit.ly/2sTJZiC>. Acesso em: 10 maio 2018.

HERNÁNDEZ SAMPIERI, Roberto; FERNÁNDEZ COLLADO, Carlos; BAPTISTA LUCIO, María del Pilar. **Metodologia de pesquisa**. 5. ed. Porto Alegre: Penso, 2013.

HEY, Tony; HEY, Jessie. e-Science and its implications for the library community. **Library Hi Tech**, Bingley, v. 24, n. 4, p. 515-528, 2006. Disponível em: <https://bit.ly/2WvJeOC>. Acesso em: 10 maio 2018.

HEY, Tony; TREFETHEN, Anne. **Cyberinfrastructure for e-Science**. Science, Washington, v. 308, n. 5723, p. 817-821, maio. 2005. Disponível em: <https://bit.ly/30UNrKT>. Acesso em: 10 maio 2018.

HILÁRIO, Carla Mara; GRÁCIO, Maria Cláudia Cabrini; GUIMARÃES, José Augusto Chaves. Aspectos éticos da coautoria em publicações científicas. **Em Questão**, Porto Alegre, v. 24, n. 2, p. 12-36, 2018. Acesso em: 10 out. 2018. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/76312>. Acesso em: 10 jan. 2019.

HOHENDORFF, Jean Von. Como escrever um artigo de revisão de literatura. *In*: KOLLER, Sílvia H.; COUTO, Maria Clara P. de Paula; HOHENDORFF, Jean Von (org.). **Manual de produção científica**. Porto Alegre: Penso, 2014. cap. 2, p. 39-54.

HURD, Julie M. The Transformation of Scientific Communication: A Model for 2020. **Journal of the American Society for Information Science**, [S.l.], v. 51, n. 14, p. 1279-1283, 2000.

INGWERSEN, Peter; JÄRVELIN, Kalervo. **The turn: integration of information seeking and retrieval in context**. New York: Springer-Verlag, 2005.

INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH. **Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle**. 5th ed. Ann Arbor: ICPRS, 2012.



INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã**. Brasília: Ibict, 2016. Disponível em: <https://bit.ly/2wuE4Us>. Acesso em: 10 set. 2018.

JENSEN, K. B.; JANKOWSKI, N. M. **Metodologias cualitativas de investigación en comunicación de masas**. Barcelona: Bosch, 1993.

JENKINS, Paul O. **Faculty-Librarian Relationships**. Oxford: Chandos Publishing, 2005.

JOINT TASK FORCE ON LIBRARIANS' COMPETENCIES IN SUPPORT OF E-RESEARCH AND SCHOLARLY COMMUNICATION. **Librarians' Competencies for E-Research and Scholarly Communication**. In: Facilitate Open Science Training for European Research (FOSTER). 2016. Disponível em: <https://goo.gl/kUVJai>. Acesso em: 1 set. 2017.

JUNQUEIRA, Eduardo S. **Tutores em EAD: teorias e práticas**. Fortaleza: Dummar, 2018.

KNÖCHELMANN, Marcel. Open Humanities: Why Open Science in the Humanities is not Enough. **LSE Blog**, [S.l.], 2020. Disponível em: <https://blogs.lse.ac.uk/impactofsocialsciences/2020/02/25/open-humanities-why-open-science-in-the-humanities-is-not-enough/>. Acesso em: 25 fev. 2020

KOZINETS, Robert V. **Netnografia: Realizando Pesquisa Etnográfica Online**. Porto Alegre: Penso, 2014.

KOZINETS, Robert V. **Netnography 2.0: Handbook of Qualitative Research Methods in Marketing**. Cheltenham, UN e Northampton, MA: Edward Elgar Publishing, 2006. p.129-142.

KOZINETS, Robert. Utopian enterprise: Articulating the meanings of Star Trek's culture of consumption. **Journal of Consumer Research**, [S.l.], v. 28, n. 1, p. 67, jun. 2002.

KRAKER, P. *et al.* **The Vienna Principles: A Vision for Scholarly Communication in the 21st Century**. 2016. Zenodo. Disponível em: <https://bit.ly/2JNyCEE>. Acesso em: 10 mar. 2019.

KRAKER, Peter *et al.* **Open Knowledge Maps: A Visual Interface to the World's Scientific Knowledge**. 2017. Disponível em: <https://bit.ly/2K9PNzE>. Acesso em: 10 maio 2019.

KUHLTHAU, Carol. Inside the Search Process: Information Seeking from the User's Perspective. **Journal of the American Society for Information Science**, New York, v. 42, n. 5, p. 361-371, 1991. Disponível em: <https://tinyurl.com/y2nmr7yn>. Acesso em: 7 set. 2018.

KUHLTHAU, Carol. **Seeking meaning**: a process approach to library and information services. Norwood: Ablex, 1993.

LADRIÈRE, Jean. Prefácio. *In*: BRUYNE, Paul de; HERMAN, Jacques; SCHOUTHEETE, Marc de. **Dinâmica da pesquisa em Ciências Sociais**: os pólos da prática metodológica. 2. ed. Rio de Janeiro: Francisco Alves, 1977. p. 17.

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília: Briquet de Lemos, 2004.

LEAGUE OF EUROPEAN RESEARCH UNIVERSITIES. **Open Science and its role in universities**: A roadmap for cultural change. Leuven: LERU, 2018. Disponível em: <https://www.leru.org/files/LERU-AP24-Open-Science-full-paper.pdf>. Acesso em: 10 mar. 2019.

LEHMKUHL, Karyn Munky; GAVRON, Edson Mário; ALVES, Maria Bernardete Martins; MORAES DE BEM, Roberta. Suporte à Pesquisa e Gestão de Dados Científicos em Bibliotecas Universitárias: experiência da BU/UFSC. *In*: GRANTS, Andréa Figueiredo Leão; MORAES DE BEM, Roberta (org.). **A Construção de Saberes**: Protagonismo compartilhado em serviços e inovações na Biblioteca Universitária da UFSC. Florianópolis: BU Publicações, 2018. cap.7, p. 142-158.

LÉVI-STRAUSS, Claude. Aula Inaugural. *In*: ZALUAR, Alba (org.). **Desvendando máscaras sociais**. Rio de Janeiro: Francisco Alves, 1975. p. 211-244.

LIMA, Juliana Soares; FARIAS, Maria Giovanna Guedes. Autoria em produções científicas: conceitos, critérios, integridade na pesquisa e responsabilidade na colaboração. **Investig. bibl**, México, v. 34, n. 82, p. 103-139, mar. 2020. Disponível em: <https://bit.ly/35k1DzH>. Acesso em: 20 mar. 2020.

LITTO, Frederic; FORMIGA, Marcos (org.). **Educação a distância**: o estado da arte. São Paulo: Pearson, 2009.

LIU, David; SALGANIK, Matthew. Successes and struggles with computational reproducibility: Lessons from the Fragile Families Challenge. **SocArXiv**, 2019. DOI: 10.31235/osf.io/g3pdb.

LOGAN, Robert K. **Que é informação?**: a propagação da informação na biosfera, na simbolosfera, na tecnosfera e na econosfera. Rio de Janeiro: Contraponto, 2012.

LUCE, Richard E. E-Science and Research Libraries: Na Agenda for action. *In*: MARCUM, Deanna B.; GEORGE, Gerald. **The data deluge**: can libraries cope with E-Science? Santa Barbara: ABC-CLIO, 2010. cap. 1.

LÜDKE, Menga.; ANDRÉ, Marli E. D. A. **Pesquisa em educação**: abordagens qualitativas. São Paulo: EPU, 1986.

MARCONDES, Carlos Henrique; SAYÃO, Luís Fernando. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a

proposta da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v. 30, n. 3, p. 24-33, dez. 2001. Disponível em: <https://bit.ly/2MpkV14>. Acesso em: 10 set. 2017.

MARCUM, Deanna B.; GEORGE, Gerald. **The data deluge**: can libraries cope with E-Science? Santa Barbara: ABC-CLIO, 2010.

MARQUES, Fabrício. Mecanismos de autocorreção da ciência. **Pesquisa FAPESP**, São Paulo, v. 280, p. 8-10, jun. 2019. Disponível em: <https://revistapesquisa.fapesp.br/2019/06/07/mecanismos-de-autocorrecao-da-ciencia/>. Acesso em: 5 ago. 2019.

MARTÍNEZ-URIBE, Luis; MACDONALD, Stuart. Un nuevo cometido para los bibliotecarios académicos: data curation. **El Profesional de la Información**, [S.l.], v. 17, n. 3, p. 273–280, 2008. DOI: 10.3145/epi.2008.may.03.

MARTÍN-MARTÍN, A. *et al.* Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. **Journal of Informetrics**, [S.l.], v. 12, n. 4, p. 1160-1177, nov. 2018. DOI: 10.1016/J.JOI.2018.09.002. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1751157718303249>. Acesso em: 27 nov. 2019.

MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da Investigação Científica para Ciências Sociais Aplicadas**. 3. ed. São Paulo: Atlas, 2018.

MAUSS, Marcel. Ensaio sobre a dádiva. *In*: MAUSS, Marcel. **Sociologia e antropologia**. São Paulo: Cosac Naify, 2003. p. 183-314.

MAYER, Katja. **Open Innovation and the Creation of Commons**. [S.l.]: Creative Commons, 2016. Disponível em: <https://creativecommons.org/2016/06/21/open-innovation-creation-commons/>. Acesso em: 10 ago. 2019.

MCKIERNAN, Erin C. *et al.* How open science helps researchers succeed. **eLife**, [S.l.], 2016. DOI: 10.7554/eLife.16800.

MEADOWS, Arthur J. **A comunicação científica**. Brasília: Briquet de Lemos, 1999.

MEDEIROS, Jackson da Silva; CAREGNATO, Sônia Elisa. Compartilhamento de dados de e-science: explorando um novo conceito para a comunicação científica. **Liinc em Revista**, Rio de Janeiro, v. 28, n. 2, p. 311-322, set. 2012.

MEKSENAS, P. Aspectos Teóricos e Metodológicos da pesquisa empírica: a contribuição da obra de Paulo Freire. *In*: SEMANA DE ESTUDOS FREIRIANOS, 1., 2007, Joinville. **Anais [...]** Joinville: [s.n.], p. 6-15.

MENDONÇA, J. R. C. de. **Interacionismo Simbólico**: uma sugestão metodológica para a pesquisa em administração. 2001. Disponível em: <https://bit.ly/2WupNWh>. Acesso em: 10 maio 2019.

MENEZHINI, R. Avaliação da produção científica e o Projeto SciELO. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 219-20, 1998.

MERTON, Robert King. **On the shoulders of giants**: a Shandean postscript. Chicago: University of Chicago Press, 1993.

MERTON, Robert King. **Sociologia**: teoria e estrutura. São Paulo: Mestre Jou, 1968.

MERTON, Robert King. The Matthew Effect in Science. *In*: MERTON, Robert King. **Sociology of science**: theoretical and empirical investigations. Chicago: University of Chicago Press, 1973a. Cap. 20.

MERTON, Robert King. The normative structure of science. *In*: MERTON, Robert King. **The Sociology of Science**: Theoretical and Empirical Investigations, University of Chicago Press, Chicago 1973b. Cap. 13.

MESGUICH, Véronique *et al.* Enjeux et dimensions. **Documentaliste-Sciences de l'Information**, [S.l.], v. 49, n. 1, p. 24-45, 2012. DOI: 10.3917/docsi.491.0024. Disponível em: <https://is.gd/jLHmq3>. Acesso em: 11 fev. 2019.

MINAYO, Maria Cecília de Souza. **O desafio do conhecimento**: pesquisa qualitativa em saúde. Rio de Janeiro: Hucitec, 2008.

MINAYO, Maria Cecília de Souza; MINAYO-GÓMEZ, C. Díficeis e possíveis relações entre métodos quantitativos e qualitativos nos estudos de problemas de saúde. *In*: GOLDENBERG, P.; MARSIGLIA, R. M. G.; GOMES, M. H. A. (org.). **O clássico e o novo**: tendências, objetos e abordagens em ciências sociais e saúde. Rio de Janeiro, Fiocruz, 2003. p. 117-142.

MIRANDA, Roberto Campos da Rocha. O uso da informação na formulação de ações estratégicas pelas empresas. **Ciência da Informação**, Brasília, v. 28, n. 3, p. 286-292, dez. 1999. Disponível em: <https://bit.ly/3c1X0gc>. Acesso em: 10 jan. 2018

MIRANDA, Silvânia. Como as necessidades de informação podem se relacionar com as competências informacionais. **Ciência da Informação**, Brasília, v. 35, n. 3, p. 99-114, 2006. Disponível em: <https://www.scielo.br/pdf/ci/v35n3/v35n3a10>. Acesso em: 10 jan. 2018.

MON, Lorri; HARRIS, Lydia Eato. The death of the anonymous librarian. **The Reference Librarian**, v. 52, p. 352-364, 2011. DOI: 10.1080/02763877.2011.585279.

MORGAN, Ann; DUFFIELD, Nel; HALL, Liz Walkley. Research Data Management Support: Sharing Our Experiences. **Journal of the Australian Library and Information Association**, [S.l.], v. 66, n. 3, p. 299-305, 2017. Disponível em: <https://bit.ly/2MKH0m4>. Acesso em: 10 fev. 2019.

MORIN, Edgar. **Ciência com consciência**. Rio de Janeiro: Bertrand Brasil, 1998.

MORIN, Edgar. **Os sete saberes necessários à educação do futuro**. São Paulo: Cortez; Brasília: UNESCO, 2001.

MORSE, J. Designing Funded Qualitative Research. *In*: DENZIN, N. K.; LINCOLN, Y. S. **Handbook of Qualitative Research**. Thousand Oaks: Sage, 1994. p. 220-235.

MOSER, Gabriel. Psicologia Ambiental. **Estudos de Psicologia (Natal)**, Natal, v. 3, n. 1, p. 121-130, jun. 1998. DOI: <http://dx.doi.org/10.1590/S1413-294X1998000100008>. Disponível em: <http://twixar.me/SHST>. Acesso em: 10 maio 2019.

MOULIER BOUTANG, Yann. Interview to Gaele Krikorian. *In*: KRIKORIAN, Gaele; KAPCZYNSKI, Amy. **Access to knowledge in the Age of Intellectual Property**. New York: Zone Book, 2010. p. 575-596.

MOURA, Maria Aparecida. Prefácio. *In*: FERREIRA, Valdinéia Barreto. **E-Science e Políticas Públicas: ciência, tecnologia e inovação no Brasil**. Salvador: EDUFBA, 2018. 256 p.

MUELLER, Suzana Pinheiro Machado. O impacto das tecnologias de informação na geração do artigo científico: tópicos para estudo. **Ciência da Informação**, Brasília, v. 23, n. 3, p. 309-317, 1994. Disponível em: <https://tinyurl.com/y67m8xyt>. Acesso em: 10 dez. 2018.

MURRAY-RUST, Peter *et al.* **Panton Principles**: Principles for Open Data in Science. 2010. Disponível em: <https://pantonprinciples.org/>. Acesso em: 10 mar. 2019.

NATIONAL INSTITUTES OF HEALTH. **Guidelines for Scientific Record Keeping in the Intramural Research Program at the NIH**. [S.l.]: NIH, 2008. Disponível em: [https://oir.nih.gov/sites/default/files/uploads/sourcebook/documents/ethical\\_conduct/guidelines-scientific\\_recordkeeping.pdf](https://oir.nih.gov/sites/default/files/uploads/sourcebook/documents/ethical_conduct/guidelines-scientific_recordkeeping.pdf). Acesso em: 10 maio 2019.

NELSON, B. Data sharing: Empty archives. **Nature**, [S.l.], v. 461, n. 7261, p. 160-163, 2009. DOI: 10.1038/461160a.

NELSON, R. R. The simple economics of basic scientific research. **Journal of Political Economy**, [S.l.], v. 67, p. 297-306, 1959.

NICHOLAS, D. *et al.* Peer review: still king in the digital age. **Learned Publishing**, [S.l.], v. 28, p. 15-21, 2015. DOI:10.1087/20150104.

NIELSEN, Hans Jørn; HJØRLAND, Birger. Curating research data: the potential roles of libraries and information professionals. **Journal of Documentation**, [S.l.], v. 70, n. 2, p.221-240, 2014. DOI: <https://doi.org/10.1108/JD-03-2013-0034>.

NIELSEN, Michael. **Open Science Now!** 2011a. Palestra proferida durante o TEDx Waterloo. Disponível em: <https://bit.ly/2wpbML5>. Acesso em: 22 abr. 2011.

NIELSEN, Michael. **Reinventing Discovery**: The New Era of Networked Science. New Jersey: Princeton University Press, 2011b.

NOSEK, Brian; CHAMBERS, Chris. The first imperative: Science that isn't transparent isn't science. **The Guardian**, News, 2015. Não paginado. Disponível em: <https://www.theguardian.com/science/head-quarters/2015/jun/25/the-first-imperative-science-that-isnt-transparent-isnt-science>. Acesso em: 10 ago. 2019.

OLIVEIRA, Adriana Silva de Oliveira. **Desvendando a autoraldade colaborativa na e-science sob A ótica dos direitos de propriedade intelectual**. Orientador: Guilherme Ataíde Dias. 2016. 300 f. Dissertação (Doutorado em Ciência da Informação) - Universidade Federal da Paraíba, João Pessoa, 2016. Disponível em: <https://bit.ly/2HJ9fBH>. Acesso em: 10 abr. 2019.

OLIVEIRA, Djalma de Pinho Rebouças. **Sistemas, organizações e métodos: uma abordagem gerencial**. 13. ed. São Paulo, 2002.

OLIVEIRA, Eloísa da Conceição Príncipe de. Percursos digitais da comunicação científica. In: BRAGA, Gilda Maria; PINHEIRO, Lena Vânia Ribeiro (org.). **Desafios do impresso ao digital: questões contemporâneas de informação e conhecimento**. Brasília: Ibict, 2009. p. 289-312.

OLIVEIRA, Marcos Barbosa de. Formas de autonomia da ciência. **Sci. stud.**, São Paulo, v. 9, n. 3, p. 527-561, 2011. Disponível em: <https://bit.ly/2JI0uKo>. Acesso em: 10 jan. 2019.

OLSON, G. M.; ZIMMERMAN, A.; BOS, N. **Scientific collaboration on the Internet**. Cambridge: MIT Press, 2008.

OPEN GALLERIES, LIBRARIES, ARCHIVES AND MUSEUMS. **OpenGLAM**. [S.l. : s.n.], 2019. Disponível em: <https://openglam.org/>. Disponível em: 10 maio 2019.

OPEN KNOWLEDGE INTERNATIONAL FOUNDATION. **Open data handbook**. 2017. Disponível em: <https://bit.ly/2WyYk5T>. Acesso em: 7 set. 2017.

OPEN KNOWLEDGE INTERNATIONAL FOUNDATION. **What is Open?** Disponível em: <https://okfn.org/opendata/>. Acesso em: 7 set. 2018.

OPEN RESPONSIBLE RESEARCH AND INNOVATION TO FURTHER OUTSTANDING KNOWLEDGE. **Analysis and Benchmarking: Self-assessment. Questionnaire**. 2017. Disponível em: <https://bit.ly/2MCN4lt>. Acesso em: 7 set. 2018.

OPEN RESPONSIBLE RESEARCH AND INNOVATION TO FURTHER OUTSTANDING KNOWLEDGE. **What is Open Science?** 2018. Disponível em: <https://bit.ly/2XmQbi7>. Acesso em: 7 set. 2018.

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. **Making Open Science a Reality**. OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing: Paris, 2015. DOI: <https://doi.org/10.1787/5jrs2f963zs1-en>.

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding**.

Paris: OECD, 2007. Disponível em: <http://www.oecd.org/sti/sci-tech/38500813.pdf>. Acesso em: 20 jul. 2017.

ORION OPEN SCIENCE. **Public attitudes to life sciences research in six European countries**. Stockholm: European Union, 2018. Disponível em: <https://bit.ly/35iVS5F>. Acesso em: 13 mar. 2019.

OTLET, P. **Traité de documentatation**: le livre sur le livre: théorie et pratique. Bruxelles: Editiones Mundaneum, 1934.

