

**UNIVERSIDADE FEDERAL DA PARAÍBA  
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS  
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO  
BACHARELADO EM BIBLIOTECONOMIA**

**RENATA LEMOS DOS ANJOS**

**DESVENDANDO AS PERCEPÇÕES E PRÁTICAS ADOTADAS PELOS  
PESQUISADORES DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM  
ECOLOGIA DO NORDESTE BRASILEIRO NA GESTÃO DOS DADOS  
CIENTÍFICOS**

**João Pessoa  
2016**

**RENATA LEMOS DOS ANJOS**

**DESVENDANDO AS PERCEPÇÕES E PRÁTICAS ADOTADAS PELOS  
PESQUISADORES DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM  
ECOLOGIA DO NORDESTE BRASILEIRO NA GESTÃO DOS DADOS  
CIENTÍFICOS**

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciência da Informação do Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba, para a obtenção do grau de Bacharel em Biblioteconomia.

**Orientador: GUILHERME ATAÍDE DIAS**

**João Pessoa  
2016**

A599d Anjos, Renata Lemos dos.  
Desvendando as percepções e práticas adotadas pelos  
pesquisadores dos programas de pós-graduação em Ecologia do  
Nordeste brasileiro na gestão dos dados científicos / Renata Lemos dos  
Anjos. – João Pessoa, 2016.  
63f. : il.

Orientador: Prof. Dr. Guilherme Ataíde Dias.  
Trabalho de Conclusão de Curso (Graduação em  
Biblioteconomia) – UFPB/CCSA.

1. Produção científica. 2. Dados científicos. 3. Gestão de dados.  
4.e-Science. I. Título.

UFPB/CCSA/BS

CDU: 02(043.2)

**RENATA LEMOS DOS ANJOS**

**DESVENDANDO AS PERCEPÇÕES E PRÁTICAS ADOTADAS PELOS  
PESQUISADORES DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM  
ECOLOGIA DO NORDESTE BRASILEIRO NA GESTÃO DOS DADOS  
CIENTÍFICOS**

Trabalho de Conclusão de Curso  
apresentado ao Departamento de  
Ciência da Informação do Centro de  
Ciências Sociais Aplicadas da  
Universidade Federal da Paraíba, para a  
obtenção do grau de Bacharel em  
Biblioteconomia.

Aprovado em 30 de Novembro de 2016.

---

Guilherme Ataíde Dias

Universidade Federal da Paraíba / Departamento de Ciência da Informação

---

Marckson Roberto Ferreira de Sousa

Universidade Federal da Paraíba / Departamento de Ciência da Informação

---

André Luiz Dias de França

Universidade Federal da Paraíba / Departamento de Comunicação Social

## RESUMO

Baseado na premissa de que os dados oriundos de investigações científicas costumam ficar armazenados nos computadores pessoais dos pesquisadores, ou são descartados e perdidos após a publicação dos resultados em artigos, impossibilitando a produção de novas investigações, este estudo buscou desvendar as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro na gestão dos dados científicos, baseando-se no fenômeno emergente nacional e internacionalmente, *e-Science*, que transformou a forma como os dados de pesquisa podem ser produzidos, disseminados, gerenciados, compartilhados e (re) utilizados a partir do processo de curadoria dos dados. Diante do exposto, indagou-se quais as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro na gestão de dados científicos. Trata-se de uma pesquisa exploratória, visando a ampliação do conhecimento acerca do tema, e de abordagem qualitativa e quantitativa para análise dos dados que foram coletados por meio da aplicação de um questionário. Concluímos que a realidade do compartilhamento de dados por meio do processo de curadoria dos mesmos é emergente, visto que alguns dos sujeitos desta pesquisa já utilizam ferramentas que auxiliam no plano de gestão de dados, e outros, tem imenso interesse em aumentar seus conhecimentos, em relação às práticas de gestão de dados científicos para posterior compartilhamento.

**Palavras-chave:** *e-Science*. Dados científicos. Gestão de dados.

## ABSTRACT

Based on the premise that data from scientific investigations are usually stored on the researchers' computers, discarded or even lost when the results are published in articles, making it impossible to produce new investigations, the present study sought to uncover the perceptions and practices adopted by the researchers from the postgraduation programs of the Brazilian Northeast ecology in the management of scientific data, based on the emerging national and international phenomenon, e-Science, that transformed the way research data can be produced, disseminated, managed, shared and reused taking start at the data curation process. In the view of the above, the researchers from the postgraduation programs of the Brazilian Northeast ecology were questioned about the perceptions and practices adopted in the management of scientific data. This study is an exploratory research, that aims to increase the knowledge about the subject, and a qualitative and quantitative approach to analyze the data that were collected through the application of a questionnaire. It was concluded that the reality of data sharing through the process of curatorship is emerging, since some of the subjects of this research already use tools that help with the data management plan, and others have a great interest in increasing their knowledge about the practices of management of scientific data for later sharing.

**Keywords:** e-Science. Scientific data. Data management.

## LISTA DE FIGURAS

<b>FIGURA 1</b> – O ciclo de vida de dados DataONE .....	20
<b>FIGURA 2</b> – Ciclo de vida de dados para Ciência da Informação .....	23
<b>FIGURA 3</b> – Habilidades do bibliotecário no plano de gestão de dados.....	26
<b>FIGURA 4</b> – Dados quantitativos de programas recomendados e reconhecidos .....	28
<b>FIGURA 5</b> – Envio do questionário aos pesquisadores.....	30
<b>FIGURA 6</b> – Análise manual dos dados .....	30
<b>FIGURA 7</b> – Origem dos dados .....	37

## LISTA DE QUADROS

<b>QUADRO 1</b> – Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro .....	29
<b>QUADRO 2</b> – Por que não faz a gestão dos dados científicos? .....	34
<b>QUADRO 3</b> – Por que não desenvolve plano de gestão de dados ao longo da condução das pesquisas?.....	38
<b>QUADRO 4</b> – Ferramentas utilizadas na criação do plano de gestão de dados.....	40
<b>QUADRO 5</b> – Padrões de metadados utilizados .....	40
<b>QUADRO 6</b> – Ferramentas de <i>software</i> utilizadas para exploração, visualização ou análise de dados .....	42
<b>QUADRO 7</b> – Outras razões para o não compartilhamento de dados.....	46
<b>QUADRO 8</b> – Período de tempo razoável para embargo aos dados de pesquisa.....	49
<b>QUADRO 9</b> – Comentários pertinentes à temática da pesquisa e para a pesquisa .....	49



## LISTA DE GRÁFICOS

<b>GRÁFICO 1</b> - Participantes do questionário .....	30
<b>GRÁFICO 2</b> - Questionários completos .....	32
<b>GRÁFICO 3</b> - Gostaria de responder alguns questionamentos pertinentes a temática introduzida.....	33
<b>GRÁFICO 4</b> - Faixa etária dos sujeitos da pesquisa .....	32
<b>GRÁFICO 5</b> - Quanto tempo os sujeitos atuam na Pós-Graduação.....	34
<b>GRÁFICO 6</b> - Faz a gestão dos dados científicos coletados nas suas pesquisas .....	35
<b>GRÁFICO 7</b> - Natureza de análise atribuída aos dados de pesquisa.....	36
<b>GRÁFICO 8</b> - Desenvolve plano de gestão de dados.....	37
<b>GRÁFICO 9</b> - Uso de ferramenta automatizada para auxílio no plano de gestão de dados...	38
<b>GRÁFICO 10</b> - Nível de concordância com declarações acerca de planos de gestão de dados .....	39
<b>GRÁFICO 11</b> - Uso de metadados .....	40
<b>GRÁFICO 12</b> - Uso de ferramenta de <i>software</i> para exploração, visualização ou análise de dados .....	42
<b>GRÁFICO 13</b> - Nível de concordância com declarações acerca do compartilhamento de dados .....	43
<b>GRÁFICO 14</b> - Nível de concordância com declarações referente ao compartilhamento de dados .....	45
<b>GRÁFICO 15</b> - Razões para o não compartilhamento de dados.....	46
<b>GRÁFICO 16</b> - Locais de armazenamento dos dados coletados .....	47
<b>GRÁFICO 17</b> - Pesquisadores devem definir um período de embargo sobre seus dados?....	47
<b>GRÁFICO 18</b> - Período de tempo razoável para embargo aos dados de pesquisa .....	48