PACKER, Abel Laerte. A construção coletiva da Biblioteca Virtual em Saúde. **Interface**, Botucatu, v. 9, n. 17, p. 249-72, mar./ago. 2005.

PADOVEZE, Clóvis Luís. **Sistemas, organizações e métodos**: uma abordagem gerencial. 13. ed. São Paulo, 2002.

PERKINS, David. Regulating the Internet Giants: The world's most valuable resource is no longer oil, but data. **The Economist**, [S.l.], v. 413, n. 9035, 2017. Não paginado. Disponível em: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. Acesso em: 10 mar. 2019.

PHILLIP, Ryan. **Keeping a Lab Notebook**: Basic Principles and Best Practices. [S.l. : s.n.], [200-?]. Disponível em: <http://twixar.me/FJXT>. Acesso em: 10 maio 2019.

PIDGEON, N. Grounded Theory: theoretical background. *In*: RICHARDSON, J. **Handbook of Qualitative Research Methods**. Oxford: BPS Blackwell, 1996. Chapter 6.

PINFIELD, S.; COX, A. M.; SMITH, J. Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. **PLoS ONE**, v. 9, n. 12, p. e114734, 2014. DOI: <https://doi.org/10.1371/journal.pone.0114734>.

PINO SIRGADO, A. O social e o cultural na obra de Vigotski. **Revista Educação e Sociedade**, Campinas, v. 21, n. 71, p. 45-78, jul. 2000. DOI: <http://dx.doi.org/10.1590/S0101-73302000000200003>.

POLANYI, Michael; ZIMAN, John; FULLER, Steve. The republic of science: Its political and economic theory. **Minerva**, [S.l.], v. 38, n. 1, p. 1-32, 2000.

POPPER, Karl. **Conhecimento objetivo**: uma abordagem evolucionária. Belo Horizonte: Itatiaia, 1999.

PRINS, A. A. M. *et al.* Using Google Scholar in research evaluation of humanities and social science programs: A comparison with Web of Science data. **Research Evaluation**, [S.l.], v. 25, n. 3, p. 264-270, jul. 2016. DOI: 10.1093/reseval/rvv049. Disponível em: <https://academic.oup.com/rev/article/25/3/264/2364634>. Acesso em: 27 nov. 2019.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico**: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed.

Novo Hamburgo: Universidade Feevale, 2013. Disponível em: <https://bit.ly/2K9Hy6V>. Acesso em: 11 maio 2019.

REDE DE DADOS DE PESQUISA BRASILEIRA. **Sobre nós**. 2019. Disponível em: [https://dadosdepesquisa.rnp.br/sobre\\_nos/](https://dadosdepesquisa.rnp.br/sobre_nos/). Acesso em: 10 maio 2019.

RESEARCH DATA NETHERLANDS. Essentials 4 Data Support. **Competencies**. 2019. Disponível em: <https://bit.ly/2WwNWeT>. Acesso em: 10 mar. 2019.

RICHARDSON, Roberto Jarry. **Pesquisa Social: métodos e técnicas**. 4. ed. São Paulo: Atlas, 2017.

ROYAL SOCIETY OF LONDON. Royal Society Science Policy Centre. **Science as an open enterprise: open data for open science**. London: Royal Society of London, 2012. Disponível em: <https://bit.ly/217vgcB>. Acesso em: 20 jul. 2017.

ROZADOS, Helen Beatriz Frota. Indicadores como ferramenta para gestão de serviços de informação tecnológica. Porto Alegre, 2004. 238 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004. Disponível em: <https://www.lume.ufrgs.br/handle/10183/5668>. Acesso em: 20 jul. 2018.

RUSSO, Mariza. Dado x Informação x Conhecimento. *In*: RUSSO, Mariza. **Fundamentos em Biblioteconomia e Ciência da Informação**. Rio de Janeiro: E-Papers, 2010. p. 14-21.

SALDANHA, Gustavo Silva. O “fabuloso” antílope de Suzanne Briet: a análise e a crítica da análise neodocumentalista. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 13., 2012. **Anais [...]**. Rio de Janeiro: Fiocruz, 2012. Disponível em: <https://is.gd/mlqRKQ>. Acesso em: 7 set. 2018.

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 - Enancib, 19., 2018. **Anais [...]**. Londrina: Enancib, 2018. Não paginado. Disponível em: <https://bit.ly/2WafXcZ>. Acesso em: 10 maio 2019.

SALES, Luana Farias. **Integração semântica de publicações científicas e dados de pesquisa: proposta de modelo de publicação ampliada para a área de ciências nucleares**. Rio de Janeiro, 2014. 268 f. Tese (Doutorado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: <https://bit.ly/30Skpf3>. Acesso em: 26 mai. 2018.

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, dez. 2016. Disponível em: <https://bit.ly/2XhwhFo>. Acesso em: 26 maio 2019. DOI: <http://dx.doi.org/10.5433/1981-8920.2016v21n2p116>.



SANTOS, Anne Danielle Soares Clinio dos. **Novos cadernos de laboratório e novas culturas epistêmicas**: entre a política do experimento e o experimento da política. Rio de Janeiro, 2016. 240 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Escola de Comunicação. Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, 2016. Disponível em: <https://bit.ly/2HLMetb>. Acesso em: 11 maio 2019.

SANTOS, Boaventura de Sousa. **A gramática do tempo**: para uma nova cultura política. 2. ed. São Paulo: Cortez, 2008. v. 4.

SANTOS, Boaventura de Sousa. A Praxis da Ecologia de Saberes: entrevista de Boaventura de Sousa Santos. **Tempus: actas de Saúde Coletiva**, Brasília, v. 8, n. 2, p. 331-338, 2014. Entrevista concedida a Fernando Ferreira Carneiro, Noemi Margarida Krefta e Cleber Adriano Rodrigues Folgado. Disponível em: <https://bit.ly/2MjuheJ>. Acesso em: 10 jan. 2018.

SANTOS, Boaventura de Sousa. Para além do pensamento abissal: das linhas globais a uma ecologia de saberes. **Novos estud. - CEBRAP**, São Paulo, n. 79, p. 71-94, nov. 2007. doi: 10.1590/S0101-33002007000300004. Disponível em: <https://bit.ly/2VV5A7O>. Acesso em: 10 jan. 2018.

SANTOS, Boaventura de Sousa. Um ocidente não-ocidentalista? A filosofia à venda, a douda ignorância e a aposta de Pascal. *In*: Santos, Boaventura de Sousa, Meneses, Maria Paula (org.). **Epistemologias do Sul**. São Paulo: Cortez, 2010.

SANTOS, Francisco Edvander Pires; LIMA, Juliana Soares; ARAÚJO, Irlana Mendes de. **Mediação de bibliotecários liaison no ambiente universitário**. 2019. No prelo.

SANTOS, Solange Maria. Rankings Internacionais de Universidades: Comparação e Desempenho por Áreas. *In*: MARCOVITCH, Jacques (org.). **Repensar a universidade**: desempenho acadêmico e comparações internacionais. São Paulo: Com-Arte; Fapesp, 2018. Parte 1, cap. 4, p. 63-92. Disponível em: <https://bit.ly/2VWKx4R>. Acesso em: 10 abr. 2019.

SANTOS NETO, João Arlindo dos; ALMEIDA JÚNIOR, Oswaldo Francisco de. A competência em informação e o bibliotecário mediador da informação na biblioteca universitária. *In*: BELLUZZO, Regina Célia Baptista; FERES, Glória Georges; VALENTIM, Marta Lígia Pomim (org.). **Redes de conhecimento e competência em informação**: interfaces da gestão, mediação e uso da informação. Rio de Janeiro: Interciência, 2015. cap. 12, p. 359-376.

SAYÃO, Luís Fernando; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Inf. & Soc.:Est.**, João Pessoa, v. 22, n. 3, p. 179-191, set./dez. 2012.

SAYÃO, Luís Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. 90 p. Disponível em: <https://bit.ly/2HDiO1I>. Acesso em: 20 jul. 2017.

SCHELIGA, Kaja; FRIESIKE, Sascha. Putting open science into practice: A social dilemma?. **First Monday**, v. 19, n. 9, 2014. Não paginado. DOI: 10.5210/fm.v19i9.5381.

SCHIELD, Milo. Information Literacy, Statistical Literacy and Data Literacy. **IASSIST Quarterly**, [S.l.], Summer/Fall, p. 6-11, 2004. Disponível em: <https://bit.ly/30SFzK5>. Acesso em: 10 maio 2018.

SCHILLER, Dan. Capitalismo digital em crise. *In*: MACIEL, Maria Lucia; ALBAGLI, Sarita. **Informação, conhecimento e poder**: mudança tecnológica e inovação social. Rio de Janeiro: Garamond, 2011. p. 133-149.

SCHMIDT, Birgit *et al.* Time to Adopt: Librarians' New Skills and Competency Profiles. *In*: LOIZIDES, Fernando; SCHMIDT, Birgit. **Positioning and Power in Academic Publishing: Players, Agents and Agendas**. Amsterdam; Berlin; Washington: IOS Press, 2016. p. 1-8. Disponível em: <https://bit.ly/2QvGknE>. Acesso em: 7 set. 2017.

SCHMIDT, Birgit; SHEARER, Kathleen. **Librarians' Competencies Profile for Research Data Management**. 2016. Disponível em: <https://bit.ly/2Wrpk7t>. Acesso em: 20 jul. 2017.

SETZER, Valdemar W. Dado, Informação, Conhecimento e Competência. **DataGramaZero - Revista de Ciência da Informação**, Brasília, v. 0, n. 0, dez. 1999. Não paginado. Disponível em: <https://tinyurl.com/y38yu8vo>. Acesso em: 10 jun. 2018.

SHAREWEB LEARNING & NETWORK. **Knowledge Map**. 2019. Disponível em: <https://bit.ly/2VZzVSw>. Acesso em: 10 maio 2019.

SHEARER, Kathleen. **Comprehensive Brief on Research Data Management Policies**. Canada Government, 2015. Disponível em: <https://portagenetwork.ca/wp-content/uploads/2016/03/Comprehensive-Brief-on-Research-Data-Management-Policies-2015.pdf>. Acesso em: 10 jan. 2018.

SHERA, Jesse H. Toward a theory of Librarianship and information science. **Ciência da Informação**, Brasília, v. 2, n. 2, dez. 1973. Disponível em: <http://revista.ibict.br/ciinf/article/view/30>. Acesso em: 2 jun. 2018.

SILVA, Andréa Soares Rocha da *et al.* Possibilidades de uso do Google Classroom como ferramenta de apoio à educação permanente em saúde. *In*: LIMA, Tamires Feitosa de *et al.* (org.). **Abordagens metodológicas não convencionais em pesquisa**. Paraná: Booknando, 2019. *Ebook*. Não paginado. Locations 75%.

SILVA, Armando Malheiro da. Mediações e mediadores em Ciência da Informação. **Prisma.com**, Porto, n. 9, 2009. Disponível em: <https://bit.ly/2WCAUwn>. Acesso em: 10 jun. 2018.

SILVA, Armando Malheiro da; RIBEIRO, Fernanda. **Paradigmas, serviços e mediações em Ciência da Informação**. Recife: Nectar, 2011.

SILVA, Fabiano Couto Corrêa da. O papel dos bibliotecários na gestão de dados científicos. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 14, n. 3, p. 387, 2016. DOI:10.20396/rdbci.v14i3.8646333.

SILVA, Jonathas Luiz; FARIAS, Maria Giovanna Guedes. Abordagens conceituais e aplicativas da mediação nos serviços de informação. **InCID: Revista de Ciência da Informação e Documentação**, v. 8, n. 2, p. 106-123, 4 out. 2017. DOI: 10.11606/issn.2178-2075.v8i2p106-123.

SODERBERG, Johan. **Hacking Capitalism: the Free and Open Source Software Movement**. New York and London: Routledge, 2008.

SOEHNER, Catherine; STEEVES, Catherine; WARD, Jennifer. **E-science and data support services: a study of ARL member institutions**. Washington, DC: Association of Research Libraries, 2010. Disponível em: <https://bit.ly/2EKiLmh>. Acesso em: 10 maio 2018.

SOLLA-PRICE, D. J. de. **O Desenvolvimento da Ciência: análise histórica, filosófica, sociológica e econômica**. Rio de Janeiro: Livros Técnicos e Científicos, 1976.

SQUAZZONI, F. *et al.* **Data Sharing and Research on Peer Review: A Call to Action**, [S.l. : s.n.], 2019. Disponível em: [osf.io/preprints/socarxiv/sr6eg](https://osf.io/preprints/socarxiv/sr6eg). DOI: 10.31235/osf.io/sr6eg.

STERNBERG, R. J. A balance theory of wisdom. **Review of General Psychology**, [S.l.], v. 2, n. 4, p. 347-365, 1998.

STEPHAN, Paula E. The Economics of Science. **Journal of Economic Literature**, [S.l.], v. 34, n. 3, p.1199-1235,1996.

STRAUSS, Anselm L.; CORBIN, Juliet. **Basics of qualitative research: Grounded theory procedures and techniques**. Newbury Park, CA: Sage, 1990.

STRAUSS, Anselm L.; CORBIN, Juliet. **Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento da teoria fundamentada**. 2. ed. Porto Alegre: Artmed, 2008.

SUBER, Peter. 2007. Creating an intellectual commons through open access. *In*: HESS, Charlotte; OSTROM, Elinor. **Understanding knowledge as a commons: from theory to practice**. Cambridge: MIT Press, 2007. p. 171-208. Disponível em: [http://www.wtf.tw/ref/hess\\_ostrom\\_2007.pdf](http://www.wtf.tw/ref/hess_ostrom_2007.pdf). Acesso em: 03 dez. 2018.

SWARTZ, Aaron. **Guerrilla Open Access Manifesto**. 2008. Disponível em: <https://tinyurl.com/lwbmuz2>. Acesso em: 7 set. 2018.

TARGINO, Maria das Graças. Orientador ou tutor é autor?. **Informação & Informação**, Londrina, v. 15, n. 1, p. 145-156, 2010. DOI: 10.5433/1981-8920.2010v15n1espp145.

TAROZZI, M. **O que é a Grounded Theory? Metodologia de pesquisa e de teoria fundamentada nos dados**. Petrópolis: Vozes, 2011.

TENNANT, Jonathan *et al.* **Foundations for Open Scholarship Strategy Development**. 2019. MetaArXiv. DOI:10.31222/osf.io/b4v8p. Disponível em: <https://osf.io/preprints/metaarxiv/b4v8p>. Acesso em: 10 abr. 2019.

TENNANT, Jonathan. **Open Science is just good science**. 2018. Figshare. Presentation. DOI: <https://doi.org/10.6084/m9.figshare.7097738.v1>.

TENNANT, Michele R.; CATALDO, Tara Tobin. Development and Assessment of Specialized Liaison Librarian Services, **Medical Reference Services Quarterly**, [S.l.], v. 21, n. 2, p. 21-37, 2002. DOI: [http://dx.doi.org/10.1300/J115v21n02\\_03](http://dx.doi.org/10.1300/J115v21n02_03).

TENOPIR, Carol; BIRCH, Ben; ALLARD, Suzie. **Academic Libraries and Research Data Services: Current Practices and Plans for the Future An ACRL White Paper**. Chicago: Association of College & Research Libraries, 2012. 55 p. Disponível em: <https://bit.ly/2D8viQG>. Acesso em: 10 maio 2018.

TENOPIR, Carol *et al.* Data Sharing by Scientists: Practices and Perceptions. **PLOS ONE**, [S.l.], v. 6, n. 6, p. 1-21, 2011. Disponível em: <https://bit.ly/2yYdaIR>. Acesso em: 10 maio 2018.

THIOLLENT, M. **Metodologia da pesquisa-ação**. 11. ed. São Paulo: Cortez, 2009.

TOFFLER, Alvin. **A terceira onda**. 21. ed. Rio de Janeiro, RJ: Record, 1995.

TRIPATHI, M.; SHUKLA, A.; SONKAR, S. Research Data Management Practices in University libraries: A study. **DESIDOC Journal of Library & Information Technology**, [S.l.], v. 37, n. 6, p. 417-424, 7 nov. 2017. Disponível em: <https://bit.ly/2EHM5g>. Acesso em: 10 fev. 2019.

TRIVIÑOS, Augusto Nivaldo Silva. **Introdução à pesquisa em Ciências Sociais: a pesquisa qualitativa em educação**. São Paulo: Atlas, 1987.

UNITED NATIONS EDUCATIONAL, SCIENCE AND CULTURAL ORGANIZATION (Unesco). **Recursos educacionais abertos (REA)**. [S.l.], 2017. Disponível em: <https://bit.ly/2JLA4aF>. Acesso em: 11 maio 2018.

UNITED NATIONS ORGANIZATION. Independent Expert Advisory Group Secretariat. Data Revolution Group. **A World that Counts: Mobilising the data revolution for sustainable development**. Geneva; New York, 2014, p. 2. Disponível em: <https://bit.ly/1wApBjP>. Acesso em: 7 set. 2017.

UNIVERSIDADE DE SÃO PAULO. Pró-Reitoria de Pesquisa. **Gestão de dados científicos**. 2019. Disponível em: <https://prp.usp.br/gestao-de-dados-cientificos/>. Acesso em: 10 mar. 2019.

UNIVERSITY OF LEEDS. Library. **Research data management explained**. 2019. Disponível em: <https://tinyurl.com/y65hx4va>. Acesso em: 10 mar. 2019.

VAN DEN EYNDEN, V.; CORTI, L. Advancing research data publishing practices for the social sciences: from archive activity to empowering researchers. **International Journal of Digital Library**, [S.l.], v. 18, n. 2, p. 113-121, 2017. Disponível em: <https://bit.ly/30VbGss>. Acesso em: 10 fev. 2019.

WALSTAD, A. Science as a market process. **The Independent Review**, [S.l.], v. 7, n. 1, p. 5-45, 2002. Disponível em: <https://bit.ly/2Wbg2gu>. Acesso em: 17 jun. 2018.

WERSIG, Gernot. Information science: the study of postmodern knowledge usage. **Information Processing and Management: an International Journal**, Tarrytown, New York, v. 29, n. 2, p. 229-239, mar./abr. 1993. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/030645739390006Y?via%3Dihub>. Acesso em: 17 jun. 2018.

WEXLER, Mark N. The who, what and why of knowledge mapping. **Journal of Knowledge Management**, [S.l.], v. 5, n. 3, p. 249-264, 2001. DOI: <https://bit.ly/2YWTrkQ>.

WIKIVERSIDADE. **Portal: Ciência Aberta**. 2017. Disponível em: <https://bit.ly/2YWbW8T>. Acesso em: 10 set. 2018.

WORLD WIDE WEB CONSORTIUM (W3C). **Data on the Web Best Practices**. 2017. Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em: 7 set. 2017.

XIA, Jingfeng; WANG, Minglu. Competencies and Responsibilities of Social Science Data Librarians: An Analysis of Job Descriptions. **College & Research Libraries**, [S.l.], v. 75, n. 3, p. 362-388, maio 2014. Disponível em: <https://bit.ly/2Z3sxl7>. Acesso em: 12 fev. 2019.

YOON, Ayoung; SCHULTZ, Teresa. Research Data Management Services in Academic Libraries in the US: A Content Analysis of Libraries' Websites. **College & Research Libraries**, [S.l.], v. 78, n. 7, p. 920, out. 2017. Disponível em: <https://bit.ly/30VelTa>. Acesso em: 10 fev. 2019.

ZIMAN, J. **Real Science: What it is, and what it means**. Cambridge University Press, Cambridge, 2000.

ZINS, Chaim. Conceptual approaches for defining 'data', 'information', and 'knowledge'. **Journal of the American Society for Information Science and Technology**, v. 58, n. 4, p. 479-493, 2007.

ZINS, Chaim. Redefinindo a Ciência da Informação: da “ciência da informação” para a “ciência do conhecimento”. **Inf. & Soc.:Est.**, João Pessoa, v. 21, n. 3, p. 155-167, set./dez. 2011. Disponível em: <https://bit.ly/2MIVhu5>. Acesso em: 10 set. 2018.

## APÊNDICE A - QUESTIONÁRIO

### Gestão de Dados de Pesquisa e Ciência Aberta: Percepção dos Pesquisadores da Universidade Federal do Ceará

Prezado(a) Pesquisador(a),

Você está sendo convidado(a) a participar da pesquisa intitulada "Gestão de Dados de Pesquisa e Ciência Aberta: Percepção dos Pesquisadores da Universidade Federal do Ceará". Este estudo está sendo desenvolvido por mim, Juliana Soares Lima, mestranda no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará (UFC), sob a orientação da Professora Doutora Maria Giovanna Guedes Farias. O objetivo geral se concentra em investigar a percepção, as práticas e as necessidades informacionais concernentes ao gerenciamento de dados científicos e à ciência aberta, a fim de propor a criação de serviços e de um programa de suporte e apoio ao pesquisador na Universidade Federal do Ceará. Este questionário tem como público-alvo os docentes, discentes e técnico-administrativos da Universidade Federal do Ceará que estão cursando ou já concluíram uma Pós-graduação Stricto Sensu (Mestrado, Doutorado e/ou Pós-Doutorado).

Suas respostas serão utilizadas exclusivamente para fins científicos. Asseguro manter a integridade e o sigilo quanto às informações prestadas, e, seus dados pessoais ou de contato não estão sendo coletados, portanto, não há quaisquer dados que possam identificá-lo(a). Ainda assim, garanto que os dados serão anonimizados e não serão divulgadas informações que estejam relacionadas à sua intimidade. Caso aceite participar, não haverá qualquer prejuízo para você em face das informações fornecidas.

Riscos: O risco relacionado à sua participação na pesquisa refere-se ao sentimento de desconforto que poderá decorrer da manifestação sobre assuntos pessoais. Nesse sentido, caso considere muito pessoal ou confidencial, ressalto que você tem toda liberdade para retirar o seu consentimento e interromper as respostas a qualquer momento, mesmo que tenha concordado anteriormente. Saliento que os dados e resultados desta pesquisa estarão sempre sob sigilo ético.

Benefícios: Sua participação e colaboração são imprescindíveis, além de permitir ampliar a compreensão sobre o tema em estudo. Inclusive, contribuirá para a criação de novos serviços e um programa de apoio e suporte ao pesquisador em práticas da ciência aberta e no gerenciamento de dados científicos na UFC.

Os resultados desta pesquisa serão publicados em minha dissertação de mestrado a ser defendida na UFC, e, após a defesa, esta estará disponível no Repositório Institucional (<http://www.repositorio.ufc.br/>). Outras formas, tais como comunicação a ser apresentada em congresso e publicações científicas (artigos, livros e capítulos) poderão ser produzidas a partir dos resultados deste estudo.

Em caso de dúvidas, coloco-me à disposição para prestar esclarecimentos sobre a pesquisa a qualquer tempo, por e-mail: [juliana.lima@ufc.br](mailto:juliana.lima@ufc.br) ou [julia10br@gmail.com](mailto:julia10br@gmail.com). Endereço: Avenida da Universidade, 2683, Biblioteca de Ciências Humanas, Benfica, Fortaleza, Ceará, CEP: 60.020-970. Esta pesquisa foi aprovada pelo Comitê de Ética da Universidade Federal do Ceará, conforme Parecer N°. 3.512.152/2019.

Desde já agradeço a sua disponibilidade e atenção.

\*Obrigatório

#### 1. Termos e condições de aceitação: \*

Aceite os termos e condições para participar da pesquisa.  
*Marque todas que se aplicam.*

- Declaro que fui devidamente esclarecido(a) sobre a pesquisa e dou o meu consentimento.
- Estou ciente de que os dados coletados nesta pesquisa serão usados exclusivamente para fins científicos. Nenhum de meus dados pessoais estão sendo coletados e não serão divulgados, garantindo a confidencialidade.

#### Perfil dos usuários

As perguntas desta seção permitirão traçar o perfil dos respondentes.

**2. Categoria: \***

É possível marcar mais de uma opção se necessário.  
*Marque todas que se aplicam.*

- Docente
- Discente
- Técnico-Administrativo em Educação (TAE)
- Outro: \_\_\_\_\_

**3. Titulação: \***

Informe apenas a sua maior titulação já obtida.  
*Marcar apenas uma oval.*

- Graduação
- Especialização
- Mestrado
- Doutorado

**4. Área do Conhecimento \***

*Marcar apenas uma oval.*

- Ciências Exatas e da Terra
- Ciências Biológicas
- Engenharias
- Ciências da Saúde
- Ciências Agrárias
- Ciências Sociais Aplicadas
- Ciências Humanas
- Linguística, Letras e Artes

**5. Qual é o seu curso? \***

Exemplo: Administração, Biblioteconomia,  
Ciência da Informação, Educação, Engenharia  
Elétrica, Física, Matemática, Medicina,  
Odontologia, Psicologia etc.

\_\_\_\_\_

**6. Gênero: \***

*Marcar apenas uma oval.*

- Feminino
- Masculino
- Prefiro não informar

**7. Faixa etária: \***

É possível marcar apenas uma opção.

*Marcar apenas uma oval.*

- 18 a 20 anos
- 21 a 30 anos
- 31 a 40 anos
- 41 a 50 anos
- 51 anos ou mais

**Delineamento das Pesquisas e Tipos de dados produzidos**

Informe nas questões a seguir quais são os tipos de dados produzidos e a natureza das pesquisas que você desenvolve.

**8. Geralmente, suas pesquisas possuem abordagem: \***

*Marcar apenas uma oval.*

- Qualitativa
- Quantitativa
- Quantitativa e qualitativa



**9. Quais são os tipos de dados gerados em suas atividades de pesquisas? \***

É possível marcar quantas opções desejar.

Marque todas que se aplicam.

- Documentos (Texto)
- Planilhas (Dados numéricos)
- Slides
- Fotografias / Imagens
- Materiais audiovisuais (áudio e/ou vídeo)
- Animações e/ou Filmes
- Entrevistas (transcrições)
- Questionários
- Respostas de testes
- Testes Psicológicos
- Diário de campo
- Caderno de laboratório
- Livros de códigos (Codebooks)
- Aplicativos e/ou softwares
- Bancos de dados e/ou Bases de dados
- Arquivos de dados (descrevem outros dados ou conjuntos de dados)
- Códigos, algoritmos, scripts
- Conteúdos de uma aplicação específica (registros de entrada, saída, arquivos de log para análise de software, esquemas)
- Artefatos, espécimes, amostras
- Simulações e/ou experimentos
- Modelos
- Metodologias e fluxos de trabalho
- Procedimentos operacionais e protocolos padrão
- Mapas afetivos
- Outro: \_\_\_\_\_

**10. O que você entende por 'Dados de Pesquisa'? \***

Informe aqui qual a sua percepção a respeito.

---

---

---

---

---

**11. Como você costuma organizar, armazenar, manter e preservar os seus dados de pesquisa para uso futuramente? \***

É possível marcar quantas opções desejar.  
Marque todas que se aplicam.

- Computador pessoal
- Computador do trabalho
- Servidores / Datacenters da instituição
- Pendrive, HD externo, Cartão de memória, CD e outras mídias e dispositivos portáteis
- Nuvem (Amazon Cloud, Dropbox, Google Drive etc.)
- Papel (costumo manter cópias impressas dos dados)
- Outro: \_\_\_\_\_

**12. Com que frequência você realiza o backup de seus dados e arquivos? \***

Apenas uma opção pode ser selecionada.  
Marcar apenas uma oval.

- Diária
- Semanal
- Quinzenal
- Mensal
- Não há uma frequência específica
- Não realizo backup

**13. Você já passou por alguma experiência problemática no armazenamento de seus dados? \***

É possível marcar quantas opções desejar.  
Marque todas que se aplicam.

- Dados perdidos ou corrompidos em computador pessoal defeituoso.
- Dados perdidos ou corrompidos em computador do trabalho.
- Servidores / Datacenters da instituição perderam os dados da pesquisa.
- Pendrive, HD externo, CD e outras mídias e dispositivos portáteis defeituosas.
- Dados perdidos por causa de vírus e outras pragas virtuais.
- Perda de acesso aos serviços de hospedagem na nuvem (Amazon Cloud, Dropbox, Google Drive etc.)
- Papel (costumo manter cópias impressas dos dados, porém, os papéis se deterioraram ou foram jogados fora).
- Nunca tive problemas com o armazenamento dos meus dados.
- Outro: \_\_\_\_\_

**14. Você conhece ou já utilizou os serviços do Centro Nacional de Processamento de Alto Desempenho - CENAPAD UFC em sua instituição para o armazenamento de dados de pesquisa em sua instituição? \***

Apenas uma opção pode ser selecionada.  
Marcar apenas uma oval.

- Sim
- Não

**15. Você já elaborou um Plano de Gestão de dados (PGD) para documentar as suas pesquisas? \***

Plano de Gestão de Dados (PGD): Documento formal que descreve como serão tratados os dados de pesquisa, desde a fase inicial até a etapa final do projeto.

*Marque todas que se aplicam.*

- Sim, por solicitação / exigência de publicação em periódico científico.
- Sim, por solicitação / exigência de minha instituição.
- Sim, por solicitação / exigência de minha agência financiadora / fomento.
- Sim, já elaborei por iniciativa própria.
- Não, nunca elaborei.
- Outro: \_\_\_\_\_

**16. Se você já elaborou um Plano de Gestão de dados (PGD), qual ferramenta utilizou?**

Responda esta questão apenas se você já elaborou um PGD. É possível marcar quantas opções desejar.

*Marque todas que se aplicam.*

- DMPonline
- DMPtool
- Pagoda - PGDonline
- Outro: \_\_\_\_\_

### **Práticas e princípios da Ciência Aberta**

Informe nas questões a seguir qual a sua familiaridade com os itens abaixo.

**17. Marque o quanto você conhece e está familiarizado com as seguintes práticas e processos: \***

É exigida uma resposta por linha.

Marcar apenas uma oval por linha.

	Conheço bastante.	Conheço pouco.	Desconheço.
Compartilhamento de Dados de Pesquisa Abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gestão de Dados de Pesquisa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dados Governamentais Abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reprodutibilidade da Pesquisa e Análise de dados Aberta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Códigos, Softwares de Pesquisa Abertos e Open Source.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hardware abertos (Impressoras 3D e outros equipamentos em Makerspaces / Hackerspaces).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acesso Aberto às Publicações Científicas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avaliação por pares aberta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Métodos, protocolos e Instrumentos científicos abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cadernos de Laboratório Abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recursos Educacionais Abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Colaboração Aberta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advocacy e engajamento público em prol da Ciência.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ciência Cidadã.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Promoção de políticas governamentais e mandatos institucionais favoráveis à Ciência Aberta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inovação aberta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Artes e processos de criação artística abertos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**18. O que você entende por 'Ciência Aberta'? \***

Sinta-se livre para descrever com suas palavras ou definir com apenas uma palavra.

---



---



---



---



---

19. Na sua opinião, o quanto você considera importante que os pesquisadores adotem os princípios da ciência aberta? \*

Marcar apenas uma oval por linha.

	Pouco importante	Relativamente importante	Importante	Muito importante	Não sei responder / Não tenho certeza
Diversidade: inclusão de grupos e subgrupos pouco representados na ciência (raça, gênero, cultura etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Novas e inovadoras possibilidades econômicas: crowdfunding, economia criativa, novos fundadores etc.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eficiência: compartilhamento de dados, procedimentos para otimizar a ciência	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Equidade: acesso para todos aos resultados científicos, métodos, softwares, independentemente da capacidade econômica e afiliação institucional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ética: A Ciência Aberta está alinhada com os princípios de integridade da pesquisa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Justiça: A ciência é um bem público, financiada pela sociedade, portanto, seus resultados devem estar disponíveis para todos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Impacto: para superar métricas tradicionais de impacto científico, maior audiência e engajamento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rigor: acesso aberto, dados abertos, para tornar mais fácil a replicabilidade e reprodutibilidade da pesquisa, facilitando os processos de revisão	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**20. Você costuma acessar outros repositórios ou sites para obter dados de pesquisa de outros pesquisadores / grupos de pesquisa? \***

Apenas uma opção pode ser marcada.

*Marcar apenas uma oval.*

Sim

Não

**21. Você compartilha seus dados de pesquisa? \***

Observação: Dados de pesquisa aqui entendidos como os dados brutos ou dados que sustentam os resultados apresentados em um artigo, dissertação, tese ou outras publicações. Por exemplo: planilha em Excel com dados coletados em uma pesquisa/levantamento, modelo de questionário, modelo de formulário, corpus textual, cálculos referenciais, poligonais, cadernos de laboratório, anotações ou notas de campo, scripts, códigos de programação etc. É possível marcar quantas opções desejar.

*Marque todas que se aplicam.*

Sim, compartilho livremente e em acesso aberto os resultados de minha pesquisa por iniciativa própria em repositório institucional, repositório de dados e outras plataformas

Sim, compartilho apenas com os colegas e/ou alunos mais próximos e/ou que fazem parte do mesmo grupo de pesquisa

Sim, já compartilhei a pedido da revista em que publiquei, pois, a disponibilização dos dados é uma das condições exigidas para a publicação do manuscrito, ou porque esse serviço foi oferecido pelo publicador em plataforma própria

Sim, por solicitação ou exigência de meu financiador / agência de fomento

Sim, por solicitação ou exigência de minha instituição

Não, nunca compartilhei meus dados

Outro: \_\_\_\_\_

**22. Se você costuma compartilhar seus dados, informe a seguir de que forma:**

Responda apenas se você compartilha seus dados. Se não compartilha, passe para a próxima questão.

*Marque todas que se aplicam.*

Repositório institucional

Repositórios temáticos / específicos da área (ArrayExpress, caNanoLab, HEPdata, ITIS, KIMoSys, NCBI Taxonomy, NDEx, NoMaD, Omics, PANGAEA, PsychData, SEAONE etc. )

Repositórios de dados (Exemplo: Dataverse, Dryad, Figshare, Mendeley Data, Open Science Framework, Zenodo)

Repositório da revista / editor (Exemplo: IEEE Dataport e Mendeley Data)

Portal Brasileiro de Dados Abertos

Outro: \_\_\_\_\_

**23. Se você não compartilha, informe o(s) motivo(s):**

Responda apenas se você NÃO compartilha seus dados. Se compartilha, passe para a próxima questão.

*Marque todas que se aplicam.*

- Desconhecimento, não sabia que posso fazer isso e/ou não sei como
- Não compartilho porque não tenho certeza sobre quais os tipos de licenças e direitos para o uso e indicações de restrição posso aplicar para que outras pessoas se utilizem dos meus dados
- Não compartilho porque não sei como os meus dados serão utilizados por outras pessoas
- Não compartilho porque tenho receio de que a minha pesquisa perca o ineditismo, seja plagiada/copiada e utilizada indevidamente
- São dados confidenciais, não posso divulgar
- Não tenho permissão dos participantes da pesquisa para divulgar os dados
- Os dados são objeto de patente, não posso compartilhar
- Os dados não estão devidamente documentados
- Os dados estão em formato proprietário ou que dificulta o compartilhamento e uso
- Falta de padronização dos dados
- Falta de financiamento
- Falta de infraestrutura
- Não é uma exigência em minha instituição
- Não é uma exigência de meu financiador / agência de fomento
- Não compartilho porque preciso publicar os resultados da pesquisa antes
- Não tenho tempo para fazer isso
- Não tenho interesse em disponibilizar meus dados de pesquisa
- Outro: \_\_\_\_\_

## 24. Sobre o compartilhamento: \*

Informe se concorda ou discorda das assertivas a seguir.

Marcar apenas uma oval por linha.

	Concordo	Discordo	Não sei responder
Os dados de pesquisa devem ser compartilhados livremente e disponibilizados em repositórios de dados e plataformas especializadas para todos que busquem e queiram usá-los.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Os dados devem ser compartilhados com os meus colegas e comunidade acadêmica.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Compartilharia os meus dados apenas se isso fosse obrigatório	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Na minha área / disciplina o compartilhamento de dados é uma prática comum e tradicional de cooperação entre os pares.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A disponibilização dos dados é importante para a verificação e reprodutibilidade da pesquisa científica.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estou disposto(a) a me dedicar em realizar os processos de organização e disponibilização dos meus dados de pesquisa para que sejam reutilizados por outros.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se eu compartilhar, meus dados podem ser mal utilizados ou mal interpretados por outros.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ao compartilhar meus dados minhas ideias serão copiadas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se compartilhar meus dados posso perder vantagem em processos de inovação e patentes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posso perder a oportunidade de publicações futuras sobre o assunto de minha pesquisa se eu compartilhar os dados.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Compartilhar dados científicos contribui para a divulgação, alcance e impacto de minha pesquisa, tornando-me cada vez mais reconhecido em minha área e pelos meus pares.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Compartilhar dados científicos podem me render citações pelos meus conjuntos de dados.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Considero apropriado que outros pesquisadores possam reutilizar dados de outros, criando o seu próprio conjunto de dados a partir dos dados que foram compartilhados.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Considero disponibilizar meus dados de pesquisa desde que estejam sob uma licença apropriada e no estabelecimento de termos de acesso e uso.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



	Concordo	Discordo	Não sei responder
O compartilhamento de dados deve ser uma exigência institucional e governamental, especialmente em pesquisas financiadas com o dinheiro público	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. **Você utiliza algum identificador digital para autores? Informe abaixo: \***

É possível marcar mais de uma opção  
*Marque todas que se aplicam.*

- ID Lattes
- ORCID (Open Researcher and Contributor Identifier)
- ResearcherID
- Scopus Author ID
- Authenticus
- Não utilizo / Não sei responder
- Outro: \_\_\_\_\_

26. **Marque quais são as mídias sociais, redes sociais acadêmicas, blogs ou outros recursos que você utiliza para a divulgação de seus estudos e para ampliar o impacto de suas pesquisas: \***

Informe se você possui perfil em redes sociais acadêmicas tais como Academia.edu, ResearchGate, Mendeley ou outras; perfil no Google Acadêmico; site pessoal, blog etc. É possível marcar quantas opções desejar.  
*Marque todas que se aplicam.*

- Academia.edu
- ResearchGate
- Mendeley
- Google Acadêmico (Perfil)
- Facebook
- Instagram
- LinkedIn
- Twitter
- Publons
- Impact Story
- Blog
- Site pessoal ou profissional
- Newsletter
- Nenhuma
- Outro: \_\_\_\_\_

**27. Você costuma publicar em revistas de acesso aberto ou em revistas com restrição de acesso? \***

Apenas uma opção pode ser selecionada.  
*Marcar apenas uma oval.*

- Sim, só publico em revistas de acesso livre e irrestrito.
- Não, publico em revistas de acesso restrito.
- Parcialmente, já publiquei /publico em revistas de acesso livre e em periódicos de acesso restrito.
- Não sei responder

**28. Você costuma acessar e ler preprints (pré-publicação não avaliada ou aguardando avaliação por pares de um artigo) de outros autores na internet ou outros repositórios? \***

Apenas uma opção pode ser marcada. Caso não saiba diferenciar uma versão preprint de um artigo publicado, então, marque 'Não sei'.  
*Marcar apenas uma oval.*

- Sim
- Não
- Não sei

**29. Sobre o preprint: \***

Marque a seguir apenas as assertivas com as quais você concorda.  
*Marque todas que se aplicam.*

- Leio as versões preprint dos artigos porque é a forma mais rápida de acompanhar as novidades em minha área.
- Leio as versões preprint dos artigos porque além de ser uma forma de manter-me atualizado(a) também posso acompanhar os comentários dos colegas e posso contribuir com sugestões para os autores do manuscrito.
- Leio as versões preprint dos artigos, porém, nunca publiquei e/ou disponibilizei um até o momento.
- Publico as versões preprint dos meus artigos porque acredito ser a melhor forma de obter um feedback, sugestões e críticas dos pares a fim de melhorar meu manuscrito antes de enviá-lo para a avaliação de uma revista.
- Não costumo publicar versões preprint de meus artigos.
- Não confio em versões que não foram revisadas por pares.

**30. Você considera que a sua instituição deveria oferecer serviços de apoio ao pesquisador sobre coleta, organização, arquivamento, preservação e compartilhamento de dados? \***

Apenas uma alternativa pode ser selecionada.  
*Marcar apenas uma oval.*

- Sim
- Não

31. Quem você considera estar qualificado ou julga ser importante no apoio ao pesquisador e deve oferecer serviços de coleta, organização, arquivamento, preservação e compartilhamento de dados? \*

É possível marcar quantas opções desejar.