## LISTA DE SIGLAS

<b>CAPES</b>	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<b>CVD-CI</b>	Ciclo de Vida de Dados para Ciência da Informação
<b>IBICT</b>	Instituto Brasileiro de Informação em Ciência e Tecnologia
<b>INPA</b>	Instituto Nacional de Pesquisa da Amazônia
<b>NSB</b>	<i>National Science Board</i>
<b>NSF</b>	<i>National Science Foundation</i>
<b>OECD</b>	<i>Organisation for Economic Co-Operation and Development</i>
<b>PGD</b>	Plano de gestão de dados
<b>PPBio</b>	Programa de Pesquisa em Biodiversidade
<b>TDICs</b>	Tecnologias Digitais de Informação e Comunicação
<b>USGS</b>	<i>United States Geological Survey</i>

# SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>11</b>
<b>2 OBJETIVOS .....</b>	<b>16</b>
<b>3 A EMERGÊNCIA DOS DADOS DE PESQUISA .....</b>	<b>17</b>
3.1 DADOS .....	17
3.2 CURADORIA DE DADOS .....	19
3.3 REFLEXÃO ACERCA DA CIÊNCIA ORIENTADA A DADOS.....	24
<b>4 O AMBIENTE DE <i>E-SCIENCE</i> E O PROFISSIONAL BIBLIOTECÁRIO .....</b>	<b>26</b>
<b>5 PROCEDIMENTOS METODOLÓGICOS.....</b>	<b>28</b>
<b>6 APRESENTAÇÃO E ANÁLISE DOS DADOS.....</b>	<b>32</b>
<b>7 CONSIDERAÇÕES FINAIS .....</b>	<b>51</b>
<b>REFERÊNCIAS.....</b>	<b>53</b>
<b>APÊNDICES .....</b>	<b>566</b>

## 1 INTRODUÇÃO

Desde a época em que foram publicados os primeiros periódicos científicos, a saber, *Philosophical Transactions of the Royal Society* na Inglaterra, e *Journal des Sçavans* na França, ambos em 1665, a comunicação científica permanece nos mesmos *modus operandi*, em que, a forma de se pensar, fazer, e divulgar o conhecimento continua tradicionalmente direcionada à publicação em periódicos científicos, a priori em periódicos impressos, mais recentemente, a partir do final do século XX, em periódicos científicos eletrônicos.

A maioria das investigações científicas encerra o seu ciclo a partir da publicação dos resultados obtidos, em artigos científicos. A partir da publicação do conhecimento produzido novos estudos podem ser gerados de forma a beneficiar toda a comunidade.

Os resultados de pesquisa são produto da análise dos dados oriundos dos processos de investigação, dados estes que tradicionalmente não são publicados, e demonstram ser uma fonte primária de recursos científicos que podem ser utilizados e reutilizados por outros pesquisadores na investigação de outros fenômenos e até para replicar um determinado experimento com fins de atestar a sua validade. Os dados oriundos de investigações científicas costumam ser descartados ou perdidos após a conclusão da pesquisa e da consequente publicação dos resultados em artigos científicos produtos do trabalho de investigação, como resultado nos deparamos com uma infinidade de dados de pesquisas, perdidos na rede, ou armazenados nos computadores pessoais dos pesquisadores, o que impede a produção de novas investigações. Visando a preservação destes dados e a origem de novas investigações, muitos pesquisadores de todas as áreas do conhecimento, em todas as regiões do globo estão dando especial atenção para o processo de curadoria dos dados científicos, oriundos de suas pesquisas. Ao invés de um simples descarte, estes dados estão sendo coletados, gerenciados, preservados e compartilhados através de repositórios especificamente construídos para este fim (BORGMAN, 2015).

Com o avanço das Tecnologias Digitais da Informação e Comunicação (TDICs), ocorreram mudanças na forma de se fazer e divulgar a ciência, já que as mesmas possibilitam um mundo conectado e interligado. As dinâmicas associadas ao processo de geração do conhecimento acadêmico e científico utilizam o escopo da comunicação científica formal como um dos meios para a disseminação, uso e acesso à pesquisa científica. Este escopo está se reconfigurando, abrindo novos espaços de interação e conectividade mediados pelas TDICs.

No contexto atual, novas possibilidades relacionadas ao uso dos dados de pesquisas estão emergindo (BORGMAN, 2015). A convergência das tecnologias digitais, e do armazenamento *online* por meio das redes de alto desempenho, transformou a forma como os dados de pesquisa podem ser produzidos, disseminados, gerenciados, compartilhados e usados.

Gray (2005 apud Sales et al, 2013) introduz o conceito de *e-Science*, onde a “ciência é produzida a partir do uso, processamento, análise e compartilhamento de dados de pesquisa”, ou seja, o processamento computacional intensivo e a curadoria dos dados de pesquisa é um fenômeno conhecido como o quarto paradigma científico, ou *e-Science*.

Sayão e Sales (2014, p.79), afirmam que:

Esse fenômeno contemporâneo cria oportunidades sem precedentes para acelerar a pesquisa científica e gerar riquezas com base na exploração desse acúmulo de dados; abre a possibilidade de que a imensidão de dados gerados pela pesquisa científica contemporânea possa ser coletada, comparada e analisada, engendrando novos conhecimentos e novas questões de pesquisas. Ferramentas avançadas de software e de mineração de dados ajudam a interpretar e transformar os dados brutos em configurações ilimitadas de informação e conhecimento.

De acordo com a citação apresentada, a *e-Science*, faz a utilização de dados brutos através das TDICs, fazendo com que a comunicação científica tenha uma nova forma de fazer ciência, uma ciência orientada por dados, voltada para o processamento dos mesmos em larga escala, compartilhados globalmente, por meio da socialização e colaboração (SALES; SAYÃO, 2015).

Ainda abrangendo o novo fenômeno da comunicação científica, Gordon (2009, *online*, tradução nossa) afirma que,

Novo quarto paradigma da ciência é baseado em dados de computação intensiva. Em tal pesquisa científica, nós estamos em um estágio de desenvolvimento que é análogo a quando a imprensa foi inventada. A impressão levou milhares de anos para se desenvolver e evoluir para as muitas formas que assume hoje<sup>1</sup>.

No âmbito da Ciência da Informação, a *e-Science* possibilita que façamos alguns questionamentos. Como, por exemplo, qual o papel do profissional da informação no tratamento dos dados primários oriundos de investigações científicas. De acordo com Costa e Cunha (2014, p. 194) “os dados produzidos pela *e-Science* trazem impactos profundos sobre a ciência e, portanto, exige um exame acurado das funções das instituições empenhadas no

---

<sup>1</sup> New fourth paradigm for science based on data intensive computing. In such scientific research, we are at a stage of development that is analogous to when the printing press was invented. Printing took a thousand years to develop and evolve into the many forms it takes today.

avanço da ciência e no apoio aos cientistas, dentre elas as bibliotecas”. As bibliotecas precisam apoiar a *e-Science*, e para isso, é necessário que elas incorporem uma série de práticas e serviços além dos que já realizam, como também, devem fazer investimentos na segurança, preservação, gestão, acesso aos dados, controle de metadados, além de treinamentos dos profissionais da informação, para assumirem satisfatoriamente as funções inerentes à gestão de dados (CUNHA, 2010, p. 12).

Os bibliotecários podem auxiliar os pesquisadores em seus processos de pesquisa, em um amplo escopo, oferecendo apoio aos mesmos durante as fases do ciclo de vida dos dados científicos, por meio de um Plano de Gestão de Dados (PGD), em que estão explicitadas ações para documentar, organizar e armazenar os dados de forma segura, tornando a assistência do bibliotecário ao pesquisador, fundamental, visto que podem auxiliar na organização dos variados tipos de dados, e na tomada de decisão sobre o acesso, e sobre a preservação de dados para posteriores projetos (CORRÊA, 2016), ou seja, os bibliotecários possuem habilidades que podem ser empregadas em todos os estágios do Plano de Gestão de Dados, além de oferecer ampla colaboração aos pesquisadores neste sentido.