Marque todas que se aplicam.

- Bibliotecas
- Agências de fomento
- Entidades governamentais
- Considero importante criar um setor especializado com uma equipe multidisciplinar
- Não considero que seja necessário apoio
- Outro: \_\_\_\_\_

32. Informe o quanto você considera importante a assistência e apoio ao pesquisador por parte do bibliotecário sobre orientações em pesquisa, ciência aberta e serviços de coleta, organização, arquivamento, preservação e compartilhamento de dados: \*

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco importante	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito importante

33. Comentários adicionais.

Utilize esse espaço para registrar qualquer comentário geral, experiência ou relato que possa complementar as respostas dadas nesse questionário.

---

---

---

---

---

## APÊNDICE B - ROTEIRO DE ENTREVISTA



### UNIVERSIDADE FEDERAL DO CEARÁ PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO MESTRADO EM CIÊNCIA DA INFORMAÇÃO

#### ROTEIRO DE ENTREVISTA

NOME:

---

CATEGORIA: ( ) Docente ( ) Discente ( ) Técnico-Administrativo em Educação (TAE)

FAIXA ETÁRIA:

( ) 18 a 20 anos ( ) 21 a 30 anos ( ) 31 a 40 anos

( ) 41 a 50 anos ( ) 51 anos ou mais

GÊNERO: ( ) Masculino ( ) Feminino ( ) Prefiro não informar

1) Qual o tipo de abordagem você utiliza em suas pesquisas (quantitativa, qualitativa, quanti-qualitativa) e quais os tipos de dados são gerados (slides, textos, questionário, entrevista etc.)?

2) Como você organiza, armazena, mantém e preserva os seus dados para uso futuramente? (Faz backup? Tem uma frequência de backup? Já teve problemas com o armazenamento e/ou perda de dados e arquivos?)

3) O que você entende por Dados de Pesquisa? (Você sabe o que é?)

4) Você conhece / está ciente de que existem repositórios de dados de pesquisa generalistas e/ou em sua área? (Você envia ou já enviou dados para estes repositórios? Se sim, cite pelo menos um. Ou apenas conhece o Repositório Institucional da UFC?)

5) Você costuma documentar suas pesquisas elaborando um Plano de Gestão de Dados (PGD) ou utiliza algum outro instrumento para este fim? (Sim ou não? Em caso de resposta positiva para a pergunta acima, qual ferramenta você utiliza para a elaboração do PGD? a) DMPtool; b) DMPonline; c) Modelo sugerido pela minha instituição; d) Outros. Informe qual: Por quê?).

6) Quais estratégias e/ou meios você utiliza para divulgar as suas pesquisas? (Mídias sociais, Redes Sociais Acadêmicas, Blog ou site pessoal/profissional etc.)

7) A publicação de artigos e a disponibilização dos dados de pesquisa associados ao manuscrito é uma das tendências da Ciência Aberta. Nesse caso, pergunto: você já publicou em revistas que exigem o envio e compartilhamento dos dados referentes ao manuscrito submetido? (Se sim, conte a experiência. Se não, pergunto: você se sente preparado para publicar um artigo e disponibilizar os dados se assim a revista exigir?).

8) Ainda em relação à pergunta anterior: no papel de avaliador, você se considera preparado para avaliar tanto os artigos como os conjuntos de dados associados ao manuscrito?

9) Outra tendência da Ciência Aberta que tem se consolidado é a revisão por pares aberta em periódicos científicos. Qual a sua opinião sobre isso?

10) Há vários estudos relacionados à Ciência Aberta que apontam existir uma crise de reprodutibilidade e integridade da pesquisa em todo o mundo. Inclusive, estes são alguns dos motivos pelos quais têm sido exigidas cada vez mais transparência e divulgação dos resultados de pesquisas, especialmente aquelas financiadas com o dinheiro público. Qual a sua opinião sobre isso? (Você é a favor ou contra o compartilhamento de dados de pesquisa?)

11) A sua área tem como tradição o compartilhamento de dados de pesquisa com outros pares/colegas, alunos ou até mesmo trabalham em cooperação com outras áreas nesse sentido?

12) Você compartilha seus dados de pesquisa? (Compartilha dados com outros colegas ou alunos? Se sim, onde e como, em que condições? Não compartilha? Por quê? Compartilharia apenas se fosse uma exigência?).

13) Para você, qual seria o maior desafio no compartilhamento de dados de pesquisa?

14) O que você entende por Ciência Aberta? (Você sabe o que é?).

15) Você considera que, futuramente, a biblioteca / os bibliotecários devem oferecer serviços de suporte e apoio em Ciência Aberta e em gerenciamento de dados de pesquisa? Explique.

16) Você acredita que a capacitação e a sensibilização poderiam ajudar na adoção dessas práticas?

17) Você gostaria de ser avisado sobre a oferta de cursos presenciais ou online sobre estes temas que venham a ser ofertados futuramente? (Aceitaria e teria disponibilidade para participar de um minicurso online sobre o tema?)

( ) Sim      ( ) Não

18) Há algo mais que gostaria de acrescentar sobre essa entrevista?

## APÊNDICE C – MODELO DO MINICURSO ONLINE



UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE HUMANIDADES  
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO  
MESTRADO EM CIÊNCIA DA INFORMAÇÃO

### PLANO DE CURSO

#### 1 DADOS DE IDENTIFICAÇÃO

##### CURSO

Dados de Pesquisa: dicas para gerenciar, organizar, compartilhar e preservar dados.

##### DISCIPLINA

Não se aplica.

##### MODALIDADE

Minicurso online, Educação a Distância (EAD).

##### INÍCIO E TÉRMINO DO CURSO

Setembro de 2019 a Novembro de 2019.

##### CARGA HORÁRIA

15 horas / aula.

##### PÚBLICO-ALVO

Docentes e Discentes dos cursos de Pós-Graduação *Stricto Sensu* da Universidade Federal do Ceará.

##### MÉDIA DE APROVAÇÃO

Sem nota. As atividades de fixação do conteúdo serão opcionais.

---

**PLATAFORMA**

Google Classroom.

**SETOR DE ESTUDO**

Universidade Federal do Ceará.

**PROFESSORA / TUTORA / PESQUISADORA RESPONSÁVEL**

Juliana Soares Lima (Mestranda).

**ORIENTADORA / PESQUISADORA RESPONSÁVEL**

Maria Giovanna Guedes Farias (Orientadora).

**2 JUSTIFICATIVA**

Atualmente, a gestão de dados de pesquisa tem sido incorporada ao cotidiano dos pesquisadores pelo fato de representar responsabilidade na condução de projetos de pesquisa e por atender às boas práticas na ciência. Assim, a necessidade de se essenciais para que possam conhecer as principais práticas no gerenciamento de dados de pesquisa, tais como a elaboração de um Plano de Gestão de Dados (PGD), bem como a produção da documentação que acompanha os dados a fim de inseri-los em um contexto para que possam ser compreendidos, contextualizados, arquivados, preservados e acessados facilmente no presente e futuramente, tanto pelo pesquisador quanto por outras pessoas, viabilizando também o compartilhamento desses dados oriundos de pesquisa científica. Nesse sentido, o(a) bibliotecário pode auxiliar os pesquisadores, mediando ações de informação / intervenção que promovam a competência em informação e a competência em dados, uma vez que a partir da conscientização e conhecimento dessas práticas os pesquisadores de diversas áreas se sentirão mais capazes e confiantes para cumprir as etapas de coleta, armazenamento, gerenciamento e compartilhamento dos dados provenientes da pesquisa científica.

---

### 3 EMENTA

Compreensão dos conceitos e tipologia dos dados de pesquisa, assim como conhecer as principais práticas no gerenciamento de dados de pesquisa e dicas na elaboração de um Plano de Gestão de Dados (PGD), documento formal que descreve que tipos de dados foram coletados, como serão armazenados, processados, preservados, arquivados e compartilhados. Reflexão sobre o papel do pesquisador na gestão de seus dados de pesquisa como um agente capaz de contribuir para as boas práticas na ciência, assim como repensar as suas práticas, comportamento e atitudes que visam contribuir com o movimento de acesso livre ao conhecimento por meio da disponibilização dos dados. Adicionalmente, serão apresentados alguns conceitos introdutórios e básicos sobre a Ciência Aberta e sua relação com os dados de pesquisa no campo científico.

### 4 OBJETIVOS DO CURSO

#### 4.1 Objetivo Geral

Criar um ambiente de estudo em que o pesquisador possa conhecer as práticas envolvidas na gestão de dados de pesquisa, refletir sobre o seu papel e a importância da adoção do gerenciamento de dados científicos, assim como pode contribuir para o avanço da ciência por meio da organização, disponibilização e compartilhamento dos dados.

#### 4.1 Objetivos Específicos

- Apresentar os principais conceitos e tipologia dos dados de pesquisa, compartilhamento, gestão de dados de pesquisa;
- Indicar fontes de pesquisa e repositórios de dados para os pesquisadores;
- Mostrar aos pesquisadores como gerenciar e documentar seus dados de pesquisa para preservação e uso futuramente;
- Capacitar os pesquisadores no preenchimento e elaboração de um Plano de Gestão de Dados (PGD);



- Introduzir o conceito de Ciência Aberta e destacar o seu impacto no *modus operandi* do desenvolvimento das pesquisas na atualidade.

## 5 METODOLOGIA DE ENSINO

Método expositivo com o uso de slides, videoaulas e indicação de outros vídeos, inclusive de fontes externas, entre outros recursos audiovisuais. Estratégias de interação entre a instrutora do minicurso e os cursistas no Ambiente Virtual de Aprendizado (AVA). Estrutura flexível, com a sugestão de atividades opcionais, indicação de referências para leitura e estímulo ao debate e troca de ideias sobre os temas abordados, tendo como foco a participação e interação dos pesquisadores.

## 6 RECURSOS DIDÁTICOS

- Apresentação de slides;
- Recursos integrados à Plataforma online Google Classroom;
- Videoaulas, vídeos e podcasts (fonte externa e interna) e recursos audiovisuais diversos;
- Indicações de ferramentas, websites e outros recursos úteis.

## 7 PROPOSTAS DE AVALIAÇÃO

- Participação e interação dos cursistas;
- Apresentação dos participantes;
- Busca por repositórios de dados no re3data;
- Busca no Google Dataset Search;
- Dicas para a elaboração de um Plano de Gestão de Dados (PGD);
- Sugestão de cadastro em repositórios de dados;
- Sugestão de cadastro em redes sociais acadêmicas e afins para contribuir com a divulgação e impacto da pesquisa: Academia.edu, ResearchGate, Mendeley; Criar perfil no Google Acadêmico; Impact Story; Open Science Framework; ORCID; Publons; Twitter; Facebook; LinkedIn, entre outras;
- Sugestão de criação de conta em ferramentas de workflow e para trabalho em equipe nas pesquisas: Artifact, GitHub, Open Science Framework, Protocols e outros.

## 8 PLANO DE AULA - META DE APRENDIZAGEM

<b>APRESENTAÇÃO</b>	Apresentação e Ambientação na Plataforma Google Classroom.
<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentação dos participantes e Ambientação (Plataforma Google Classroom; regras de boa convivência / netiqueta em ambientes virtuais e disposições gerais sobre o curso).
<b>TEMÁTICA</b>	Apresentação dos participantes e Ambientação
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Apresentação de slides sobre os recursos oferecidos no Google Classroom e dicas de utilização.
<b>PRÁTICAS</b>	Apresentação dos participantes do curso.
<b>AValiação</b>	Como estratégia avaliativa será ponderada a apresentação do cursista, a interação com os colegas de turma e a tutora em troca de mensagens no mural do curso.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	1h/aula.
<b>AULA 1</b>	
<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentar as práticas envolvidas na gestão de dados de pesquisa, refletir sobre o seu papel e a importância da adoção do gerenciamento de dados científicos, assim como pode contribuir para o avanço da ciência por meio da organização, disponibilização e compartilhamento dos dados.
<b>TEMÁTICA</b>	Dados de Pesquisa: Introdução
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Videoaula, vídeo de 1 minuto sobre a importância do gerenciamento dos dados e slides.
<b>PRÁTICAS</b>	Nenhuma atividade opcional. Enquete aberta para a votação de melhor dia e horário para Hangouts com a turma.
<b>AValiação</b>	Todas as atividades do curso são opcionais, mas, será ponderada a interação com os colegas de turma e a tutora em troca de mensagens no mural do curso.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	2h/aula.
<b>AULA 2</b>	
<b>OBJETIVOS DE APRENDIZAGEM</b>	Dados de Pesquisa: onde buscar e compartilhar?

<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentar o RE3DATA e vários repositórios para a busca e compartilhamento de dados.
<b>TEMÁTICA</b>	Repositórios de dados, busca de datasets e compartilhamento de dados
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Vídeoaula; Indicação de sites e repositórios de dados; Wiki de repositórios disciplinares; Lista de Data Journals; Repositórios de Pré-prints; Google Dataset Search; Slides.
<b>PRÁTICAS</b>	Atividades opcionais: Pesquisa no Google Dataset Search e Pergunta sobre a opinião/percepção a respeito das versões pré-print de artigos.
<b>AValiação</b>	Interação entre os participantes e a tutora; comentários sobre as atividades e feedback sobre a experiência de busca no Google Dataset Search e em outros repositórios.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	3h/aula.
<b>AULA 3</b>	
<b>AULA 3</b>	Plano de Gestão de Dados (PGD)
<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentar o Plano de Gestão de Dados (PGD) como uma das estratégias de documentação da pesquisa e como um documento formal que descreve de que maneira os dados serão coletados, tratados, preservados e disponibilizados.
<b>TEMÁTICA</b>	Elaboração de um PGD.
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Vídeoaula; Exemplos de PGD; Slides; Slides sobre cadastro no DMPonline; Notas de aula (Princípios FAIR); Indicação de um periódico que publica PGD's (RIO Journal).
<b>PRÁTICAS</b>	Atividades opcionais sugeridas: Consultar exemplos de PGD; Cadastrar-se no DMPonline.
<b>AValiação</b>	Interação entre os participantes e a tutora; comentários sobre as atividades e feedback sobre a execução das atividades recomendadas.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	3h/aula.
<b>AULA 4</b>	
<b>AULA 4</b>	Gestão de Dados: dicas de organização dos arquivos, medidas de segurança dos dados e licenças.
<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentar estratégias de organização e nomeação de arquivos de dados científicos; Conhecer algumas medidas de segurança para os dados e tipos de licenças que podem ser utilizadas..
<b>TEMÁTICA</b>	Estratégias de organização de arquivos, medidas de segurança dos dados e licenças

<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Vídeoaula; Slides; Notas de aula (Formatos de arquivos recomendados e tipos de licenças); Material disponível na internet sobre dados sensíveis; Slides sobre ORCID; Slides sobre as licenças Creative Commons.
<b>PRÁTICAS</b>	Atividades sugeridas opcionais: Cadastro no ORCID.
<b>AValiação</b>	Interação entre os participantes e a tutora; comentários sobre as atividades e feedback sobre a execução das atividades recomendadas.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima
<b>CARGA HORÁRIA</b>	2h/aula
<b>AULA 5</b>	Ciência Aberta
<b>OBJETIVOS DE APRENDIZAGEM</b>	Apresentar os conceitos básicos e introdutórios sobre a Ciência Aberta e suas relações com os dados de pesquisa.
<b>TEMÁTICA</b>	Ciência Aberta
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Podcast; Slides; Notas de aula; Tira-dúvidas online (Hangouts); Indicação de redes sociais acadêmicas e demais Plataformas relacionadas à Ciência Aberta; Indicação de leitura; Lista de cursos sobre Dados de Pesquisa e Ciência Aberta.
<b>PRÁTICAS</b>	Atividades opcionais sugeridas: Participação no Hangouts; Cadastro em redes sociais acadêmicas e nas plataformas apresentadas.
<b>AValiação</b>	Interação entre os participantes e a tutora; comentários sobre as atividades e feedback sobre a execução das atividades recomendadas.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	4h/aula.
<b>AValiação DO CURSO</b>	Avaliação sobre o curso.
<b>OBJETIVOS DE APRENDIZAGEM</b>	Não se aplica.
<b>TEMÁTICA</b>	Não se aplica
<b>ESTRATÉGIAS FACILITADORAS DA APRENDIZAGEM</b>	Não se aplica.
<b>PRÁTICAS</b>	Questionário de Avaliação do minicurso online e Lousa virtual (Jamboard)

<b>AVALIAÇÃO</b>	Sem nota. Verificação das respostas do questionário de avaliação enviadas pelos cursistas e as mensagens escritas na Jamboard.
<b>PROFESSOR / TUTOR</b>	Juliana Soares Lima.
<b>CARGA HORÁRIA</b>	Não se aplica.

## 9 DESCRIÇÃO DO CONTEÚDO PROGRAMÁTICO / UNIDADES / TÓPICOS

ORGANIZAÇÃO CURRICULAR DAS AULAS VIRTUAIS				
AULAS	PERÍODO	TIPO	DESCRIÇÃO	CARGA HORÁRIA
Apresentação	11/09/2019 a 01/11/2019	Acompanhamento a distância	1 Apresentação do curso 1.1 Ambientação: Conhecendo o Google Classroom 1.2 Apresentação dos participantes	1h
Referências	LIMA, Juliana Soares; SANTOS, Izabel Lima dos; SANTOS, Francisco Edvander Pires. Google Classroom como ferramenta para treinamentos a distância: um relato de experiência em bibliotecas universitárias. <b>Revista Brasileira de Biblioteconomia e Documentação</b> , São Paulo, v. 13, n. esp., p. 1511-1535, dez. 2017. Disponível em: <a href="https://rbbd.febab.org.br/rbbd/article/view/1044">https://rbbd.febab.org.br/rbbd/article/view/1044</a> . Acesso em: 30 abr. 2018.			
Aula 1: Dados de Pesquisa	11/09/2019 a 01/11/2019	Acompanhamento a distância	2 Dados de Pesquisa 2.1 Dados de Pesquisa: Introdução 2.1.1 Videoaula introdutória 2.1.2 Gestão de Dados de Pesquisa 2.1.3 Centro Nacional de Processamento de Alto Desempenho (Cenapad)	2h
Referências	<p>ARAÚJO, Ana Carolina. <b>Universidade públicas realizam mais de 95% da Ciência no Brasil</b>. Alfenas: Universidade Federal de Alfenas, 2019. Disponível em: <a href="https://bit.ly/2wspyMK">https://bit.ly/2wspyMK</a>. Acesso em: 18 abr. 2019.</p> <p>AVENTURIER, Pascal; ALENCAR, Maria de Cléofas Faggion. Os desafios dos dados de pesquisas abertos. <b>RECIIS</b>, Rio de Janeiro, v. 10, n. 3, jul./set. 2016. Disponível em: <a href="https://bit.ly/2WaOAIv">https://bit.ly/2WaOAIv</a>. Acesso em: 10 abr. 2019.</p> <p>CORTI, Louise <i>et al.</i> <b>Managing and sharing data: best practice for researchers</b>. 3. ed. Colchester: UK Data Archive, 2011. Disponível em: <a href="https://data-archive.ac.uk/media/2894/managingsharing.pdf">https://data-archive.ac.uk/media/2894/managingsharing.pdf</a>. Acesso em: 10 mar. 2019.</p> <p>CRUZ, Carlos Henrique de Brito. <b>100 universidades e institutos no Brasil com mais artigos científicos</b>. 2019. Extraído da Base de dados InCites da Clarivates. Disponível em: <a href="https://bit.ly/2VZ4IPu">https://bit.ly/2VZ4IPu</a>. Acesso em: 18 abr. 2019.</p>			