Corrêa (2016, p. 405) explica ainda que:

(...) o bibliotecário possui o perfil necessário para auxiliar pesquisadores no desenvolvimento de um projeto de gestão de dados e indicar as ferramentas necessárias para este procedimento. Além disso, é imprescindível que ocorra a seleção adequada de um repositório conforme os requisitos de determinados conjuntos de dados, processo no qual ambos, bibliotecários e pesquisadores podem trabalhar juntos.

O novo cenário da comunicação científica mostra-se oportuno para os bibliotecários, já que dentre inúmeras capacidades, ressaltamos que os mesmos têm habilidades em lidar com o processo de gestão de dados científicos, bem como, realizar pesquisas em repositórios de dados e promover um ‘diálogo’ entre os pesquisadores. De acordo com Sales e Sayão (2015, p. 43) esclarecem que os bibliotecários podem participar na pré e pós-publicação das pesquisas científicas,

Através da colaboração próxima com os grupos de pesquisa, iniciada nas primeiras fases do ciclo de vida dos projetos, os bibliotecários podem se envolver como cocriadores e a biblioteca como um laboratório. Isto pode ser concretizado por meio do estabelecimento, teste e difusão de padrões, tecnologias e boas práticas na garantia da qualidade na geração dos dados; no desenvolvimento de repositórios dinâmicos que apresentem modelos de dados flexíveis capazes de apoiar o fluxo de trabalho da fase de pré-publicação, por meio de ambientes interativos de submissão e de colaboração, que possam também ser estendidos para apoio aos estágios pós-publicação, como integração, análise e visualização.

Neste contexto, a comunicação científica se encontra em desenvolvimento no que tange ao cenário científico brasileiro. A iniciativa de preservar e reutilizar dados científicos não está associada apenas a ação de pesquisadores isolados, mas a um empreendimento que é fomentado por instituições de pesquisas de vários países.

No Brasil, por exemplo, o Instituto Nacional de Pesquisa da Amazônia (INPA), mantém o Programa de Pesquisa em Biodiversidade (PPBio), que disponibiliza um Repositório de Dados de Estudos Ecológicos criado com “o objetivo de intensificar estudos sobre a biodiversidade no Brasil, descentralizar a produção científica dos centros desenvolvidos academicamente, integrar atividades de pesquisa e divulgar os resultados” (PPBio, *online*). Ainda como exemplo, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) mantém um repositório digital de dados de projetos de pesquisa, que foi implementado através do *software* Dataverse, que realiza o compartilhamento, preservação e arquivamento desses dados para todo o mundo, no qual os dados desta pesquisa estarão disponíveis. Dando maior ênfase a este fenômeno, o IBICT, lançou recentemente o Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã, que tem como objetivo:

[...] demonstrar o seu valor estratégico e informacional, estimular e apoiar movimentos e iniciativas para Ciência Aberta no Brasil, traduzidos pelo amplo e irrestrito acesso a fontes primárias de pesquisa utilizadas por pesquisadores e outros segmentos sociais, possibilitando o compartilhamento, reprodutibilidade, verificação, avaliação, reutilização e redistribuição em novos contextos e em pesquisas colaborativas e interdisciplinares. (IBICT, *online*, 2016).

Como visto em ambos os exemplos, os repositórios de dados conforme previamente mencionados, abrangem qualquer tipo de dado, de todas as áreas do conhecimento. Notadamente, a grande área das Ciências Exatas da Natureza, se configura, a partir das suas investigações científicas, como uma área geradora de muitos dados, já que analisam fenômenos distintos e realizam muitos experimentos baseados em dados a fim de atestar sua validade.

Para realização da pesquisa, delimitamos a área do conhecimento a ser pesquisada, de acordo com a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), em que selecionamos a grande área de avaliação Biodiversidade, que abrange 04 áreas específicas (Botânica, Ecologia, Oceanografia e Zoologia) das quais, selecionamos a área de Ecologia, por está sempre em desenvolvimento, o que a faz uma área necessitada do uso e preservação

dos dados, já que os mesmos se configuram como uma fonte primária, podendo ser utilizados e reutilizados em pesquisas posteriores, a fim de compreender fenômenos e certificá-los.

As reflexões emergem das tendências contemporâneas desse paradigma científico do *e-Science* e com diferentes perspectivas de interação, compartilhamento de dados de pesquisa e colaboração entre os pesquisadores, nos faz indagar: **Quais as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro na gestão dos dados científicos?**

O estudo está estruturado em seis capítulos, onde o seguinte, Capítulo II, serão apresentados os objetivos gerais e específicos do estudo. No Capítulo III, o estudo irá tratar sobre a emergência dos dados de pesquisa, conceituando os tipos de dados e tratando sobre a curadoria de dados. Subsequente, no Capítulo IV, o estudo explanará sobre o ambiente de *e-Science*. Nos Capítulos V e VI, serão expostos os procedimentos metodológicos realizados na pesquisa, bem como, a apresentação e análise dos dados que foram coletados, respectivamente. E por fim, no Capítulo VII, as conclusões serão apresentadas.

De acordo com o exposto, a presente pesquisa abrange uma temática emergente no que tange à comunicação científica internacional e nacional. A investigação possibilitará identificar de que forma a gestão de dados científicos está se configurando no Brasil, notadamente na grande área do conhecimento Biodiversidade, com ênfase em Ecologia. Além de possibilitar um panorama de como os pesquisadores científicos brasileiros, tomando como amostra os pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste, realizam a gestão de seus dados científicos.

A realização da pesquisa foi motivada por interesses pessoais no assunto, a fim de obter mais informação acerca do quarto paradigma científico, *e-Science*, além de saber e constatar como este fenômeno está se configurando, nesta população específica do cenário científico brasileiro. Voltada também, à relevância do tema no que diz respeito ao exercício da função dos bibliotecários, e como o mesmo pode influenciar e alterar as práticas e os serviços dos mesmos.



## 2 OBJETIVOS

No que tange ao objetivo geral, esta pesquisa busca **desvendar as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro na gestão de dados científicos**. Desta forma, para atender ao objetivo geral proposto neste estudo, delineou-se os seguintes objetivos específicos:

- a) Traçar o perfil dos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro;
- b) Identificar as ferramentas adotadas na realização do plano de gestão de dados científicos de pesquisa pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro;
- c) Averiguar a natureza e origem dos dados científicos coletados pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro;
- d) Verificar se os pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro fazem uso de Metadados;
- e) Identificar as práticas de compartilhamento de dados empregadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro;
- f) Identificar as ações tomadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro, no que diz respeito ao armazenamento e preservação (curadoria) dos dados oriundos de suas investigações científicas;

### 3 A EMERGÊNCIA DOS DADOS DE PESQUISA

Conforme o exposto, os dados trazem uma inovação na forma de se fazer ciência. Dados podem assumir diversos tipos, por assumirem diferentes características no que diz respeito a sua natureza e origem, como apresentaremos a seguir.

#### 3.1 DADOS

Existem várias definições para dados, visto que estes entendimentos podem variar notadamente entre pesquisadores e as suas áreas do conhecimento, já que são gerados para propósitos específicos, por diferentes processos e diferentes comunidades acadêmicas e científicas.

Dentre dezenas de possibilidades para o estabelecimento de um conceito de dados, concordamos com o pensamento de Setzer (2015, *online*), que define dado como “uma sequência de símbolos quantificados ou quantificáveis”. De acordo com Borgman (2015, p. 42, tradução nossa) “o resumo mais abrangente é dizer que os dados são representações de observações, objetos ou outras entidades utilizadas como evidência de fenômenos para fins de investigação ou atividades acadêmicas<sup>2</sup>”. Dados são oriundos de análises durante as investigações científicas, portanto, dados podem ser números, imagens, textos, vídeos, áudio, software, algoritmos, equações, animações, modelos, simulações (SAYÃO; SALES, 2014). Ou seja, no nosso contexto, podemos assumir que dados representam aquilo que foi extraído a partir das investigações científicas.

Em virtude da heterogeneidade dos dados faz-se necessário criar políticas que identifiquem e sustentem os vários tipos de dados e sua natureza (SAYÃO; SALES, 2014). Para o presente estudo, entendemos que é importante abordar a definição dos dados pessoais, bem como, a definição e a natureza dos dados de pesquisa, de uma forma ampla.

Dado pessoal, de acordo com o anteprojeto de lei para a proteção de dados pessoais, que dispõe sobre o tratamento de dados pessoais para proteger a personalidade e a dignidade da pessoa natural, considera-se:

Art. 5º Para os fins desta Lei considera-se:

I - dado pessoal: dado relacionado à pessoa natural identificada ou identificável, inclusive números identificativos, dados locais ou identificadores eletrônicos quando estes estiverem relacionados a uma pessoa. (Ministério da Justiça, *online*).

---

<sup>2</sup> The most inclusive summary is to say that data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

Os dados pessoais são de importância fundamental quando consideramos o mundo dos negócios. Os dados pessoais oferecem aos negócios a visão do perfil dos seus potenciais clientes, para que seus produtos possam atender de maneira satisfatória a necessidade dos mesmos, e conseqüentemente aumentar a procura por eles.

Dados de pesquisa são aqueles dados oriundos de investigações científicas, aceitos na comunidade científica e necessários para validar os resultados das investigações (DUDZIAK, *online*). Segundo o DataONE (*online*, p. 2, Tradução nossa), “dados de pesquisa são produtos valiosos do empreendimento científico<sup>3</sup>”. De acordo com a definição da *Organisation for Economic Co-Operation and Development* (OECD, 2007, p.13), dados de pesquisa, são:

(...) registros factuais (scores numéricos, documentos textuais, imagens e sons) usados como fonte primária para a pesquisa científica e que são comumente aceitos pelos pesquisadores como necessários para validar os resultados do trabalho científico<sup>4</sup>. (Tradução nossa)

Dados de pesquisa podem assumir diversas tipologias de acordo com a forma de como são obtidos, conforme a *National Science Board* (NSB) da *National Science Foundation* (NSF), os dados de pesquisa podem ser distinguidos pela sua natureza ou sua origem em: dados observacionais, computacionais e experimentais (NSB, 2005, *online*).

Os dados de pesquisa observacionais são aqueles obtidos por meio de observações associadas a lugares e tempos específicos, configurando-se como registros históricos não podendo ser coletados uma segunda vez. Os dados de pesquisa computacionais são frutos da execução de modelos computacionais ou de simulações. Por sua vez, os dados de pesquisa experimentais, são oriundos de situações controladas, por meio de experimentos (SALES; SAYÃO, 2013).

Não importa o tipo de dado de pesquisa, a preservação faz-se crucial. No primeiro tipo, devido ao fato de que não podem ser coletados uma segunda vez. Por sua vez, os dados computacionais, demandam um grande número de informações e dependências técnicas para reproduzi-los. E os dados experimentais, visto que nem sempre as condições experimentais serão idênticas. A preservação é indispensável à comunicação científica, visto que, segundo Poliakoff apud Agência FAPESP (2013, *online*):

Os dados que coletamos hoje podem ser usados no futuro de forma que ainda não conseguimos imaginar. Os exploradores de antigamente que coletavam

<sup>3</sup> Research data are valuable products of the scientific enterprise (...).

<sup>4</sup> (...) “research data” are defined as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings.

espécimes de plantas e animais não sabiam nada sobre DNA e hoje as amostras são submetidas a esse tipo de investigação. Quando você coleta os seus dados, reúne informações que, no futuro, poderão ser analisadas de formas muito diferentes. São coisas que terão um valor enorme para cientistas que ainda nem nasceram.

Dados de pesquisa são considerados o alicerce do conhecimento, precisando ser estruturados, organizados, descritos e disponibilizados para atuais e futuras compreensões (DUDZIAK, *online*). Assim, ressalta-se a importância da prática da curadoria de dados. Que será explanada a seguir.

### 3.2 CURADORIA DE DADOS

A curadoria de dados envolve a gestão atuante e a preservação dos dados de pesquisa durante todo o seu ciclo de vida, visando o atendimento a gerações atuais e futuras de usuários. Em linhas gerais, o processo de curadoria de dados “envolve a gestão de dados de pesquisa desde o seu planejamento, assegurando a sua preservação por longo prazo, descoberta, interpretação e reuso” (SAYÃO; SALES, 2012, p. 179).

O gerenciamento de dados possibilita que os dados sejam organizados, protegidos, acessíveis, e passíveis de reprodutibilidade. Segundo Harvey (2010 apud COSTA; CUNHA, 2014, p. 202), curadoria diz respeito à:

(...) gestão ativa de dados durante o tempo em que ele continua a ser: acadêmico, científico, pesquisável, administrável e/ou de interesse pessoal, com objetivos de apoiar sua reprodutibilidade, reutilização e agregando valor a esses dados. Gerenciando-os do momento de sua criação, até que eles sejam determinados como não úteis e garantindo a sua acessibilidade em longo prazo, assim como a sua preservação, autenticidade e integridade.

Esta prática envolve todo o processo de gerenciamento de dados científicos primários, permitindo a preservação e manutenção desses dados, tornando possível o compartilhamento e reuso dos mesmos (MEDEIROS; CAREGNATO, 2012).

De acordo com Sayão e Sales (2012, p. 185), a curadoria de dados:

(...) assegura a sustentabilidade dos dados para o futuro, não deixando, entretanto, de conferir valor imediato a eles para os seus criadores e para os seus usuários. Os recursos estratégicos, metodológicos e as tecnologias envolvidas nas práticas da curadoria digital facilitam o acesso persistente a dados digitais confiáveis por meio da melhoria da qualidade desses dados, do seu contexto de pesquisa e da checagem de autenticidade. Dessa forma, a curadoria contribui para assegurar a esses dados validade como registros arquivísticos, significando que eles podem ser usados no futuro como evidência legal.

O processo de curadoria de dados permite que pesquisadores, professores, estudantes, entre outros, possam realizar e dar continuidade aos seus trabalhos fazendo o uso de dados de pesquisa primários. Ou seja, novos estudos, podem analisar ou acrescentar novos elementos a esses dados, fazendo com que outros projetos sejam reutilizados, reiniciando um novo ciclo. Existem várias propostas de modelo de ciclo de vida de dados de pesquisa, cada um com suas particularidades, esses modelos são compostos por estruturas que se efetivadas garantem usabilidade otimizada e estendida aos dados (SAYÃO; SALES, 2015).

Dentre esses modelos de ciclo de vida de dados de pesquisa, Sayão e Sales (p. 11, 2015) destacam os seguintes modelos: *Digital Curation Centre (DCC) Curation Lifecycle Model*<sup>5</sup>; *DataONE Data Lifecycle*<sup>6</sup>; *DDI Combined Lifecycle model*<sup>7</sup> e; *UK Data archive data lifecycle*<sup>8</sup>.

Dentre os modelos mencionados, destacamos o ciclo de vida desenvolvido pela iniciativa DataONE, o qual, segundo o mesmo, tem como objetivo:

O objetivo do gerenciamento de dados é produzir conjuntos de dados de auto-descrição. Se você dá a seus dados para um cientista ou um colega que não está envolvido com o projeto, eles serão capazes de fazer sentido? Será que eles vão ser capazes de usá-lo de forma eficaz e adequadamente? Nós descrevemos algumas práticas de gestão de dados fundamentais que lhe permitirá desenvolver um plano de gestão de dados, bem como a forma de criar de forma eficaz, organizar, gerenciar, descrever, preservar e compartilhar dados<sup>9</sup>. (DataONE, *online*, p.1) (Tradução nossa).

O ciclo de vida dos dados DataONE, é constituído por 08 (oito) componentes. Conforme ilustrado na Figura 1, a seguir.