	<p>DIGITAL SCIENCE. Figshare. <i>The State of Open Data 2017: a selection of analyses and articles about open data</i>, curated by Figshare. London: Holtzbrinck Publishing Group: Springer Nature: Wiley, 2017. DOI: 10.6084/m9.figshare.5481187. Disponível em: <a href="https://is.gd/YicWRa">https://is.gd/YicWRa</a>. Acesso em: 10 mar. 2010.</p> <p>GRAHAM, Dan. Academic Publishing Survey of funders supports the benign Open Access outcome priced into shares. <i>In: HSBC BANK. HSBC Global Research</i>. Europe: HSBC Bank, 2013. Disponível em: <a href="https://bit.ly/2VYZxyX">https://bit.ly/2VYZxyX</a>. Acesso em: 18 abr. 2019.</p> <p>GREEN, Ann; MACDONALD, Stuart; RICE, Robin. <b>Policy-making for Research Data in Repositories: A Guide</b>. Edinburgh: EDINA and University Data Library, 2009. Disponível em: <a href="https://goo.gl/sdyd2p">https://goo.gl/sdyd2p</a>. Acesso em: 20 jul. 2017.</p> <p>HERREMA, Auke. [Data sharing]. [S.l. : s.n.], 2019. Disponível em: <a href="http://wpmu.mah.se/nmict181group4/2018/02/15/standing-shoulder-giants-returning-favor/">http://wpmu.mah.se/nmict181group4/2018/02/15/standing-shoulder-giants-returning-favor/</a>. Acesso em: 10 ago. 2019.</p> <p>HERREMA, Auke. [Data tree: FREE picking. The future?]. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.fosteropenscience.eu/content/cartoonopen-data">https://www.fosteropenscience.eu/content/cartoonopen-data</a>. Acesso em: 10 ago. 2019.</p> <p>HERREMA, Auke. [No use, can not drink it. Data ocean]. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.fosteropenscience.eu/content/cartoon-data">https://www.fosteropenscience.eu/content/cartoon-data</a>. Acesso em: 10 ago. 2019.</p> <p>HERREMA, Auke. [Publications and data]. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.fosteropenscience.eu/content/cartoonpublication-and-data">https://www.fosteropenscience.eu/content/cartoonpublication-and-data</a>. Acesso em: 10 ago. 2019.</p> <p>HERREMA, Auke. [The data are mine!]. [S.l. : s.n.], 2019. Disponível em: <a href="http://depts.washington.edu/relpov/wordpress/wp-content/uploads/2015/03/data-are-mine.jpg">http://depts.washington.edu/relpov/wordpress/wp-content/uploads/2015/03/data-are-mine.jpg</a>. Acesso em: 10 ago. 2019.</p> <p>OLIVEIRA, Djalma de Pinho Rebouças. <b>Sistemas, organizações e métodos: uma abordagem gerencial</b>. 13. ed. São Paulo, 2002.</p> <p>POPPER, Karl. <b>Conhecimento objetivo: uma abordagem evolucionária</b>. Belo Horizonte: Itatiaia, 1999.</p> <p>PIZZA DE DADOS: Episódio 023: Bioinformática. Entrevistadores: Leticia Portella, Jéssica Temporal e Gustavo Coelho. Entrevistada: Thais Sarraf Sabedot. [S.l. : s.n.], 2019. Podcast. Disponível em: <a href="https://youtu.be/TmQgn-KYjhE?t=723">https://youtu.be/TmQgn-KYjhE?t=723</a>. Acesso em: 10 ago. 2019.</p> <p>ROYAL SOCIETY OF LONDON. <b>Royal Society Science Policy Centre. Science as an open enterprise: open data for open science</b>. London: Royal Society of London, 2012. Disponível em: <a href="https://bit.ly/2l7vgcB">https://bit.ly/2l7vgcB</a>. Acesso em: 20 jul. 2017.</p> <p>SALES, Luana Farias; SAYÃO, Luís Fernando. <b>Dados de pesquisa: quem ama cuida</b>. Ilustração de: Joana Sales Marques. Brasília, DF : Comissão Nacional de Energia Nuclear (CNEN) : IBICT, 2019. Cartilha de Gestão da Dados de Pesquisa para Pesquisadores. (Série: Divulgação em</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>Ciência Aberta). Disponível em: <a href="https://is.gd/UX9kwL">https://is.gd/UX9kwL</a>. Acesso em: 10 ago. 2019.</p> <p>SANTOS, Solange Maria. Rankings Internacionais de Universidades: Comparação e Desempenho por Áreas. In: MARCOVITCH, Jacques (org.). <b>Repensar a universidade</b>: desempenho acadêmico e comparações internacionais. São Paulo: Com-Arte; Fapesp, 2018. Parte 1, cap. 4, p. 63-92. Disponível em: <a href="https://bit.ly/2VWKx4R">https://bit.ly/2VWKx4R</a>. Acesso em: 10 abr. 2019.</p> <p>SAYÃO, Luís Fernando; SALES, Luana Farias. <b>A ciência invisível</b>: revelando os dados da cauda longa da pesquisa. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - Enancib, 19., 2018. <b>Anais [...]</b>. Londrina: Enancib, 2019. Não paginado. Disponível em: <a href="https://bit.ly/2WafXcZ">https://bit.ly/2WafXcZ</a>. Acesso em: 10 maio 2019.</p> <p>SAYÃO, Luís Fernando; SALES, Luana Farias. <b>Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores</b>. Rio de Janeiro: CNEN/EN, 2015. Disponível em: <a href="https://is.gd/pOSKQO">https://is.gd/pOSKQO</a>. Acesso em: 20 jul. 2017.</p> <p>SILVA, Fabiano Couto Corrêa da. <b>Gestão de dados científicos</b>. Rio de Janeiro: Interciência, 2019.</p> <p>SUBER, Peter. 2007. Creating an intellectual commons through open access. In: HESS, Charlotte; OSTROM, Elinor. <b>Understanding knowledge as a commons</b>: from theory to practice. Cambridge: MIT Press, 2007. p. 171-208. Disponível em: <a href="http://www.wtf.tw/ref/hess_ostrom_2007.pdf">http://www.wtf.tw/ref/hess_ostrom_2007.pdf</a>. Acesso em: 10 ago. 2019.</p>			
Aula 2: Dados de Pesquisa: onde buscar e compartilhar?	11/09/2019 a 01/11/2019	Acompanhamento a distância	<p>2.2.1 Dados de Pesquisa: onde buscar e compartilhar?</p> <p>2.2.2 Dicas de sites para pesquisar repositórios de dados</p> <p>2.2.3 Wiki de repositórios disciplinares</p> <p>2.2.4 Figshare</p> <p>2.2.5 Lista de Data Journals</p> <p>2.2.6 Repositórios de Preprints</p> <p>2.2.7 Editoras e periódicos científicos: uso dos preprints e avaliação aberta</p> <p>2.2.8 Google Dataset Search</p>	3h
Referências	<p>AVENTURIER, Pascal; ALENCAR, Maria Cleofas de. Os desafios de dados de pesquisa abertos. <b>Revista Eletrônica de Comunicação, Informação e Inovação em Saúde</b>, Rio de Janeiro, v. 10, n. 3, set. 2016. ISSN 1981-6278. Disponível em: <a href="https://www.reciis.icict.fiocruz.br/index.php/receis/article/view/1069">https://www.reciis.icict.fiocruz.br/index.php/receis/article/view/1069</a>. Acesso em: 10 ago. 2019.</p> <p>GOFAIR. <b>FAIR Principles</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.go-fair.org/fair-principles/">https://www.go-fair.org/fair-principles/</a>. Acesso em: 10 ago. 2019.</p>			
Aula 3: Plano de Gestão de Dados (PGD)	11/09/2019 a 01/11/2019	Acompanhamento a distância	<p>2.3 Plano de Gestão de Dados (PGD)</p> <p>2.3.1 Videoaula sobre Plano de Gestão de Dados: dicas de elaboração</p> <p>2.3.2 Plano de Gestão de Dados: dicas de elaboração</p> <p>2.3.3 DMPonline</p> <p>2.3.4 Notas de aula</p> <p>2.3.5 Exemplos de PGD</p> <p>2.3.6 RIO Journal: Exemplo de periódico científico que publica Planos</p>	3h

			de Gestão de Dados	
Referências	<p>AVENTURIER, Pascal. Plano de Gestão de Dados : uma introdução. <i>In</i>: Publicação científica: <b>Blog do Pascal Aventurier sobre as publicações científicas e os dados de pesquisa</b>, [S.l. : s.n.], 2017. Disponível em: <a href="https://pubcient.hypotheses.org/1660">https://pubcient.hypotheses.org/1660</a>. Acesso em: 10 ago. 2019.</p> <p>DATAONE. <b>Data Life Cycle</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.dataone.org/data-life-cycle">https://www.dataone.org/data-life-cycle</a>. Acesso em: 10 ago. 2019.</p> <p>DIRETRIZES do CSE para promover integridade em publicações de periódicos científicos. São Paulo: ABEC Brasil, 2017. Disponível em: <a href="https://www.abecbrasil.org.br/arquivos/whitepaper_CSE.pdf">https://www.abecbrasil.org.br/arquivos/whitepaper_CSE.pdf</a>. Acesso em: 10 ago. 2019.</p> <p>GOFAIR. <b>FAIR Principles</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.go-fair.org/fair-principles/">https://www.go-fair.org/fair-principles/</a>. Acesso em: 10 ago. 2019.</p> <p>MICHENER, William K. Ten Simple Rules for Creating a Good Data Management Plan. <b>PLOS Computational Biology</b>, [S.l.], v. 11, n. 10, e1004525, 2015. DOI: <a href="https://doi.org/10.1371/journal.pcbi.1004525">https://doi.org/10.1371/journal.pcbi.1004525</a>. Disponível em: <a href="https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004525">https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004525</a>. Acesso em: 20 jul. 2019.</p> <p>SAYÃO, Luís Fernando; SALES, Luana Farias. <b>Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores</b>. Rio de Janeiro: CNEN/IN, 2015. 90 p. Disponível em: <a href="https://is.gd/pOSKQO">https://is.gd/pOSKQO</a>. Acesso em: 20 jul. 2017.</p> <p>SILVA, Fabiano Couto Corrêa da. <b>Gestão de dados científicos</b>. Rio de Janeiro: Interciência, 2019.</p> <p>UNIVERSITY OF SANTA CRUZ. University Library. <b>Research Data Management Lifecycle</b>. California: UC, 2019. Disponível em: <a href="https://guides.library.ucsc.edu/datamanagement">https://guides.library.ucsc.edu/datamanagement</a>. Acesso em: 10 ago. 2019.</p> <p>WIKIPÉDIA. [Fair Principles]. [S.l. : s.n.], 2016. Disponível em: <a href="https://pt.wikipedia.org/wiki/Ficheiro:FAIR_data_principles.jpg">https://pt.wikipedia.org/wiki/Ficheiro:FAIR_data_principles.jpg</a>. Acesso em: 10 set. 2019.</p>			
Aula 4: Gestão de Dados: Dicas de organização dos arquivos, medidas de segurança e licenças	11/09/2019 a 01/11/2019	Acompanhamento a distância	2.4 Gestão de Dados: dicas de organização dos arquivos, medidas de segurança dos dados e licenças 2.4.1 Notas de aula 2.4.2 Dados pessoais, dados sensíveis e dados pessoais sensíveis 2.4.3 ORCID - Identificador digital para autores 2.4.4 Licenças Creative Commons	2h
Referências	<p>BRASIL. <b>Lei Nº 13.709</b>, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <a href="http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm">http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm</a>. Acesso em: 10 ago. 2019.</p> <p>CESSDA Training Working Group (2017 - 2018). <b>CESSDA Data Management Expert Guide</b>. Bergen, Norway: CESSDA ERIC, [2017]. Disponível em: <a href="https://www.cessda.eu/DMGuide">https://www.cessda.eu/DMGuide</a>. Acesso em: 10 ago. 2019.</p> <p>DATAONE. <b>Investigator Toolkit</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.dataone.org/investigator-toolkit">https://www.dataone.org/investigator-toolkit</a>. Acesso em: 10 ago. 2019.</p> <p>DATAONE. <b>Software Tools Catalog</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.dataone.org/software_tools_catalog">https://www.dataone.org/software_tools_catalog</a>. Acesso em: 10 ago. 2019.</p>			



	<p>EDWARDS, Morgan. <i>In: Twitter. #overlyhonestmethods</i>. [S.l. : s.n.], 2013. Disponível em: <a href="https://twitter.com/mangoedwards/status/288694359451389953">https://twitter.com/mangoedwards/status/288694359451389953</a>. Acesso em: 10 ago. 2019.</p> <p>SAYÃO, Luís Fernando; SALES, Luana Farias. <b>Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores</b>. Rio de Janeiro: CNEN/ IEN, 2015. 90 p. Disponível em: <a href="https://is.gd/pOSKQO">https://is.gd/pOSKQO</a>. Acesso em: 20 jul. 2017.</p> <p>SILVA, Fabiano Couto Corrêa da. <b>Gestão de dados científicos</b>. Rio de Janeiro: Interciência, 2019.</p> <p>UNITED KINGDOM DATA ARCHIVE. <b>Recommended formats</b>. [S.l. : s.n.], 2019. Disponível em: <a href="https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats">https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats</a>. Acesso em: 10 ago. 2019.</p>			
Aula 5: Ciência Aberta	11/09/2019 a 01/11/2019	Acompanhamento a distância	<p>3 Ciência Aberta</p> <p>3.1 Podcast: Ciência Aberta e Gestão de Dados de Pesquisa</p> <p>3.1.1 Ciência Aberta: apresentação</p> <p>3.1.2 Notas de aula</p> <p>3.2 Tira-dúvidas online</p> <p>3.3 Redes Sociais Acadêmicas e Afins</p> <p>3.4 Plataformas colaborativas online no contexto da Ciência Aberta</p> <p>3.5 Plataformas de Escrita Colaborativa</p> <p>3.6 Ciência Aberta para Editores Científicos</p> <p>3.7 Avaliação por Pares Aberta</p> <p>3.8 Métricas Alternativas</p> <p>3.9 Projetos de Reprodutibilidade</p> <p>3.9.1 Caderno de Laboratório Eletrônico</p> <p>3.9.2 Dicas de leitura sobre Ciência Aberta</p> <p>3.9.3 Massively Open Online Papers</p> <p>3.9.4 Lista de cursos sobre Dados de Pesquisa e Ciência Aberta</p>	4h
Referências	<p>ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (org.). <b>Ciência aberta, questões abertas</b>. Brasília: Ibict; Rio de Janeiro: Unirio, 2015. Disponível em: <a href="https://bit.ly/2o2b6c4">https://bit.ly/2o2b6c4</a>. Acesso em: 10 maio 2018.</p> <p>BEZJAK, Sonja et al. <b>Open Science Training Handbook</b>. [S.l. : s.n.], 2018. Disponível em: <a href="https://foster.gitbook.io/manual-de-formacao-em-ciencia-aberta/">https://foster.gitbook.io/manual-de-formacao-em-ciencia-aberta/</a>. Acesso em: 20 abr. 2018 DOI: <a href="https://doi.org/10.5281/zenodo.1212496">https://doi.org/10.5281/zenodo.1212496</a>.</p> <p>FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER. <b>What is Open Science</b>. [S.l. : s.n.], 2015. Disponível em: <a href="https://tinyurl.com/y4l3l44g">https://tinyurl.com/y4l3l44g</a>. Acesso em: 7 set. 2018.</p> <p>FECHER, Benedikt; FRIESIKE, Sascha. Open Science: One Term, Five Schools of Thought. In: BARTLING, Sönke; FRIESIKE, Sascha. <b>Opening Science: the evolving guide on how the internet is changing research, collaboration and scholarly publishing</b>. New York: Springer, 2014. p. 17-47. Disponível em: <a href="https://goo.gl/yj2wYU">https://goo.gl/yj2wYU</a>. Acesso em: 20 jul. 2017.</p> <p>MAYER, Katja. <b>Open Innovation and the Creation of Commons</b>. [S.l.]: Creative Commons, 2016. Disponível em: <a href="https://creativecommons.org/2016/06/21/open-innovation-creation-commons/">https://creativecommons.org/2016/06/21/open-innovation-creation-commons/</a>. Acesso em: 10 ago. 2019.</p> <p>NOSEK, Brian; CHAMBERS, Chris. The first imperative: Science that isn't transparent isn't science. <b>The Guardian</b>, News, 2015. Não paginado. Disponível em: <a href="https://www.theguardian.com/science/head-quarters/2015/jun/25/the-first-imperative-science-that-is-nt-transparent-isnt-science">https://www.theguardian.com/science/head-quarters/2015/jun/25/the-first-imperative-science-that-is-nt-transparent-isnt-science</a>. Acesso em: 10 ago. 2019.</p>			

	<p>SANTOS, Anne Danielle Soares Clinio dos. <b>Novos cadernos de laboratório e novas culturas epistêmicas</b>: entre a política do experimento e o experimento da política. Rio de Janeiro, 2016. 240 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Escola de Comunicação. Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, 2016. Disponível em: <a href="https://bit.ly/2HLMetb">https://bit.ly/2HLMetb</a>. Acesso em: 11 maio 2019.</p> <p>TENNANT, Jonathan. <b>Who should own public science? Preprints, Power, and Publishers</b>. [S.l.]: Figshare, 29 nov. 2018. Disponível em: <a href="https://figshare.com/articles/Who_should_own_public_science_Preprints_Power_and_Publishers/7398506/1">https://figshare.com/articles/Who_should_own_public_science_Preprints_Power_and_Publishers/7398506/1</a>. Acesso em: 10 ago. 2019.</p> <p>UNITED NATIONS EDUCATIONAL, SCIENCE AND CULTURAL ORGANIZATION (Unesco). <b>Recursos educacionais abertos (REA)</b>. [S.l.: s.n.], 2017. Disponível em: <a href="https://bit.ly/2JLA4aF">https://bit.ly/2JLA4aF</a>. Acesso em: 11 maio 2018.</p> <p>WIKIVERSIDADE. <b>Ciência Aberta</b>. [S.l.: s.n.], 2017. Disponível em: <a href="https://pt.wikiversity.org/wiki/Portal:Ci%C3%Aancia_Aberta">https://pt.wikiversity.org/wiki/Portal:Ci%C3%Aancia_Aberta</a>. Acesso em: 10 ago. 2019.</p>			
Avaliação do curso	11/09/2019 a 01/11/2019	Acompanhamento a distância	4 Avaliação sobre o curso 4.1 Avaliação sobre o minicurso 4.2 Lousa Virtual (Jamboard)	-
<b>CARGA HORÁRIA TOTAL</b>				15h
<b>ENCONTROS PRESENCIAIS (AULAS PRESENCIAIS)</b>				
<b>DIA</b>	<b>HORÁRIO</b>	<b>PROGRAMAÇÃO</b>	<b>DESCRIÇÃO DA AULA</b>	<b>CARGA HORÁRIA</b>
-	-	-	-	-
<b>CRONOGRAMA DAS ATIVIDADES VIRTUAIS</b>				
<b>AULA</b>	<b>TIPO DE ATIVIDADE</b>	<b>PERÍODO</b>	<b>DESCRIÇÃO</b>	
Apresentação	Mural / Fórum	11/09/2019 a 01/11/2019	Tarefa: Inicialmente, é importante que todos/as os/as participantes se apresentem. Informe o seu nome, qual a sua área e o seu curso. Aproveitem e fiquem à vontade para falar quais são as suas expectativas em relação a este curso. Para isso, use o espaço dos comentários logo abaixo.	
Aula 1: Dados de Pesquisa	Enquete	11/09/2019 a 01/11/2019	Tarefa: Escolha uma das opções de horário para tirar dúvidas online via Hangouts.	
Aula 2: Dados de Pesquisa: onde buscar e compartilhar?	Pergunta / Tarefa	11/09/2019 a 01/11/2019	Tarefa: O que você pensa a respeito dos preprints?  Tarefa: Que tal fazer uma busca por dados no Google Dataset Search? Experimentem. Pesquisem sobre o assunto que quiserem e vejam o que poderão encontrar.	
Aula 3: Plano de Gestão de Dados (PGD)	Tarefa	11/09/2019 a 01/11/2019	Tarefa: Considero válido mostrar para vocês um exemplo de periódico científico que publica e/ou solicita o Plano de Gestão de Dados (PGD), em inglês, Research Data Management Plan (RDMP) ou Data Management Plan (DMP).	