<sup>5</sup> <<http://www.dcc.ac.uk/resources/curation-lifecycle-model>>

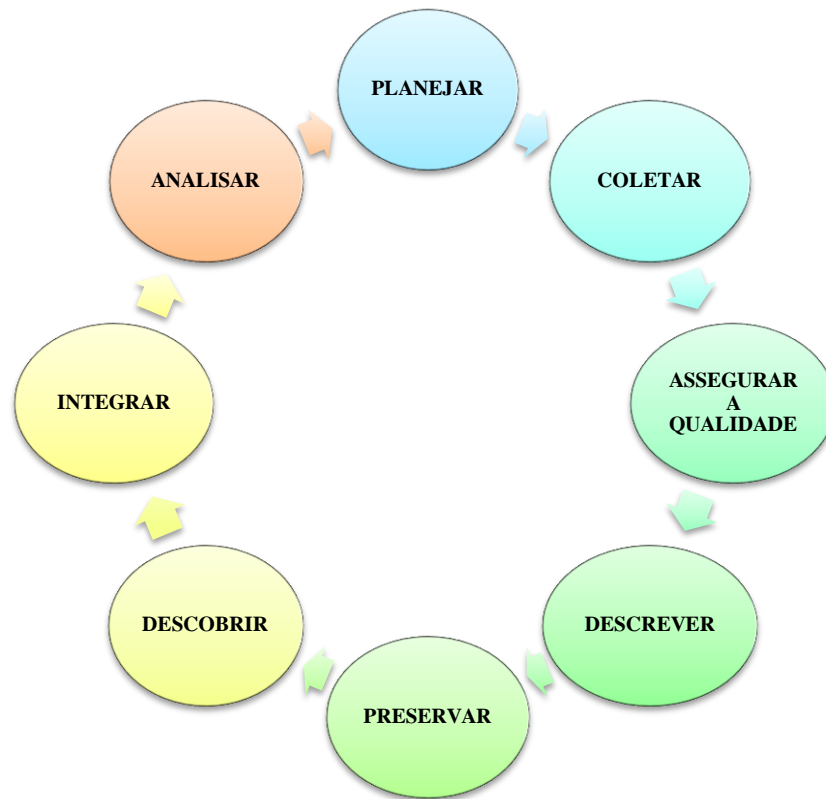
<sup>6</sup> <<http://escholarship.org/uc/item/7tf5q7n3#page-1>>

<sup>7</sup> <<http://www.ddialliance.org/Specification/DDI-Lifecycle/>>

<sup>8</sup> <<http://www.data-archive.ac.uk/create-manage/life-cycle>>

<sup>9</sup> The goal of data management is to produce self-describing data sets. If you give your data to a scientist or colleague who has not been involved with your project, will they be able to make sense of it? Will they be able to use it effectively and properly? This primer describes a few fundamental data management practices that will enable you to develop a data management plan, as well as how to effectively create, organize, manage, describe, preserve and share data.

**FIGURA 1** – O ciclo de vida de dados DataONE.



**Fonte:** DataONE (*online*). (Tradução nossa).

Strasser, et al (2015, p. 3 - 7) descrevem as etapas do ciclo de vida dos dados conforme a DataONE:

**Planejar:** descrição dos dados que serão compilados e como eles são gerenciados e disponibilizados para acesso durante o seu tempo de vida. Nesta etapa do *Data Lifecycle*, alguns fatores deverão ser considerados e adotados, dentre eles: como os dados serão gerados (metodologia, instrumentos utilizados); repositório mais adequado para armazenamento desses dados (levando em consideração a área do conhecimento a qual o repositório se destina); como esses dados serão organizados (se por meio de banco de dados, ou outros); como serão gerenciados, descritos (utilizando padrões de metadados); como serão compartilhados e preservados (estabelecer um plano para compartilhamento e preservação desses dados). (STRASSER, et al, 2015).

**Coletar:** as observações são realizadas manualmente ou por sensores ou outros instrumentos, e os dados são colocados em formas digitais, assegurando sua utilização posterior. Recomenda-se criar um modelo para coleta, assegurando que qualquer dado relevante seja recolhido, além de descrever o conteúdo dos seus dados (unidades, valores, tipos de arquivos). Fazendo-se oportuno organizar os dados dentro de um arquivo (planilha ou

base de dados), e adotar um *software* que realize o *backup* desses dados (STRASSER, et al, 2015).

**Assegurar a qualidade:** a qualidade dos dados é assegurada por meio de controles e inspeção, durante a sua coleta, entrada e análise, recomenda-se a utilização de resumos estatísticos e gráficos, a fim de verificar se há valores questionáveis (STRASSER, et al, 2015).

**Descrever:** dados são descritos usando padrões de metadados apropriados. Aconselha-se descrever alguns itens, como: o ambiente digital (nome do arquivo, data, software); descrever o pessoal e as partes interessadas (quem recolheu os dados, a que área se destina); descrever a razão científica do por que foram coletados; descrever os parâmetros utilizados. A fim de realizar uma descrição minuciosa dos dados objetivando a sua compreensão e utilização (STRASSER, et al, 2015).

**Preservar:** dados são submetidos a arquivos apropriados para preservação de longo prazo. Sugere-se que o pesquisador trabalhe em conjunto com um centro de dados, já que os mesmos podem informar e fornecer informações visando uma preservação em longo prazo, um armazenamento preciso, atinar sobre as considerações legais e políticas acerca da utilização e reutilização dos dados (STRASSER, et al, 2015).

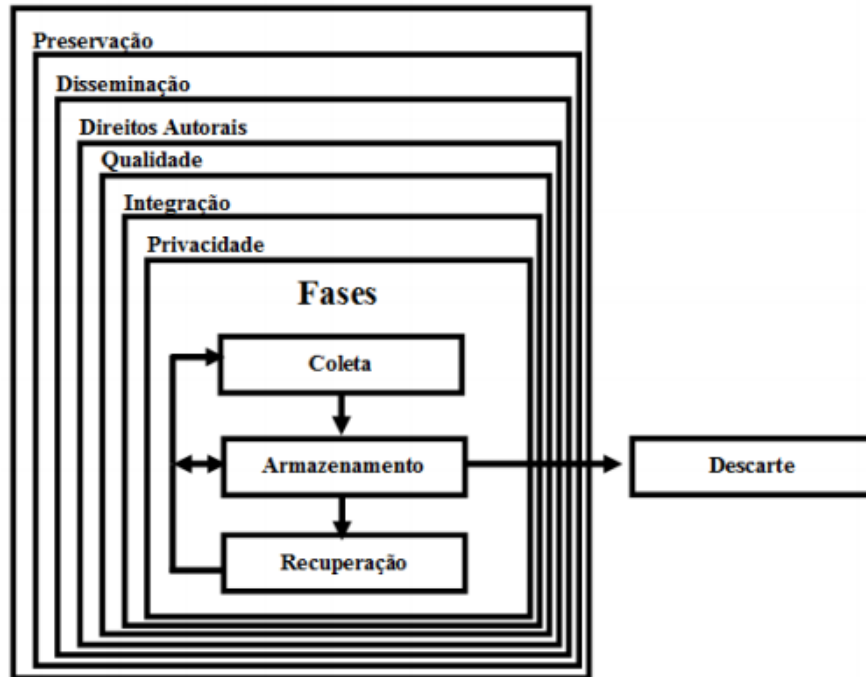
**Descobrir:** dados úteis são descobertos e acessados juntamente com os metadados (STRASSER, et al, 2015).

**Integrar:** dados diversos são combinados para formar um conjunto de dados homogêneo para serem analisados (STRASSER, et al, 2015).

**Analisar:** Dados são analisados (STRASSER, et al, 2015)

Levando em consideração todos os modelos de ciclo de vida de dados de pesquisa que foram citados anteriormente, em conjunto com outros, somando as características pertinentes à área de Ciência da Informação, Santana (2013, *online*), elaborou o modelo de Ciclo de Vida de Dados para Ciência da Informação (CVD-CI), levando em consideração o acesso aos dados.

**FIGURA 2** – Ciclo de vida de dados para Ciência da Informação.



Fonte: Santana (2013, *online*).

O CVD-CI considera como objetivos da Ciência da Informação, o estudo do comportamento da informação, seu fluxo e os meios para o seu acesso. Para compreensão do processo, considerando o acesso aos dados, o modelo CVD-CI, propõe as fases de coleta, armazenamento, recuperação e descarte. Estas, sendo mediadas pelos objetivos de privacidade, qualidade, direitos autorais, integração, disseminação e preservação (SANTANA, 2013, *online*).

As fases do modelo CVD-CI, serão descritas resumidamente a seguir, segundo Santana (2013, *online*):

**Coleta:** atividades vinculadas à definição inicial dos dados a serem utilizados, no que diz respeito ao planejamento de como serão obtidos, filtrados, organizados, e a estrutura no que tange ao formato, meios de descrição e metadados utilizados.

**Armazenamento:** atividades que visam à persistência dos dados em um suporte digital, dentre elas, processamento, transformação, inserção, modificação, migração, transmissão, entre outras.

**Recuperação:** concretiza-se o acesso aos dados. Atividades relacionadas à consulta e visualização (estruturação, filtro, tratamento, representação, refinamento, interatividade).

Em resumo, o modelo CVD-CI, de acordo com Santana (2013, *online*):

No processo descrito no CVD-CI a fase de coleta permite que seja iniciada a fase de armazenamento que por sua vez propicia a execução da fase de



recuperação e ainda pode gerar novos dados retomando ações da fase de coleta. Uma vez em execução atividades da fase de recuperação, tem-se mais uma vez a possibilidade de geração de novos dados originados nesta fase o que retroalimenta o ciclo e retoma ações da fase de coleta para situações novas ou diretamente a fase de armazenamento para dados que já são esperados como resultado das ações ocorridas na fase de recuperação. Alcançados os objetivos previamente estabelecidos e ou alcançados os limites de tempo previstos para os dados, passa-se a fase de descarte, que dependendo da situação irá gerar dados, seja por transferência para outras bases seja para efeito de preservação de informações selecionadas da base a ser descartada, originando assim o início de atividades, ora da fase coleta, ora da fase de armazenamento.

Com a realização das práticas de gestão de dados no início do ciclo de vida dos mesmos, os pesquisadores estarão cada vez mais preparados para colaborar com o conhecimento científico, através da reutilização, integração e síntese desses dados (Strasser et al, 2015, p. 7).

O ciclo de vida de dados serve como uma ferramenta que auxilia os seus usuários em todos os estágios do ciclo de vida dos seus dados. As ações associadas a cada etapa do ciclo de vida dos dados são fundamentais para o processo de curadoria. Os pesquisadores das mais diversas áreas do conhecimento deveriam ter conhecimento deste assunto.

### 3.3 REFLEXÃO ACERCA DA CIÊNCIA ORIENTADA A DADOS

Para alguns pesquisadores, a *e-Science* se configura como uma oportunidade para acelerar o desenvolvimento da comunicação científica, posto que a imensidão de dados gerados por investigações científicas podem ser coletados, comparados e analisados, transformando-se em informação e conhecimento, além de tornaram-se subsídios para novos conhecimentos e novas questões de pesquisa (SAYÃO; SALES, 2014).

Em contrapartida, outros pesquisadores fazem críticas ao modelo da ciência orientada a dados, afirmando que a mesma ignora o conhecimento externo e leva em consideração apenas os dados. Alegam ainda que a reunião, o reuso de dados equivaleria à tentativa de falsificação acarretando na replicação de pesquisas. (FRICKÉ, 2015).

Frické (2015, p.660, tradução nossa) afirma ainda que:

A capacidade de recolher grandes quantidades de dados de forma barata e fácil tem vantagens: Os tamanhos das amostras podem ser maiores, o teste das teorias pode ser melhor, pode haver avaliação contínua e assim por diante. Mas a ciência baseada em dados, o "quarto paradigma", é uma quimera. A ciência precisa de problemas, pensamentos, teorias e

experiências projetadas. Se alguma coisa, a ciência precisa de mais teorias e menos dados<sup>10</sup>.

A ciência orientada a dados, por meio do compartilhamento de dados busca promover um maior crescimento da comunicação científica. Através do uso e reuso dos dados oriundos de investigações científicas, pesquisadores podem validar determinada pesquisa, como também, realizar comparações, como forma de identificar os avanços nas mais diversas áreas do conhecimento.

---

<sup>10</sup> The ability to cheaply and easily gather large amounts of data does have advantages: Sample sizes can be larger, testing of theories can be better, there can be continuous assessment, and so on. But data-driven science, the “fourth paradigm,” is a chimera. Science needs problems, thoughts, theories, and designed experiments. If anything, Science needs more theories and less data.

#### 4 O AMBIENTE DE *E-SCIENCE* E O PROFISSIONAL BIBLIOTECÁRIO

Como visto, o quarto paradigma científico, também denominado como *e-Science*, requer um novo perfil de profissional. Este, dentre tantas habilidades requeridas, deve estar apto para trabalhar no ambiente virtual visando à preservação de informações digitais.

A Ciência da Informação sob a perspectiva da *e-Science*, encontra um amplo campo de atuação para os profissionais da informação. De modo específico, os bibliotecários, e demais profissionais da informação, precisam entender e descobrir seus papéis nesta nova forma de comunicação científica, em meio a esse dilúvio de dados, para então, poderem contribuir para uma gestão eficiente da informação (COSTA; CUNHA, 2014, p.201).

Sayão e Sales (2012, p. 182), estabelecem um novo perfil de profissional da informação apto para lidar com todos requisitos que a *e-Science* demanda, o “cientista de dados”, em que faz-se necessário que o bibliotecário apresente capacidades para lidar com o tratamento de dados oriundos da *e-Science*, realizar o diálogo entre os pares, como também atuar no período de pré e pós-publicação. Igualmente, Tenopir, Birch e Allard (2012, p.12), propuseram o ciclo do bibliotecário, no qual ponderam se o bibliotecário possui habilidades para realizar o processo de curadoria de dados, de acordo com o modelo de ciclo de vida dos dados do DataONE, o *Data Lifecycle*. Conforme demonstrado a seguir:

**FIGURA 3** – Habilidades do bibliotecário no plano de gestão de dados.



**Fonte:** Tenopir, Birch e Allard (2012, p.12) (Tradução nossa).

Baseado no ciclo do bibliotecário proposto por Tenopir, Birch e Allard (2012, p.2), mediante as ponderações realizadas e levando em consideração os procedimentos das etapas do modelo de ciclo de vida dos dados, o *Data Lifecycle*, e o modelo de Ciclo de Vida dos Dados para Ciência da Informação, o CVD-CI, que foi elaborado levando em consideração as características da Ciência da Informação, e subentende-se também que foi destinado aos profissionais da informação. Conclui-se que os bibliotecários, baseado nas práticas que os mesmos desenvolvem, tem *know-how* para desempenhar o papel de cientista de dados como foi definido por Sayão e Sales (2012).

## 5 PROCEDIMENTOS METODOLÓGICOS

A presente pesquisa configura-se como pesquisa exploratória com análise qualitativa e quantitativa.

Trata-se de uma pesquisa exploratória, que visa coletar informações sobre a temática estudada, com vistas à ampliação do entendimento acerca do fenômeno, já que as informações acerca desse assunto ainda são reduzidas. Este tipo de abordagem metodológica é indicada em situações nas quais as informações sobre o problema em estudo são insignificantes ou insuficientes. A pesquisa exploratória tem por finalidade ampliar o conhecimento sobre o tema pesquisado, esclarecendo conceitos e fornecendo subsídios para as etapas subsequentes da investigação (VERGARA, 2006).

A pesquisa assume um caráter quantitativo e qualitativo no que diz respeito às informações coletadas. Assume um caráter quantitativo, acerca da representação em forma numérica dos dados obtidos através da pesquisa, e qualitativo para interpretação e análise dos dados coletados.

Para realização da pesquisa, delimitamos a área do conhecimento a ser pesquisada, de acordo com a CAPES, onde selecionamos a área de Ecologia que integra a grande área de avaliação, Biodiversidade.

**FIGURA 4** – Dados quantitativos de Programas Recomendados e Reconhecidos.

Dados Quantitativos de Programas Recomendados e Reconhecidos										
Área de Avaliação » Área de Conhecimento		Total de Programas de pós-graduação					Totais de Cursos de pós-graduação			
Nome	Área de Avaliação	Total	ME	DO	MF	ME/DO	Total	ME	DO	MF
BOTÂNICA	BIODIVERSIDADE	26	9	0	0	17	43	26	17	0
ECOLOGIA	BIODIVERSIDADE	67	22	0	4	41	108	63	41	4
OCEANOGRAFIA	BIODIVERSIDADE	5	0	0	0	5	10	5	5	0
ZOOLOGIA	BIODIVERSIDADE	43	14	0	1	28	71	42	28	1
<b>Totais</b>		<b>141</b>	<b>45</b>	<b>0</b>	<b>5</b>	<b>91</b>	<b>232</b>	<b>136</b>	<b>91</b>	<b>5</b>

ME: Mestrado Acadêmico  
 DO: Doutorado  
 MF: Mestrado Profissional  
 ME/DO: Mestrado e Doutorado

Fonte: CAPES, online.