			Abaixo estão os links de alguns PGD's publicados no RIO Journal que selecionei para mostrar aqui. Vocês também podem navegar pela plataforma do RIO Journal e buscar outros exemplos de PGD. Se algum chamar muito a sua atenção, deixe seus comentários aqui nessa postagem e compartilhe com todos o que você notou, gostou, achou curioso, ou até mesmo o que você acha que poderia ser melhor descrito e não foi no PGD escolhido. Ou vocês podem comentar sobre estes exemplos de PGD expostos aqui. aguardo os comentários de todos.
Aula 4: Gestão de Dados: Dicas de organização dos arquivos, medidas de segurança e licenças	Tarefa	11/09/2019 a 01/11/2019	Tarefa: ORCID: Atualmente, ter um identificador digital para o autor auxilia não apenas na integração com o Currículo Lattes e outras redes sociais acadêmicas: também já faz parte da rotina de muitos periódicos científicos solicitar aos autores que se cadastrem no ORCID e utilizem esse código para submeter artigos e para constar nas informações do manuscrito após a publicação. O ORCID também tem sido bastante usado para que o autor possa vincular o seu nome aos conjuntos de dados de seus trabalhos e no depósito dos datasets em repositórios de dados. Acessem a apresentação sobre o ORCID para saber mais a respeito. Se você ainda não tem um cadastro no ORCID, faça o quanto antes, afinal, muitos pesquisadores já estão usando.
Aula 5: Ciência Aberta	Chat / Hangouts	11/09/2019 a 01/11/2019	Tarefa: Tira-dúvidas online : minicurso Gestão de Dados de Pesquisa.
Avaliação do curso	Formulários Google  Lousa Virtual (Jamboard)	11/09/2019 a 01/11/2019	Tarefa: Chegou a hora de avaliar o curso e nos informar quais foram as suas impressões em relação ao conteúdo ministrado. A sua opinião é extremamente importante para que possamos aprimorar os conteúdos e melhorar cada vez mais. Não deixe de responder ao questionário de avaliação.  Tarefa: Essa é uma lousa virtual do nosso curso. Deixarei disponível para quem quiser deixar alguma mensagem final ou recadinhos sobre este minicurso. Sugiro que cliquem na opção 'nota autoadesiva' para deixar a sua mensagem.

## 10 MATERIAL DE APOIO

ARQUIVOS DISPONIBILIZADOS PARA AS AULAS		
AULA	NOME DO ARQUIVO	TIPO
<i>Apresentação</i>	Ambientação e informações gerais sobre o curso	Slides
<i>Aula 1: Dados de Pesquisa</i>	Vídeo de 1 minuto sobre Gestão de Dados	Vídeo
	Vídeoaula 1	Vídeoaula
	Slides Aula 1	Slides

<i>Aula 2: Dados de Pesquisa: onde buscar e compartilhar?</i>	Videoaula 2 Slides Aula 2	Videoaula Slides
<i>Aula 3: Plano de Gestão de Dados (PGD)</i>	Videoaula 3 Slides Aula 3 Notas de aula (3) Slides DMPonline	Videoaula Slides Notas de aula / Texto Slides
<i>Aula 4: Gestão de Dados: Dicas de organização dos arquivos, medidas de segurança e licenças</i>	Videoaula 4 Slides Aula 4 Notas de aula (4) Slides ORCID	Videoaula Slides Notas de aula / Texto Slides
<i>Aula 5: Ciência Aberta</i>	Podcast Aula 5 Slides Aula 5 Notas de aula (5)	Podcast Slides Notas de aula / Texto
<i>Avaliação do curso</i>	Slides Jamboard Lousa virtual	Slides Lousa virtual



\*Parte integrante do minicurso online intitulado "Dados de Pesquisa: dicas para gerenciar, organizar, compartilhar e preservar dados", ministrado por Juliana Soares Lima, Bibliotecária da Biblioteca de Ciências Humanas da Universidade Federal do Ceará (UFC) e mestranda do Programa de Pós-Graduação em Ciência da Informação da UFC (PPGCI-UFC).

\*\*Este é o modelo elaborado para documentar o curso.



**5. 5 - Informe seu grau de satisfação com os itens abaixo: \***

Marcar apenas uma oval por linha.

	Ruim	Regular	Bom	Ótimo	Excelente
Clareza das explicações	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Plataforma utilizada	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Material utilizado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comunicação com a tutora	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Duração do curso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informações, conteúdos e fontes de informação indicadas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**6. 6 - Atribua uma nota para o curso de acordo com o seu nível geral de satisfação. \***

De 0 a 10, que nota você daria para o curso?

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**7. 7 - Você teria disponibilidade e/ou gostaria de estar presente em um encontro presencial para falar sobre este curso com a pesquisadora responsável? Se sim, informe o seu e-mail de contato a seguir. Se a resposta for não, pode deixar o campo em branco.**

\_\_\_\_\_

**8. 8 - Por favor, registre aqui suas impressões gerais sobre o curso e comentários adicionais (Sugestões, elogios e críticas).**

A sua opinião é fundamental para ajudar a aprimorar os conteúdos apresentados e a qualidade do curso.

\_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

## APÊNDICE E - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO



### UNIVERSIDADE FEDERAL DO CEARÁ PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO MESTRADO EM CIÊNCIA DA INFORMAÇÃO

#### Termo de Consentimento Livre e Esclarecido

Prezado(a) Pesquisador(a),

Você está sendo convidado(a) a participar da pesquisa intitulada “Gestão de Dados de Pesquisa e Ciência Aberta: Percepção dos Pesquisadores da Universidade Federal do Ceará”. Este estudo está sendo desenvolvido por mim, Juliana Soares Lima, mestranda no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará (UFC), sob a orientação da Professora Doutora Maria Giovanna Guedes Farias. O objetivo geral se concentra em investigar a percepção, as práticas e as necessidades informacionais concernentes ao gerenciamento de dados científicos e à Ciência Aberta, a fim de propor a criação de serviços e de um programa de suporte e apoio ao pesquisador na Universidade Federal do Ceará.

Destaco que a sua participação nesta pesquisa é inteiramente voluntária. Suas respostas serão utilizadas exclusivamente para fins científicos. Asseguro manter a integridade e o sigilo quanto às informações prestadas, além disso, garanto que os dados serão anonimizados e não serão divulgadas informações que estejam relacionadas à sua intimidade. Caso aceite participar, não haverá qualquer prejuízo para você em face das informações fornecidas.

**Riscos:** O risco relacionado à sua participação na pesquisa refere-se ao sentimento de desconforto que poderá decorrer da manifestação sobre assuntos pessoais. Nesse sentido, caso considere muito pessoal ou confidencial, ressalto que você tem toda liberdade para retirar o seu consentimento e interromper as respostas a qualquer momento, mesmo que tenha concordado anteriormente.

**Benefícios:** Sua participação e colaboração são imprescindíveis, além de permitir ampliar a compreensão sobre o tema em estudo. Inclusive, contribuirá para a criação de novos serviços e de um programa de apoio e suporte ao pesquisador no gerenciamento de dados

científicos e em práticas da Ciência Aberta na UFC a fim de melhorar os processos de integridade, verificabilidade e reprodutibilidade da pesquisa científica.

Os dados e resultados desta pesquisa estarão sempre sob sigilo ético. Os arquivos serão anonimizados, criptografados e armazenados no computador pessoal da pesquisadora e em serviços de armazenamento externos por um período mínimo de cinco anos, a fim de preservar o acesso indevido ao seu conteúdo e para qualquer solicitação de verificação e integridade do referido estudo. A divulgação de alguns relatórios e informes, assim como qualquer outro documento ou publicação elaborada acerca dos resultados de pesquisa serão disponibilizados em repositórios de dados de pesquisa em até dois anos após o término deste estudo. Todas as estratégias de preservação dos dados estão descritas em um Plano de Gestão de Dados (PGD), disponível publicamente nos sites: DMPonline ([https://dmponline.dcc.ac.uk/public\\_plans](https://dmponline.dcc.ac.uk/public_plans)) e DMPTool ([https://dmptool.org/public\\_plans](https://dmptool.org/public_plans)).

Os resultados deste estudo serão publicados em minha dissertação de mestrado a ser defendida na UFC, e, após a defesa, estará disponível no Repositório Institucional (<http://www.repositorio.ufc.br/>). Solicito, por meio deste termo, a sua permissão para utilizar os dados em minha dissertação de mestrado. Outras formas, tais como comunicação a ser apresentada em congresso e publicações científicas (artigos, livros e capítulos) poderão ser produzidas a partir dos resultados deste estudo.

Durante o andamento da pesquisa, se por qualquer motivo você resolver desistir, tem toda liberdade para retirar o seu consentimento. Esta proposta foi encaminhada ao Comitê de Ética em Pesquisa (CEP) da Universidade Federal do Ceará, cuja finalidade é garantir que os participantes estejam protegidos de qualquer risco ou prejuízo à sua pessoa, e foi aprovada conforme Parecer N°. 3.512.152/2019.

Em caso de dúvidas, coloco-me à disposição para prestar esclarecimentos sobre a pesquisa a qualquer tempo, por e-mail: [juliana.lima@ufc.br](mailto:juliana.lima@ufc.br) ou [julia10br@gmail.com](mailto:julia10br@gmail.com). Endereço: Avenida da Universidade, 2683, Biblioteca de Ciências Humanas, Benfica, Fortaleza, Ceará, CEP: 60.020-970. Se desejar, pode solicitar informações no Comitê de Ética e Pesquisa da Universidade Federal do Ceará no telefone: (85) 3366-8346 e endereço: Rua Coronel Nunes de Melo, 1000, Rodolfo Teófilo, Fortaleza, Ceará, CEP: 60.430-270.

Esta pesquisa está em conformidade com a Resolução CNS N°. 510, de 7 de abril de 2016, disponível em: <http://bit.ly/2GZnSzk>.

**Endereço da responsável pela pesquisa:**

Nome: Juliana Soares Lima

Instituição: Universidade Federal do Ceará

Endereço: Avenida da Universidade, 2683, Biblioteca de Ciências Humanas, Benfica, Fortaleza, Ceará, CEP: 60.020-970.

Telefone para contato: (85) 3366-7658.



**ATENÇÃO:** Se você tiver alguma consideração ou dúvida, sobre a sua participação na pesquisa, entre em contato com o Comitê de Ética em Pesquisa da UFC/PROPESQ – Rua Coronel Nunes de Melo, 1000 - Rodolfo Teófilo, fone: 3366-8346/44. (Horário: 08:00-12:00 horas de segunda a sexta-feira). O CEP/UFC/PROPESQ é a instância da Universidade Federal do Ceará responsável pela avaliação e acompanhamento dos aspectos éticos de todas as pesquisas envolvendo seres humanos.

O abaixo assinado \_\_\_\_\_,  
 \_\_\_\_\_ anos, RG: \_\_\_\_\_, declara que é de livre e espontânea vontade que está como participante de uma pesquisa. Eu declaro que li cuidadosamente este Termo de Consentimento Livre e Esclarecido e que, após sua leitura, tive a oportunidade de fazer perguntas sobre o seu conteúdo, como também sobre a pesquisa, e recebi explicações que responderam por completo minhas dúvidas. E declaro, ainda, estar recebendo uma via assinada deste termo.

Fortaleza, \_\_\_\_/\_\_\_\_/\_\_\_\_.

Nome do participante da pesquisa

Data

Assinatura

Nome do pesquisador principal

Data

Assinatura

Nome do Responsável legal/testemunha Data  
 (se aplicável)

Assinatura

Nome do profissional

Data

Assinatura  
 que aplicou o TCLE

## APÊNDICE F – PLANO DE GESTÃO DE DADOS

---

### **Gestão de Dados de Pesquisa no contexto da Ciência Aberta: percepção dos pesquisadores da Universidade Federal do Ceará**

*A Data Management Plan created using DMPonline*

**Creator:** Juliana Soares Lima

**Affiliation:** Other

**Template:** Digital Curation Centre

**ORCID iD:** 0000-0001-9399-673X

**Project abstract:**

Essencialmente, a ciência tem como princípio a colaboração entre os cientistas no desenvolvimento de pesquisas, fator crucial para o avanço do conhecimento e que promove práticas abertas, facilitando a interconexão entre os pares, aumentando a visibilidade do pesquisador e de seus trabalhos, agilizando o acesso rápido aos dados e softwares, além de criar oportunidades de interação e de contribuição em projetos. Diante disso, apresenta-se uma proposta de estudo que visa investigar qual a percepção dos pesquisadores (Professores e alunos de PósGraduação Stricto Sensu) da Universidade Federal do Ceará (UFC) em relação aos princípios da Ciência Aberta e sobre o compartilhamento e gerenciamento de dados de pesquisa. O objetivo geral se concentra em investigar as práticas e necessidades informacionais concernentes à ciência aberta e ao gerenciamento de dados científicos, a fim de propor a criação de serviços de dados e um programa de suporte e apoio ao pesquisador na Universidade Federal do Ceará (UFC). Como estratégia metodológica, adotou-se a triangulação de métodos – Teoria Fundamentada em Dados e a Netnografia –, e como instrumentos de coleta de dados serão utilizados questionários, entrevista semiestruturada individual e diário de campo eletrônico (OneNote) e o uso do caderno eletrônico de laboratório RSpace para o registro de observações e notas de campo, de forma que possibilite uma ampla visão sobre o objeto de estudo e os sujeitos investigados. Na fase final deste estudo pretende-se criar e implantar uma ação de informação que visa a capacitação dos pesquisadores, por exemplo, a demonstração de como preencher e elaborar um Plano de Gestão de Dados (PGD), bem como desenvolver um programa de apoio ao pesquisador de forma personalizada e com base nas necessidades informacionais identificadas na instituição. Esta pesquisa foi submetida ao Comitê de Ética em Pesquisa da Universidade Federal do Ceará, tendo sido aprovada, conforme o Parecer N°. 3.512.152/2019.

**Last modified:** 22-02-2020

## **Gestão de Dados de Pesquisa no contexto da Ciência Aberta: percepção dos pesquisadores da Universidade Federal do Ceará**

---

### **Data Collection**

#### **What data will you collect or create?**

Serão coletados dados textuais e observacionais, coletados por meio de entrevistas individuais semiestruturadas e questionário. Além disso, serão registradas as anotações de campo e observacionais em diário eletrônico via OneNote e caderno eletrônico de laboratório via RSpace.

#### **How will the data be collected or created?**

Os dados serão coletados por meio de entrevista individual semiestruturada, questionário e observações realizadas em ambiente virtual de aprendizagem. O software OneNote será utilizado para fazer o registro de todas as anotações observacionais e de campo, assim como auxiliar a elaboração de memorandos para posterior análise documental.

### **Documentation and Metadata**

#### **What documentation and metadata will accompany the data?**

Serão disponibilizados junto com os datasets um arquivo Readme.txt, Codebook e planilhas em formato aberto (.ods).

### **Ethics and Legal Compliance**

#### **How will you manage any ethical issues?**

A fim de cumprir as questões éticas e legais referentes à pesquisa, o estudo foi enviado ao Comitê de Ética em Pesquisa da Universidade Federal do Ceará, tendo sido aprovado, conforme o Parecer N°. 3.512.152/2019. Também será solicitado de cada participante a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE). Os dados obtidos serão anonimizados e codificados de forma que possa garantir a integridade e o sigilo dos participantes. Os dados serão armazenados em pasta protegida por criptografia e senha em computador pessoal da pesquisadora, e, uma cópia da pasta e todos os documentos serão salvos na nuvem (Dropbox e Google Drive) e em repositório privado no GitHub (neste caso, os arquivos textuais estarão redigidos em Markdown). O backup será feito diariamente. Os dados serão mantidos por até cinco anos e após esse período, os dados poderão ser excluídos e destruídos definitivamente por meio do software livre Eraser ou similar, seguidos da destruição física de todos os componentes que abrigaram os arquivos de pesquisa. A cópia impressa dos TCLEs mantidas durante os cinco anos também serão incineradas após o período pré-estabelecido, completando o ciclo do processo de destruição física dos materiais relacionados à pesquisa.

#### **How will you manage copyright and Intellectual Property Rights (IPR) issues?**

Os dados obtidos no estudo serão mantidos pela pesquisadora. Será claramente expresso aos participantes da pesquisa e nos termos de consentimento livre e esclarecido que todos os dados coletados serão utilizados apenas para os fins desta pesquisa, bem como para os produtos oriundos dela, entre artigos científicos, capítulos de livros e trabalhos apresentados em eventos. Por serem dados

que contém determinadas informações que serão coletadas a partir de entrevistas e depoimentos, poderão ser aplicadas algumas restrições de acesso e divulgação, assim como será aplicada a anonimização e codificação dos dados a fim de manter o sigilo e integridade da pesquisa. Os dados serão preparados para compartilhamento, e, portanto, serão acompanhados de uma licença flexível, como a CC-BY (creative commons). Sobre os direitos de Propriedade Intelectual, se houver qualquer problema em relação a esse quesito, deverão ser resolvidos em foro específico.

## **Storage and Backup**

### **How will the data be stored and backed up during the research?**

Os dados serão armazenados em computador pessoal em uma pasta protegida por senha e criptografia, e, além disso, essa pasta e todos os documentos contidos nela serão copiados e salvos em serviços de armazenamento na nuvem (Dropbox e Google Drive). Com a finalidade de garantir a segurança desses dados em caso de qualquer incidente, todos os dados de pesquisa também serão armazenados em repositório privado no GitHub. A frequência do backup será diária. A pesquisadora será a responsável pelo backup e recuperação dos arquivos.

### **How will you manage access and security?**

Todos os arquivos serão criptografados. O acesso aos dados será feito através de login e senha de acesso, tanto no computador pessoal como nos serviços de armazenamento na nuvem (Dropbox e Google Drive) e em repositório GitHub. Google Drive e GitHub possibilitam verificar o número de versões de um arquivo, facilitando esse controle em possíveis auditorias que possam vir a ser necessárias nos arquivos da pesquisa.

## **Selection and Preservation**

### **Which data are of long-term value and should be retained, shared, and/or preserved?**

Todos os dados obtidos por meio de entrevista, questionário, observações e notas de campo, inclusive registros oriundos do ambiente virtual de aprendizagem a ser observado serão mantidos. Todos os dados serão anonimizados e codificados de forma que possa garantir a integridade e o sigilo dos participantes. Os dados serão armazenados em pasta protegida por criptografia e senha em computador pessoal da pesquisadora, e, uma cópia da pasta e todos os documentos serão salvos na nuvem por meio do Dropbox e Google Drive e também em um repositório privado no GitHub. O backup será feito diariamente. Após essas fases, os dados serão preparados para compartilhamento em repositórios de dados como o Dataverse e Zenodo. Os dados serão retidos por até dois anos e mantidos por até cinco anos. Após esse período, os dados poderão ser excluídos do computador pessoal da pesquisadora e destruídos definitivamente por meio do software livre Eraser ou similar e por meio da destruição física de documentos impressos (Termo de Consentimento Livre e Esclarecido - TCLE), pendrives e outras mídias em que os dados estiverem armazenados.

### **What is the long-term preservation plan for the dataset?**

Os dados preparados e processados serão disponibilizados em formato aberto, em repositórios de dados como o Dataverse e Zenodo após dois anos a contar a partir da conclusão desta pesquisa. Não haverá custos ou taxas adicionais, e, se houver, serão custeados com recursos próprios.

## **Data Sharing**

**How will you share the data?**

Os dados poderão ser usados por pesquisadores da área de Ciência da Informação ou afins que investiguem temáticas semelhantes. Os dados serão disponibilizados em formato aberto, em repositórios de dados como o Dataverse e Zenodo após dois anos a contar a partir da conclusão desta pesquisa. Cada um dos datasets receberá um Digital Object Identifier (DOI) e serão compartilhados por meio de uma licença flexível CC-BY (Creative Commons). Serão acompanhados dos dados arquivos instrucionais e informativos de como utilizá-los, tais como README.txt, Codebooks e planilhas.

**Are any restrictions on data sharing required?**

Nenhum dado pessoal ou dados sensíveis serão disponibilizados. Quanto aos dados que serão compartilhados, se usados e reutilizados, deverão ser citados e os créditos de atribuição deverão ser informados. Os dados serão mantidos por dois anos a contar a partir da data do término desta pesquisa e somente após esse período será amplamente divulgado em repositórios de dados como o Dataverse e Zenodo. Qualquer outro problema, dúvida ou casos omissos deve ser informado para a pesquisadora por meio do e-mail: [juliana.lima@ufc.br](mailto:juliana.lima@ufc.br).