Dentre os Programas de Pós-Graduação em Ecologia recomendados e reconhecidos pela CAPES, a investigação optou por estudar os Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro, totalizando em 15 programas de pós-graduação, de 06 estados do nordeste brasileiro, como demonstrados a seguir:

**QUADRO 1** – Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro.

INSTITUIÇÃO	ESTADO	PROGRAMA
Fundação Universidade Federal de Sergipe - FUFSE	SE	Programa de Pós-Graduação em Ecologia e Conservação – PPEC
Universidade Estadual da Paraíba - UEPB	PB	Programa de Pós-Graduação em Ecologia e Conservação – PPGECC
Universidade Estadual de Santa Cruz - UESC	BA	Programa de Pós-Graduação em Ecologia e Conservação da Biodiversidade – PPGECCB
		Programa de Pós-Graduação em Sistemas Aquáticos Tropicais – PPGSAT
Universidade Federal da Bahia - UFBA	BA	Programa de Pós-Graduação em Ecologia e Biomonitoramento – PPGECCOBIO
Universidade Federal da Paraíba/Areia - UFPB	PB	Programa de Pós-Graduação em Biodiversidade – PPGBIO
Universidade Federal da Paraíba/Rio Tinto - UFPB	PB	Programa de Pós-Graduação em Ecologia e Monitoramento Ambiental – PPGEMA
Universidade Federal de Pernambuco - UFPE	PE	Programa de Pós-Graduação em Saúde Humana e Meio Ambiente – PPGSHMA
Universidade Federal do Ceará - UFC	CE	Programa de Pós-Graduação em Ecologia e Recursos Naturais – PPGERN
Universidade Federal do Rio Grande do Norte - UFRN	RN	Programa de Pós-Graduação em Ecologia – PGE
Universidade Federal Rural de Pernambuco - UFRPE	PE	Programa de Pós-Graduação em Biodiversidade e Conservação – PGBC
		Programa de Pós-Graduação em Ecologia – PPGE
		Programa de Pós-Graduação em Etnobiologia e Conservação da Natureza – PPGETNO
Universidade Federal Rural do Semi-Árido - UFRSA	RN	Programa de Pós-Graduação em Ecologia e Conservação - PPEC
Universidade Regional do Cariri - URCA	CE	Programa de Pós-Graduação em Bioprospecção Molecular - PPGBM

Fonte: Dados da pesquisa, 2016.

Foram identificados através dos *sites* dos programas os pesquisadores que constituem o quadro de docentes permanentes e colaboradores dos mesmos, totalizando em 303 pesquisadores que foram mantidos em sigilo. Os *e-mails* dos pesquisadores também foram pesquisados, a fim de enviar o questionário elaborado para tal pesquisa. Vale ressaltar, que 25 pesquisadores deste total, atuam em mais de um programa de pós-graduação, mas no presente estudo, foram contabilizados apenas uma vez.

O questionário intitulado “**Desvendando as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro na gestão dos dados científicos**” (APÊNDICE A), é um questionário semiestruturado, foi elaborado e executado através do *software LimeSurvey*<sup>11</sup>, e baseado no questionário intitulado

<sup>11</sup> <https://www.limesurvey.org/>

“*Research data management assessment: Building understanding of your tools, data management, and training needs*” desenvolvido pela *United States Geological Survey - USGS*.

O instrumento de pesquisa está estruturado em 03 seções, da seguinte forma: seção introdutória (com 01 questão), seção de identificação (02 questões) e, seção sobre dados (19 questões). Totalizando em 22 questões, dentre as quais, fechadas, abertas e mistas. É importante ressaltar, que algumas questões são dependentes entre si, no sentido de que, só se responde determinada questão se tiver respondido a anterior de forma afirmativa ou negativa, concluindo que alguns pesquisadores respondem as 22 questões, e outros não.

O questionário foi inicialmente elaborado e enviado por *e-mail* como forma de teste, para 07 pesquisadores de programas de pós-graduação da Universidade Federal da Paraíba (UFPB), que não possuem o mesmo perfil dos sujeitos desta pesquisa, para que os mesmos pudessem avaliar e dar sugestões de melhora acerca do instrumento de pesquisa, e verificar se o mesmo estaria compreensível.

Para envio e acesso ao questionário, códigos de acesso foram criados, objetivando que apenas pesquisadores convidados respondessem ao instrumento de pesquisa, como também, ter o controle de acesso, para garantir que os pesquisadores respondessem uma única vez o questionário.

Após estudo e acato das sugestões da fase teste, e gerados os códigos de acesso, no dia 27 de outubro de 2016 convidamos os 303 pesquisadores alvo desta pesquisa a participaram do estudo respondendo ao questionário que ficou ativo por 18 dias para acesso e preenchimento, sendo encerrado no dia 13 de novembro de 2016.

**FIGURA 5** – Envio do questionário aos pesquisadores.

<b>Sumário dos códigos de acesso</b>	
<b>Total de registros na tabela de códigos de acesso</b>	303
<b>Total sem código de acesso</b>	0 / 303
<b>Total de convites enviados</b>	303 / 303
<b>Total opted out</b>	0 / 303
<b>Total de questionários completos</b>	1 / 303

**Fonte:** Tela do *software LimeSurvey*

Após a desativação do questionário para preenchimentos por parte dos pesquisadores, houve um problema técnico no software utilizado, e foi preciso extrair os dados da pesquisa diretamente da tabela do banco de dados MySQL através da linguagem de consulta SQL. Por não ter conhecimentos suficientes sobre o manuseio desta linguagem, a tabela com os dados da

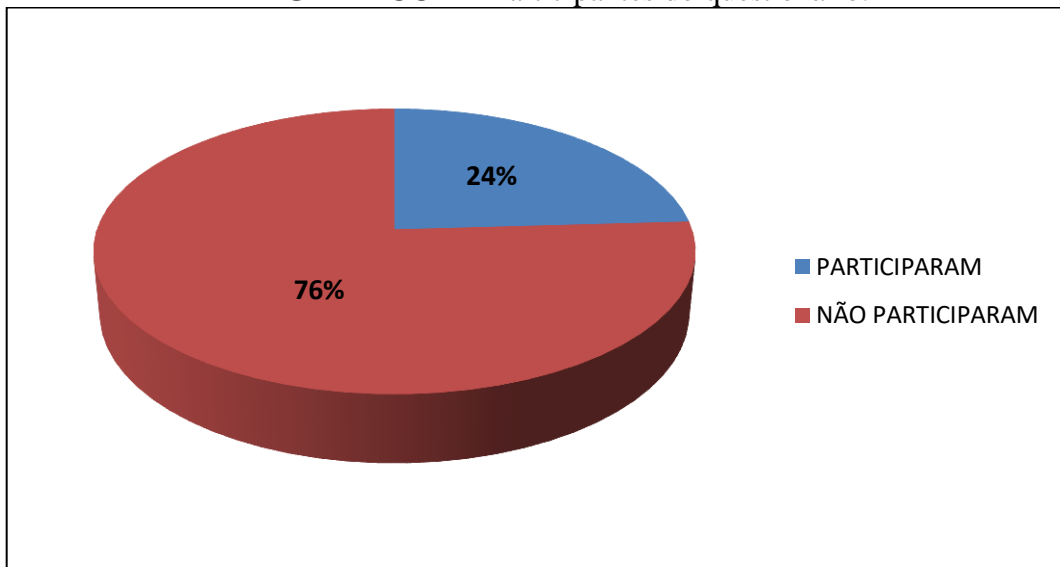
pesquisa foi impressa, e os dados desta investigação foram analisados manualmente, conforme a Figura 6.

**FIGURA 6 – Análise manual dos dados.**

Fonte: Dados da pesquisa, 2016.

Dos 303 pesquisadores convidados para a pesquisa, 73 pesquisadores aceitaram responder o questionário, representando 24% do total, como ilustrado no Gráfico 1.

**GRÁFICO 1 – Participantes do questionário.**



Fonte: Dados da pesquisa, 2016.

Apresentaremos e analisaremos os dados coletados com o instrumento de pesquisa, a seguir.