**Responsibilities and Resources****Who will be responsible for data management?**

A pesquisadora será responsável pela implementação deste Plano de Gestão de Dados e garante que este documento será revisado e revisado durante todo o processo de pesquisa. Todas as atividades e responsabilidades ficarão a cargo da pesquisadora, tais como a captura dos dados, descrição dos metadados, qualidade dos dados, armazenamento e backup, arquivamento e compartilhamento dos dados de pesquisa. Quaisquer outras dúvidas e casos omissos, entrar em contato com a pesquisadora via e-mail: [juliana.lima@ufc.br](mailto:juliana.lima@ufc.br).

**What resources will you require to deliver your plan?**

A pesquisadora será responsável por selecionar e disponibilizar os dados em repositórios de dados em acesso livre e aberto. Os softwares demandados para a aplicação da pesquisa são: OneNote (diário de campo eletrônico), RSpace (caderno de laboratório eletrônico), serviços de armazenamento na nuvem (Dropbox e Google Drive), GitHub (Repositório disponível na nuvem e também com uma versão para computador pessoal), planilhas Google Docs, Publish or Perish (software bibliométrico e que calcula o índice H para autores) utilizado para efetuar a busca para compor a revisão de literatura, Atlas.TI (software usado na análise qualitativa dos dados).

## APÊNDICE G – LISTA DE CURSOS SOBRE CIÊNCIA ABERTA E GESTÃO DE DADOS DE PESQUISA<sup>153</sup>

LISTA DE CURSOS: CIÊNCIA ABERTA, GESTÃO E COMPARTILHAMENTO DE DADOS				
INSTITUIÇÃO	CURSO	PLATAFORMA	CERTIFICAÇÃO	ACESSO
4TU Centre for Research Data e Research Data Netherlands	<a href="#">Essentials 4 Data Support</a>	Site da Instituição / Site próprio	Sim*	Módulos online são gratuitos sem certificação; Módulos online + 2 dias de curso presencial são pagos e com direito a certificado
CESSDA Training	<a href="#">Vários cursos</a>	Site da Instituição	Não	Gratuito
Data Documentation Initiative (DDI)	<a href="#">Alguns treinamentos e Workshops</a>	Site da Instituição / Site próprio	Não	Gratuito
DataONE	<a href="#">Vários módulos</a>	Site da Instituição / Site próprio	Não	Gratuito
Fiocruz / Universidade do Minho	<a href="#">Formação Modular em Ciência Aberta</a>	Site da Instituição / Site próprio / Moodle	Sim	Gratuito
FOSTER	<a href="#">Vários cursos</a>	Site da Instituição e Plataforma Moodle	Não	Gratuito sem certificação; oferece badges (a cada módulo completo o usuário conquista uma badge)
Jonathan Tennant	<a href="#">Open Science MOOC</a>	Eliademy	Sim	Gratuito, oferta certificado
Ministério da Ciência, Tecnologia e Ensino Superior da República Portuguesa	<a href="#">Mooc Ciência Aberta</a>	Site da Instituição / Site próprio	Não	Gratuito, não oferta certificado
TU Delft	<a href="#">Open Science: Sharing Your Research with the World (EdX)</a>	edX	Sim*	Gratuito sem certificação; Pago com certificado
University of Edinburgh	<a href="#">MANTRA</a>	Site da Instituição / Site próprio	Não	Gratuito, não oferta certificado

<sup>153</sup> Uma lista com essas sugestões de cursos está disponível no Figshare: <https://doi.org/10.6084/m9.figshare.9962117>.


LISTA DE CURSOS: CIÊNCIA ABERTA, GESTÃO E COMPARTILHAMENTO DE DADOS				
INSTITUIÇÃO	CURSO	PLATAFORMA	CERTIFICAÇÃO	ACESSO
University of Edinburgh e University of North Carolina at Chapel Hill	<a href="#">Research Data Management and Sharing - MOOC</a>	Coursera	Sim*	Gratuito sem certificação; Pago com certificado
University of Edinburgh, Digital Curation Centre (DCC), Research Data Netherlands, 4TU.Centre for Research Data, Data Archiving and Networked Services (DANS) and SURFsara	<a href="#">Delivering Research Data Management Services</a>	Future Learn	Sim*	Gratuito sem certificação; Pago com certificado
Aprender3C	<a href="#">MOOC de acceso abierto</a>	Site da Instituição / Site próprio	Sim	Gratuito
Eurodoc	<a href="#">Eurodoc Open Science Ambassador Training</a>	Site da Instituição / Site próprio	Não	Gratuito
Countway Medical Library, Harvard Medical School, Simmons University e Elsevier	<a href="#">Research Data Management Librarian Academy (RDMLA)</a>	Canvas Network	Sim	Gratuito com certificação; oferece badges (a cada módulo completo o usuário conquista uma badge)
UK Data Service	<a href="#">Datatree - Data Training Engaging End-users</a>	UK Data Service	Sim	Gratuito com certificação; oferece badges (a cada módulo completo o usuário conquista uma badge)
Universidade do Minho	<a href="#">O essencial da Gestão de Dados de Investigação</a>	Nau / OpenEDX	Sim	Gratuito, oferta certificado


Fonte: Elaborado pela autora (2019).

## APÊNDICE H – PROPOSTA DE PROGRAMA DE GESTÃO DE DADOS DE PESQUISA



# Programa de Gestão de Dados de Pesquisa - UFC

 **Juliana Soares Lima**  
Universidade Federal do Ceará  
Biblioteca de Ciências Humanas  
Fortaleza - CE

 **Maria Giovanna Guedes Farias**  
Universidade Federal do Ceará  
Programa de Pós-Graduação em Ciência da Informação  
Fortaleza - CE



## 1 Apresentação

O Programa de Gestão de Dados de Pesquisa (PGDP) da Universidade Federal do Ceará (UFC), constitui-se em um produto gerado a partir da dissertação de Juliana Soares Lima, sob orientação da Profa. Dra. Maria Giovanna Guedes Farias, Programa de Pós-Graduação da Universidade Federal do Ceará (PPGCI-UFC).

A proposta do programa se concentra no desenvolvimento de ações formais de propostas educacionais voltadas para a gestão de dados de pesquisa na UFC. Nesse sentido, visa promover uma série de ações de educação e informação, incluindo ainda a elaboração de políticas, diretrizes, materiais instrucionais e por fim a sugestão de criação de serviços de apoio e suporte ao pesquisador na UFC que venham a contribuir para a adoção das práticas de gestão de dados de pesquisa na instituição.

Este programa baseia-se ainda na abordagem da mediação da informação na Ciência da Informação, na qual “[...] o bibliotecário atuaria como um “fornecedor” de informações que atendem as demandas e necessidades dos usuários.” (FARIAS; FARIAS, 2017, p. 333).

Proposta por Silva (2015), Silva e Farias (2017) e também aplicada em um programa de mediação técnica da informação por Araújo (2018), a mediação da informação é caracterizada como um conjunto de práticas construtivas, um programa de ação, intervenção e interferência do bibliotecário, por exemplo, empregada nos serviços de informação, constituindo-se como uma das principais atividades técnica, pedagógica, institucional e pragmática da mediação. Ademais, a mediação é voltada para os estudos de usuários e para a busca de uma aproximação da comunidade atendida, inclusive no incentivo à autonomia dos usuários, a exemplo da realização de capacitações e treinamentos como uma das formas de os usuários apreenderem e apropriarem-se da informação, além de participarem da construção coletiva de conhecimento na universidade. No caso

---

específico deste programa, a mediação da informação contribuirá sobremaneira para a organização, gestão da informação e gerenciamento dos dados de pesquisa na instituição.

Este programa está em consonância com o Plano de Desenvolvimento Institucional da UFC (PDI 2018-2022) e baseia-se em uma série de referências internacionais que recomendam a criação e realização de cursos, formações modulares e treinamentos nesse âmbito, a exemplo do relatório elaborado pela European Commission e o Working Group on Education and Skills under Open Science (2017), intitulado “Providing researchers with the skills and competencies they need to practice Open Science”; e o documento “Implementing Research Data Management Policies Across Europe: Experiences from Science Europe Member Organisations” publicado pela Science Europe (2020); além das políticas e orientações de universidades internacionais (University of Edinburgh - Escócia; University of Queensland - Austrália; Delft University of Technology - TUDELFT - Holanda; Universidade do Minho - Portugal) e treinamentos consagrados como o MANTRA, Open Science MOOC, FOSTER, Formação Modular em Ciência Aberta da Fiocruz, entre outras.

## 1.1 Justificativa

Adota-se neste programa a perspectiva do desenvolvimento de ações de educação e informação para promover a conscientização (*awareness*) dos pesquisadores por meio de capacitações presenciais e a distância que irão ajudá-los a adquirir as competências necessárias para o gerenciamento dos dados de pesquisa através da mediação bibliotecária aplicada nessas ações a fim de que promovam a competência em informação e a competência em dados. A proposta justifica-se uma vez que a partir da conscientização e conhecimento dessas práticas os pesquisadores de diversas áreas se sentirão mais capazes e confiantes para compreender o ciclo de vida da pesquisa, o ciclo de vida dos dados, executar o

gerenciamento de seus dados de pesquisa, elaborar seus Planos de Gestão de Dados (PGD's), escolher os formatos ideais para seus arquivos, conhecer as estratégias de versionamento, segurança e licenciamento dos dados, seguir padrões e protocolos, entre outras atividades concernentes a esse processo.

## 1.2 Objetivos

1. Formular políticas e diretrizes para a Gestão dos Dados de Pesquisa na UFC;
2. Propor ações educacionais e de informação presenciais e à distância (e-learning e blended learning) sobre a Gestão dos Dados de Pesquisa para a comunidade acadêmica da Universidade Federal do Ceará;
3. Formar, capacitar, treinar e preparar a equipe de trabalho das bibliotecas que prestarão os serviços de dados, curadoria e de apoio ao pesquisador para a sua respectiva comunidade acadêmica na UFC;
4. Desenvolver serviços de dados e de apoio ao pesquisador e prestar consultoria, fornecer aconselhamento, orientação e ações de educação, treinamento, informação e congêneres à comunidade acadêmica da UFC no que concerne à gestão de dados de pesquisa e aspectos relacionados;
5. Firmar parcerias e redes de trabalho na própria instituição com outros setores e os cursos, departamentos, Pró-Reitoria de Pesquisa e Pós-Graduação e Pró-Reitoria de Graduação para promover e fortalecer o gerenciamento dos dados de pesquisa na UFC;
6. Firmar parcerias nacionais e/ou internacionais que venham a contribuir com este programa;
7. Facilitar a adoção e implementação de políticas e diretrizes que favoreçam o acesso aberto, livre e gratuito ao conhecimento científico, inclusive visando estar em conformidade com as recomendações, exigências e normas institucionais, das agências de fomento, dos financiadores e do governo.

## 1.3 Diretrizes

- Realização de ações de educação e informação sobre a gestão dos dados de pesquisa;

- Formação, capacitação, treinamento e preparação da equipe de trabalho;
- Estabelecimento de parcerias, rede de trabalho institucionais e setoriais para promover a gestão dos dados de pesquisa;
- Desenvolvimento dos serviços de dados e de apoio e suporte ao pesquisador na UFC;
- Formulação de políticas e diretrizes para a gestão dos dados de pesquisa na instituição.

## 2 Metodologia

A elaboração do programa se deu com base nos resultados da pesquisa de campo realizada durante o mestrado e sua estruturação está em conformidade com os temas abordados na dissertação (gestão e compartilhamento de dados de pesquisa no contexto da Ciência Aberta).

## 3 Estratégias

O programa será primordialmente assentado na sensibilização como um dos meios mais eficazes para comunicar e informar sobre o trabalho e a natureza da gestão de dados na UFC, incluindo a formulação de diretrizes e políticas ao longo do processo.

O conteúdo de cada curso / treinamento será adaptado para atender a quatro públicos distintos:

- 1º Nível:
  - Reitoria, Pró-Reitoria de Pesquisa e Pós-Graduação, Pró-Reitoria de Graduação, Colégio de Estudos Avançados e demais instâncias universitárias;
- 2º Nível:

- Docentes;
- 3º Nível:
  - Discentes da Pós-Graduação;
  - Discentes da Graduação;
- 4º Nível:
  - Bibliotecários.

O primeiro nível será capacitado para lidar com as questões normativas e políticas no âmbito da alta gestão. O segundo e o terceiro serão treinados e receberão informações sobre os serviços oferecidos pela biblioteca (consultoria, suporte e apoio ao pesquisador, auxílio no preenchimento do Plano de Gestão de Dados, retenção e compartilhamento dos dados de pesquisas em repositórios adequados etc.) e instruções de como proceder à gestão dos dados de suas pesquisas. O quarto nível participará de uma formação mais técnica e voltada para prestar o suporte e instaurar os serviços de dados em seu campus / faculdade / curso e assim apoiar as atividades de pesquisas de seus usuários. Precipuaemente, é preciso formar a mão de obra necessária para dar conta da demanda existente na instituição.

De maneira geral, as atividades ocorrerão por meio de reuniões, workshops, treinamentos presenciais e a distância, entre outros tipos. O intuito de cada uma dessas atividades é focar nas oportunidades e benefícios oferecidos através dos novos serviços de gerenciamento de dados de pesquisa, uma vez que a crítica dessas práticas não conduz a nenhum lugar de destaque e não contribui para o avanço na ciência.

#### **4 Estrutura do Programa**

De forma geral, o escopo desse programa integra a seguinte estrutura: Elaboração de Políticas e Diretrizes, além do estabelecimento de normas e padrões

para facilitar o processo de gestão dos dados de pesquisa na instituição; Ações de Educação e Informação; Produtos; Serviços e Gestão.



- **Políticas e diretrizes**

Formulação de uma Política de Gestão dos Dados de Pesquisa e Diretrizes para a Gestão dos dados na instituição.

- **Ações de Educação e Informação**

Elaboração e realização de capacitações, cursos e treinamentos presenciais e online sobre a Gestão dos Dados de Pesquisa. Prevê ainda a formação e preparação da equipe de trabalho para atuar na prestação dos serviços de dados e de apoio ao pesquisador na UFC (consultoria, aconselhamento, curadoria e orientações sobre a gestão dos dados de pesquisa).

- **Produtos**

Criação de produtos e um portfólio de cursos especializados, totalmente focado nas necessidades de treinamento para a Gestão dos Dados de Pesquisas Científicas desenvolvidas na universidade, além da elaboração de materiais instrucionais e afins para as capacitações a serem desenvolvidas.

- **Serviços de dados e apoio ao pesquisador**

Desenvolvimento dos serviços de dados e de apoio e suporte ao pesquisador na UFC (consultoria, aconselhamento, curadoria e orientações sobre a gestão dos dados de pesquisa).

- **Gestão**

Realização de levantamento, mapeamento e diagnóstico institucional prévio sobre a produção dos dados de pesquisas desenvolvidas na UFC. A etapa de diagnóstico verificará quais as possíveis limitações, as condições de infraestrutura e recursos humanos para as atividades de gestão dos dados de pesquisa, assim como a avaliação de desempenho do programa.

A proposta de programa prevê ainda a sua articulação com a comunidade acadêmica a fim de angariar voluntários e apoiadores da causa, promovendo a prática do *advocacy* em torno da gestão dos dados de pesquisa e fortalecendo os laços entre todos os atores envolvidos.

Dessa forma, reconhece-se o valor da 'ecologia de saberes', preconizada por Boaventura de Sousa Santos, e o quanto tem extrema relevância no cerne desta proposta, tendo em vista que "[...] promover o diálogo entre vários saberes que podem ser considerados úteis para o avanço das lutas sociais [e] pelos que nelas intervêm" (SANTOS, 2014, p. 332). Em outras palavras, trata-se de um processo coletivo de construção e produção do conhecimento. Portanto, a integração, colaboração e a pluralidade de saberes existentes na universidade e no mundo corrobora com o esforço de superação das desigualdades sociais através do compartilhamento do conhecimento científico produzido na UFC, sendo

exatamente este o papel da universidade, especialmente no cumprimento de sua responsabilidade social e de seus três pilares: ensino, pesquisa e extensão.

## 5 Avaliação

Inicialmente, será realizado um diagnóstico institucional com a finalidade de identificar e mapear os ambientes produtores de dados de pesquisa, quais as possíveis limitações, condições de infraestrutura e recursos humanos para as atividades de gestão dos dados de pesquisa; e posteriormente pela avaliação comparativa a cada ano para acompanhar o período de execução proposto para o início do programa (previsto para 2021). Durante esse processos, outras formas de avaliação poderão ser empregadas para avaliar o impacto, desempenho e resultados obtidos no decurso do programa.

## 6 Prazos

- Data de início: 06/2021.
- Data de término: 12/2023.

## Referências

ARAÚJO, Ana Rafaela Sales de. **Mediação no âmbito da organização da informação:** proposta de um programa de atuação para o Sistema de Bibliotecas da Universidade Federal do Ceará. 2018. 208 f. Dissertação (Mestrado) – Universidade Federal do Cariri, Centro de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Biblioteconomia, Mestrado Profissional em Biblioteconomia, Juazeiro do Norte, 2018. Disponível em: [http://www.repositorio.ufc.br/bitstream/riufc/38110/1/2018\\_dis\\_arsaraujo.pdf](http://www.repositorio.ufc.br/bitstream/riufc/38110/1/2018_dis_arsaraujo.pdf). Acesso em: 7 set. 2019.

EUROPEAN COMMISSION. Working Group on Education and Skills under Open Science. **Providing researchers with the skills and competencies they need to practise Open Science:** Open Science Skills Working Group Report. Brussels: European Union, 2017. Disponível em: <https://tinyurl.com/y3byn4eh>. Acesso em: 7 set. 2018.



---

FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER. **Planos de Gestão de Dados no Horizonte 2020**. 2019. Courses. Disponível em: <https://www.fosteropenscience.eu/node/2643>. Acesso em: 10 mar. 2019.

FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH – FOSTER. **What is Open Science**. 2015. Disponível em: <https://tinyurl.com/y4l3l44g>. Acesso em: 7 set. 2018.

FARIAS, Maria Giovanna Guedes; FARIAS, Gabriela Belmont de. Mediação na Ciência da Informação: uma análise bibliométrica na coleção Benancib. **RICI: R.Ibero-amer. Ci. Inf.**, Brasília, v. 10, n. 2, p. 332-349, jul./dez. 2017. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/download/2551/2278/>. Acesso em: 10 jan. 2018.

SANTOS, Boaventura de Sousa. Um ocidente não-ocidentalista? A filosofia à venda, a douta ignorância e a aposta de Pascal. *In*: Santos, Boaventura de Sousa, Meneses, Maria Paula (org.). **Epistemologias do Sul**. São Paulo: Cortez, 2010.

SANTOS, Boaventura de Sousa. A Práxis da Ecologia de Saberes: entrevista de Boaventura de Sousa Santos. **Tempus: actas de Saúde Coletiva**, Brasília, v. 8, n. 2, p. 331-338, 2014. Entrevista concedida a Fernando Ferreira Carneiro, Noemi Margarida Krefta e Cleber Adriano Rodrigues Folgado. Disponível em: <https://bit.ly/2Mjuhel>. Acesso em: 10 jan. 2018.

SCIENCE EUROPE. **Implementing Research Data Management Policies Across Europe – Experiences From Science Europe Member Organisations**. Mountain View: Science Europe, 2020. Disponível em: [https://www.scienceeurope.org/media/jikjlb2g/se\\_rdm\\_best\\_practices.pdf](https://www.scienceeurope.org/media/jikjlb2g/se_rdm_best_practices.pdf). Acesso em: 10 fev. 2020.

SILVA, Jonathas Luiz; FARIAS, Maria Giovanna Guedes. Abordagens conceituais e aplicativas da mediação nos serviços de informação. **InCID: Revista de Ciência da Informação e Documentação**, v. 8, n. 2, p. 106-123, 4 out. 2017. DOI: 10.11606/issn.2178-2075.v8i2p106-123.

SILVA, Jonathas Luiz Carvalho. Percepções conceituais sobre mediação da informação. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 6, n. 1, p. 93-108, mar./ago. 2015. Disponível em: <http://www.revistas.usp.br/incid/article/view/89731>. Acesso em: 10 jan. 2018.

UNIVERSIDADE FEDERAL DO CEARÁ. **Plano de Desenvolvimento Institucional - PDI UFC 2018-2022**. Fortaleza: UFC, [2018]. Disponível em: [http://www.ufc.br/images/files/a\\_universidade/plano\\_desenvolvimento\\_institucional/cartilha\\_pdi\\_2018\\_2022.pdf](http://www.ufc.br/images/files/a_universidade/plano_desenvolvimento_institucional/cartilha_pdi_2018_2022.pdf). Acesso em: 10 jan. 2018.

## APÊNDICE I – MINUTA DA PROPOSTA DE POLÍTICA DE GESTÃO DE DADOS DE PESQUISA DA UNIVERSIDADE FEDERAL DO CEARÁ

### 1 Disposições gerais

Este documento visa ser um dispositivo norteador e quadro geral abrangente para a Universidade Federal do Ceará (UFC) na orientação da aplicação de procedimentos específicos voltados para a pesquisa aberta e a gestão dos dados oriundos de pesquisa científica na instituição. Dispõe sobre as responsabilidades da universidade e da comunidade acadêmica interna referente à gestão e preservação dos dados científicos atuais ou futuros. Esta política aplica-se a todas as pesquisas conduzidas por servidores docentes, servidores técnico-administrativos em educação e discentes da UFC, independentemente de ser a pesquisa financiada ou não. Esta política tem a aspiração de ser implementada, entretanto, reconhece-se que poderá levar alguns anos até a sua efetiva implantação.

Diante da explosão de dados e informações hodiernamente, torna-se fator estratégico a adesão aos devidos padrões da administração e governança dos dados de pesquisa científica (pré-registro / registro, coleta, documentação, arquivamento, preservação, compartilhamento e publicação), tão necessário para acompanhar o cenário científico nacional e internacional, além de advogar cada vez mais pela adoção de processos de pesquisa sustentáveis, abertos (abertura quando possível, fechado quando necessário), verificáveis, reproduzíveis e dinâmicos.

Consideram-se dados de pesquisa todo e qualquer dado ou unidade de informação que são criados / produzidos no decurso da pesquisa, tais como planilhas, gravações de áudio e vídeo, transcrições, modelos, matrizes de dados, scripts, espécimes, coleções de imagens, fotografias, entre outros tipos de registros científicos que geram dados oriundos da pesquisa, seja em suporte físico ou digital.

Todas as propostas de pesquisa desenvolvidas na universidade devem incluir a elaboração de um Plano de Gestão de Dados, documento formal que explicita a coleta, captura, gerenciamento, integridade, confidencialidade, retenção, compartilhamento e publicação dos dados.

Em síntese, a gestão dos dados de pesquisa viabiliza:

- Envidar as boas práticas relacionadas à integridade e ética da pesquisa científica;
- Aplicar as boas práticas que assegurem a garantia dos argumentos e resultados científicos reproduzíveis a longo prazo;
- Ampliar o alcance, o reconhecimento e a divulgação das pesquisas desenvolvidas pela instituição no cenário nacional e internacional;
- Contribuir para a internacionalização e formação de parcerias educacionais;
- Contribuir ainda mais para a elevação da qualidade das pesquisas desenvolvidas pela universidade;
- Aderir à gestão responsável dos dados de pesquisa, desde o pré-registro / registro, coleta, documentação, arquivamento, preservação, compartilhamento e publicação, até o

armazenamento seguro de dados, cumprimento de normas e protocolos que não firam a ética e em respeito à lei geral de proteção de dados pessoais, além da proteção do capital intelectual e patentes desenvolvidos pelos pesquisadores da UFC;

- Atender aos requisitos das agências de fomento, financiadores e editores no que concerne ao compartilhamento e à gestão de dados de pesquisa.

A respeito dos direitos exclusivos de reutilização ou publicação dos dados de pesquisa científica desenvolvidas no âmbito da UFC, recomenda-se que não devem ser entregues para editoras ou agentes comerciais sem que esses dados tenham recebido o licenciamento adequado e disponibilizado abertamente para reutilização, exceto se esta for uma condição imposta pelo financiador da pesquisa ou da agência de fomento, ou se esbarrar em alguma condição ética, legal ou contratual.

Podem ser aplicados aos dados de pesquisa, mediante justificativa documentada no Plano de Gestão de Dados, período de embargo de até dois anos para a disponibilização e compartilhamento dos dados de pesquisa.

## 2 Papéis e responsabilidades

Sobre os papéis desempenhados e as responsabilidades esperadas concernentes à gestão de dados de pesquisa, estabelece-se:

- Biblioteca:
  - Desenvolver serviços de dados e apoio ao pesquisador;
  - Prestar consultoria, fornecer aconselhamento, orientação e ações de educação, treinamento, informação e congêneres à comunidade acadêmica da UFC no que concerne à gestão de dados de pesquisa e aspectos relacionados, a exemplo da Ciência Aberta;
  - Formar, capacitar, treinar e preparar a equipe de trabalho das bibliotecas que prestarão os serviços de dados, curadoria e de apoio ao pesquisador na UFC;
  - Intermediar soluções de armazenamento seguro para dados de pesquisa dinâmicos e estáticos em cooperação com a Secretaria de Tecnologia da Informação institucional (STI) e/ou buscar exploração de serviços de nuvem apropriados;
  - Firmar parcerias e redes de trabalho na própria instituição com outros setores e os cursos, departamentos, Pró-Reitoria de Pesquisa e Pós-Graduação e Pró-Reitoria de Graduação para promover e fortalecer o gerenciamento dos dados de pesquisa na UFC;
  - Formar e coordenar uma rede de administradores de dados (bibliotecários *liaison*, incorporados aos cursos ou campi / faculdade da universidade) em cada campus;
  - Realizar estudos com vistas a diagnosticar, mapear e sugerir melhorias para fornecer uma infraestrutura de dados na universidade.

- Secretaria de Tecnologia da Informação (STI):
  - Fornecer infraestrutura adequada, padronizada, certificada, robusta e de alta qualidade para facilitar o processo de gerenciamento dos dados de pesquisa;
  - Fornecer espaço de armazenamento exclusivamente para os dados de pesquisas desenvolvidas na UFC;
  - Fornecer o gerenciamento, controle de acesso aos dados e mecanismos de segurança da informação em conformidade com as diretrizes legais da área e do país.
  
- Programas de Pós-Graduação Stricto Sensu (Mestrado e Doutorado) e Programas de Graduação:
  - Solicitar a colaboração dos orientadores para supervisionar os seus alunos do mestrado e doutorado e assegurar o cumprimento da devida gestão dos dados de suas pesquisas desde o início ao final do projeto, principalmente com o preenchimento do Plano de Gestão de Dados (PGD);
  - Informar ao orientador e aluno de que devem garantir, que os dados de pesquisa serão repassados para a universidade antes de sua saída da instituição por motivo de conclusão do curso;
  - Informar que o aluno e o orientador são responsáveis por assegurar que os dados de terceiros, não serão repassados aos financiadores de pesquisa, agências de fomento ou qualquer outro sem a expressa autorização
  - Incentivar o seguimento das diretrizes e orientações fornecidas pelo serviço de dados e apoio ao pesquisador desempenhado pela biblioteca;
  - Sugerir a participação do corpo docente e discente em pelo menos uma capacitação e/ou treinamento presencial ou a distância sobre o gerenciamento dos dados de pesquisa oferecidos pela biblioteca.
  
- Docentes / Orientadores (Graduação, Mestrado e Doutorado):
  - Supervisionar os seus alunos do mestrado e doutorado e assegurar o cumprimento da devida gestão dos dados de suas pesquisas desde o início ao final do projeto, principalmente com o preenchimento do Plano de Gestão de Dados (PGD) e que este deve ser atualizado regularmente;
  - O orientador e aluno devem garantir que os dados de pesquisa serão repassados para a universidade antes de sua saída da instituição por motivo de conclusão do curso;
  - O aluno e o orientador são responsáveis por assegurar que os dados de terceiros não serão repassados aos financiadores de pesquisa, agências de fomento ou qualquer outro sem a expressa autorização;
  - Seguir e fazer com que sejam seguidas as diretrizes e orientações fornecidas pelo serviço de dados e apoio ao pesquisador desempenhado pela biblioteca;
  - Participar de pelo menos uma capacitação e/ou treinamento presencial ou a distância sobre o gerenciamento dos dados de pesquisa oferecidos pela biblioteca;

- Garantir que os alunos participem, em pelo menos uma capacitação e/ou treinamento presencial ou a distância sobre o gerenciamento dos dados de pesquisa oferecidos pela biblioteca;
  - Solicitar, a qualquer tempo, cursos, capacitações, treinamentos, consultoria e aconselhamento junto à biblioteca mediante agendamento;
  - Elaborar um Plano de Gestão de Dados para cada pesquisa ou projeto desenvolvido, pelo menos nos primeiros 12 meses de seu início e procedendo à sua atualização ao longo do processo até o seu término;
  - Garantir que os dados de pesquisa, códigos e quaisquer outros materiais necessários para os resultados de pesquisa sejam reproduzíveis (quando for o caso), que sigam os princípios FAIR (dados fáceis de encontrar, acessíveis, interoperáveis e reutilizáveis), documentados e compartilhados em um repositório de dados de livre escolha do pesquisador ao final do projeto, disponíveis por pelo menos 10 anos. Nos casos em que o compartilhamento não for possível, a justificativa deve constar expressamente no Plano de Gestão de Dados e ser informada no manuscrito final do trabalho;
  - Caso os dados não sejam disponibilizados em um repositório, deve ser assegurada a garantia de publicação dos metadados relevantes em um repositório de dados e em quaisquer publicações de pesquisa resultantes do projeto em que conste uma declaração expressa em que conste a explicação sobre o material existente e seus conjuntos de dados adicionais, inclusive deve informar porque o acesso é restrito, fornecer contato para possíveis pedidos de informação, informar quem poderá utilizar e em que circunstâncias.
- Discentes (Graduação, Mestrado e Doutorado):
    - Assegurar o cumprimento da devida gestão dos dados de suas pesquisas desde o início ao final do projeto, principalmente com o preenchimento do Plano de Gestão de Dados (PGD) e que este deve ser atualizado regularmente;
    - Deve garantir o repasse e entrega dos dados de pesquisa para a universidade antes de sua saída da instituição por motivo de conclusão do curso;
    - Responsabilizar-se e assegurar que os dados de terceiros não serão repassados aos financiadores de pesquisa, agências de fomento ou qualquer outro sem a expressa autorização;
    - Elaborar um Plano de Gestão de Dados para cada pesquisa ou projeto desenvolvido, pelo menos nos primeiros 12 meses de seu início e procedendo à sua atualização ao longo do processo até o seu término;
    - Garantir que os dados de pesquisa, códigos e quaisquer outros materiais necessários para os resultados de pesquisa sejam reproduzíveis (quando for o caso), que sigam os princípios FAIR (dados fáceis de encontrar, acessíveis, interoperáveis e reutilizáveis), documentados e compartilhados em um repositório de dados à livre escolha do pesquisador ao final do projeto, disponíveis por pelo menos 10 anos. Nos casos em

- que o compartilhamento não for possível, a justificativa deve constar expressamente no Plano de Gestão de Dados e ser informada no manuscrito final do trabalho;
- Caso os dados não sejam disponibilizados em um repositório, deve ser assegurada a garantia de publicação dos metadados relevantes em um repositório de dados e em quaisquer publicações de pesquisa resultantes do projeto em que conste uma declaração expressa em que conste a explicação sobre o material existente e seus conjuntos de dados adicionais, inclusive deve informar porque o acesso é restrito, além de fornecer contato para possíveis pedidos de informação, informar quem poderá utilizar e em que circunstâncias;
  - Seguir as diretrizes e orientações fornecidas pelo serviço de dados e apoio ao pesquisador desempenhado pela biblioteca;
  - Participar de pelo menos uma capacitação e/ou treinamento presencial ou a distância sobre o gerenciamento dos dados de pesquisa oferecidos pela biblioteca;
  - Solicitar, a qualquer tempo, cursos, capacitações, treinamentos, consultoria e aconselhamento junto à biblioteca e mediante agendamento.
- Responsabilidades gerais dos pesquisadores (docentes, técnico-administrativos e discentes):
    - O pesquisador irá escolher quais dados deseja compartilhar;
    - O pesquisador pode escolher o repositório de sua preferência para o depósito dos dados;
    - Não devem compartilhar dados sensíveis, registros de pacientes de qualquer natureza, ou todo o tipo de dado que possa causar algum dano a outrem, ou ainda que seja objeto de patente e afins;
    - Os pesquisadores devem respeitar as leis e aspectos éticos relacionados às pesquisas, aos seres humanos, objetos ou animais;
    - Os pesquisadores devem providenciar e garantir os meios necessários para que o armazenamento dos dados seja feito de forma segura;
    - Compete ao pesquisador assegurar-se de que estão sendo aplicados e cumpridos os termos de utilização e licenciamento para o compartilhamento dos dados;
    - Os interesses dos participantes de pesquisa devem ser assegurados, considerando as normas éticas da pesquisa, do Comitê de Ética em Pesquisa e a legislação vigente;
    - Compete aos pesquisadores publicar seus dados de pesquisa em formatos digitais apropriados (não-proprietários) sempre que possível a fim de facilitar o reuso dos dados;
    - Caso o pesquisador se desligue da universidade durante o período de realização da pesquisa ou após a sua conclusão, deverá transmitir a responsabilidade da gestão dos dados antes de sua saída expressa e documentada no Plano de Gestão de Dados. A referida transferência pode implicar em um acordo entre as partes para transferir essa responsabilidade para a outra instituição que o pesquisador está se filiando.

### 3 Relações desta política com outras normas, diretrizes e legislação vigente

- Lei de Acesso à Informação (Lei nº 12.527);
- Lei Geral de Proteção de Dados (Lei nº 13.709/2018);
- Resolução CNS nº 466/2012;
- Resolução CNS nº 510/2016.

### 4 Definições

Dados de Pesquisa: Todo e qualquer dado ou unidade de informação que são criados / produzidos no decurso da pesquisa, tais como planilhas, gravações de áudio e vídeo, transcrições, modelos, matrizes de dados, scripts, espécimes, coleções de imagens, fotografias, entre outros tipos de registros científicos que geram dados oriundos da pesquisa.

Dado pessoal: Informação relacionada a pessoa natural identificada ou identificável.

Metadados: Conjunto de informações detalhadas e autoexplicativas sobre os conjuntos de dados. Apresenta características dos dados, informações de licenças, dados de autoria, tipologia dos arquivos, mediante padrões internacionalmente aceitos e que permitem conhecer a origem dos dados. Os metadados são responsáveis por auxiliar na localização de cada conjunto de dados através dos mecanismos de busca em um repositório.

Plano de Gestão de Dados: Documento formal que explicita a coleta, captura, gerenciamento, integridade, confidencialidade, retenção, compartilhamento e publicação dos dados.

Princípios FAIR (Findable, Accessible, Interoperable, Reusable): Conjunto de princípios orientadores para tornar os dados fáceis de encontrar, acessíveis, interoperáveis e reutilizáveis. A entidade FORCE11 é a organização responsável pela criação e definição desses princípios.

## ANEXO A – PARECER CONSUBSTANCIADO DO COMITÊ DE ÉTICA EM PESQUISA

UFC - UNIVERSIDADE  
FEDERAL DO CEARÁ /



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Ciência aberta, compartilhamento e gestão de dados científicos: percepção dos pesquisadores da Universidade Federal do Ceará

**Pesquisador:** Juliana Soares Lima

**Área Temática:**

**Versão:** 1

**CAAE:** 18708719.0.0000.5054

**Instituição Proponente:** UNIVERSIDADE FEDERAL DO CEARÁ

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

**Número do Parecer:** 3.512.152

#### Apresentação do Projeto:

Este estudo busca investigar a percepção dos pesquisadores (professores e alunos de Pós-graduação Stricto Sensu) da Universidade Federal do Ceará a respeito dos princípios da Ciência Aberta e da Gestão de dados de pesquisa. Além disso, a pesquisa irá identificar quais as necessidades informacionais destes usuários em relação às temáticas a fim de realizar um diagnóstico institucional e propor diretrizes, a criação de serviços em gestão de dados de pesquisa e o estabelecimento de um Programa de Apoio ao Pesquisador. A pesquisa tem abordagem mista (qualitativa e quantitativa) e adotou como método norteador a Teoria Fundamentada em Dados

#### Objetivo da Pesquisa:

Investigar as práticas e necessidades informacionais concernentes à ciência aberta e ao gerenciamento de dados científicos, a fim de propor um programa de suporte e apoio ao pesquisador na Universidade Federal do Ceará (UFC).

#### Avaliação dos Riscos e Benefícios:

**Riscos:**

Durante a entrevista serão tomados os devidos cuidados a fim de não causar qualquer constrangimento aos entrevistados. Os dados serão anonimizados e codificados a fim de manter a integridade e o sigilo da pesquisa. Os questionários aplicados não coletarão nenhum dado que

**Endereço:** Rua Cel. Nunes de Melo, 1000

**Bairro:** Rodolfo Teófilo

**CEP:** 60.430-275

**UF:** CE

**Município:** FORTALEZA

**Telefone:** (85)3366-8344

**E-mail:** comepe@ufc.br



UFC - UNIVERSIDADE  
FEDERAL DO CEARÁ /



Continuação do Parecer: 3.512.152

identifique os participantes e garantindo o sigilo do respondente. Os registros observacionais e afins constantes no ambiente virtual de aprendizagem serão anonimizados e codificados de forma a garantir o sigilo dos participantes.

**Benefícios:**

Este estudo contribuirá para a proposta de criação de novos serviços em gestão de dados científicos e do Programa de Apoio e Suporte ao Pesquisador.

**Comentários e Considerações sobre a Pesquisa:**

O Projeto está bem escrito e segue o rigor da ética em pesquisa científica.

**Considerações sobre os Termos de apresentação obrigatória:**

Os termos obrigatório estão todos presentes conforme a resolução 466/12

**Recomendações:**

Aprovação

**Conclusões ou Pendências e Lista de Inadequações:**

Aprovado

**Considerações Finais a critério do CEP:**

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1361494.pdf	18/06/2019 23:44:14		Aceito
Outros	CVLattes.pdf	18/06/2019 23:43:45	Juliana Soares Lima	Aceito
Folha de Rosto	FolhadeRostoassinada.pdf	18/06/2019 23:29:29	Juliana Soares Lima	Aceito
Declaração de Instituição e Infraestrutura	Autorizacao_institucional.pdf	18/06/2019 23:28:59	Juliana Soares Lima	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_Dissertacao.pdf	18/06/2019 11:37:04	Juliana Soares Lima	Aceito
Outros	Termo_de_compromisso_dados.pdf	18/06/2019 11:22:56	Juliana Soares Lima	Aceito
Outros	Carta_de_solicitacao_de_apreciacao_CEP.pdf	18/06/2019 11:22:01	Juliana Soares Lima	Aceito
Declaração de Pesquisadores	Declaracao_de_concordancia.pdf	18/06/2019 11:20:25	Juliana Soares Lima	Aceito

**Endereço:** Rua Cel. Nunes de Melo, 1000

**Bairro:** Rodolfo Teófilo

**CEP:** 60.430-275

**UF:** CE

**Município:** FORTALEZA

**Telefone:** (85)3366-8344

**E-mail:** comepe@ufc.br

UFC - UNIVERSIDADE  
FEDERAL DO CEARÁ /



Continuação do Parecer: 3.512.152

Cronograma	Cronograma.pdf	18/06/2019 11:19:20	Juliana Soares Lima	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TERMO_DE_CONSENTIMENTO_LIVR E_ESCLARECIDO_pesquisaCI.pdf	18/06/2019 11:19:02	Juliana Soares Lima	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

FORTALEZA, 16 de Agosto de 2019

---

Assinado por:  
**FERNANDO ANTONIO FROTA BEZERRA**  
(Coordenador(a))

**Endereço:** Rua Cel. Nunes de Melo, 1000  
**Bairro:** Rodolfo Teófilo **CEP:** 60.430-275  
**UF:** CE **Município:** FORTALEZA  
**Telefone:** (85)3366-8344 **E-mail:** comepe@ufc.br