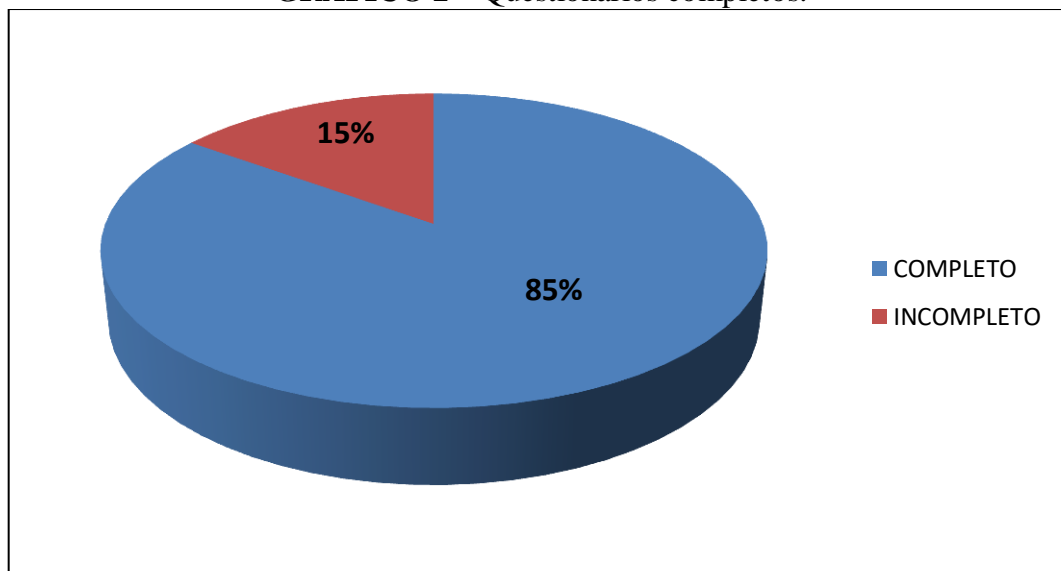


## 6 APRESENTAÇÃO E ANÁLISE DOS DADOS

Apresentaremos e analisaremos os resultados das 22 questões que compuseram o questionário.

Conforme mencionado, 73 pesquisadores aceitaram participar da pesquisa, destes, 62 completaram o questionário, enquanto que 11 pesquisadores abandonaram a pesquisa e não finalizaram o instrumento de pesquisa, como apresentado no Gráfico 2. Ressalta-se que as respostas destes 15% que não finalizaram o questionário, não serão contabilizadas e analisadas.

**GRÁFICO 2** – Questionários completos.



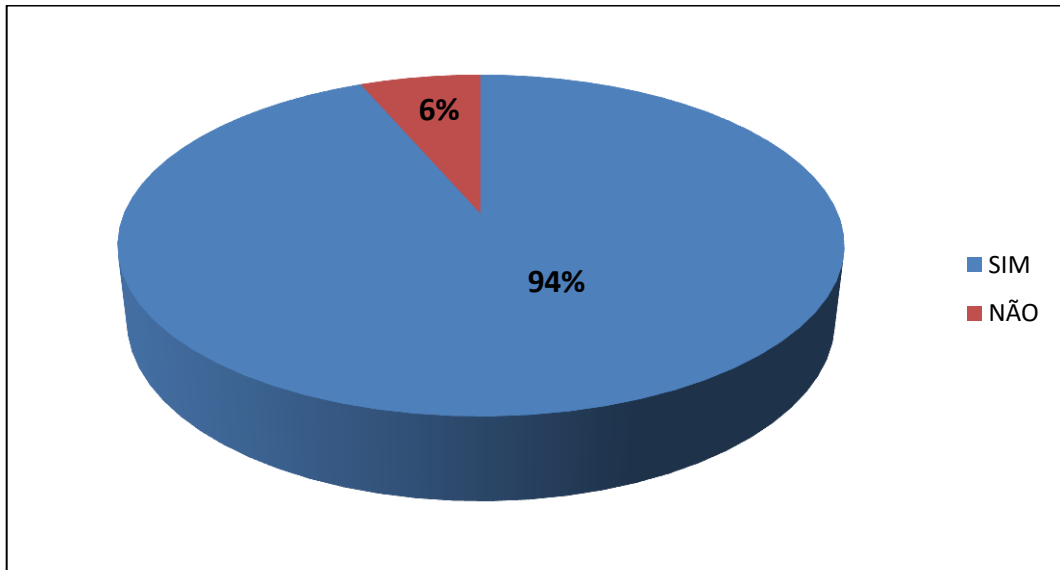
**Fonte:** Dados da pesquisa, 2016.

Temos então um total de 62 questionários completos, os quais terão suas respostas analisadas a seguir.

A primeira questão dizia respeito à Seção Introdutória, como forma de filtrar os pesquisadores que tinham mais conhecimento sobre o tema, apresentamos uma breve explanação sobre dados de pesquisa, e indagamos se o pesquisador gostaria de responder alguns questionamentos pertinentes à temática introduzida.

Do total de 62 pesquisadores, 58 afirmaram que gostariam de responder os questionamentos sobre a temática e deram prosseguimento ao preenchimento do questionário, enquanto que de maneira oposta 04 disseram que ‘NÃO’, como ilustrado no Gráfico 3. Vale ressaltar que estes 04 pesquisadores finalizaram o questionário e encerraram a participação na pesquisa aqui, visto que as outras questões tinham como pré-requisito a resposta afirmativa a esta questão, mas foram contabilizados como questionários completos.

**GRÁFICO 3** – Gostaria de responder alguns questionamentos pertinentes a temática introduzida?

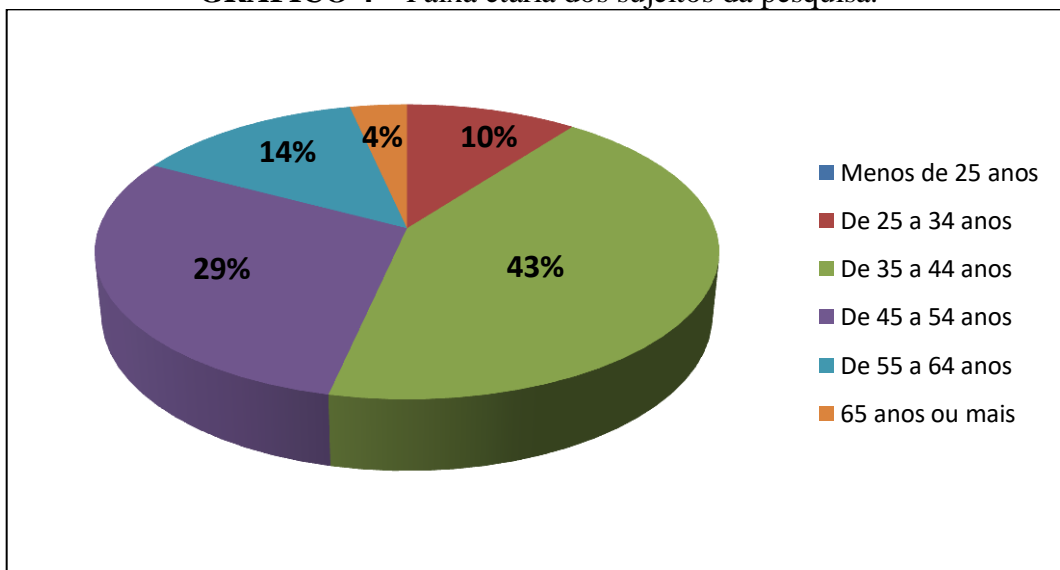


Fonte: Dados da pesquisa, 2016.

A população reduziu-se então a 58 pesquisadores, visto que 04 finalizaram o questionário anteriormente.

As próximas duas questões fazem parte da Seção Identificação, e dizem respeito ao perfil dos pesquisadores. A primeira delas objetivou traçar a faixa etária dos sujeitos desta pesquisa, em que: 06 pesquisadores têm de 25 a 34 anos; 25 pesquisadores entre 35 a 44 anos; 17 pesquisadores têm de 45 a 54 anos; 08 pesquisadores entre 55 a 64 anos; e apenas 02 pesquisadores com 65 anos ou mais. Como pode ser visto e analisado no Gráfico 4.

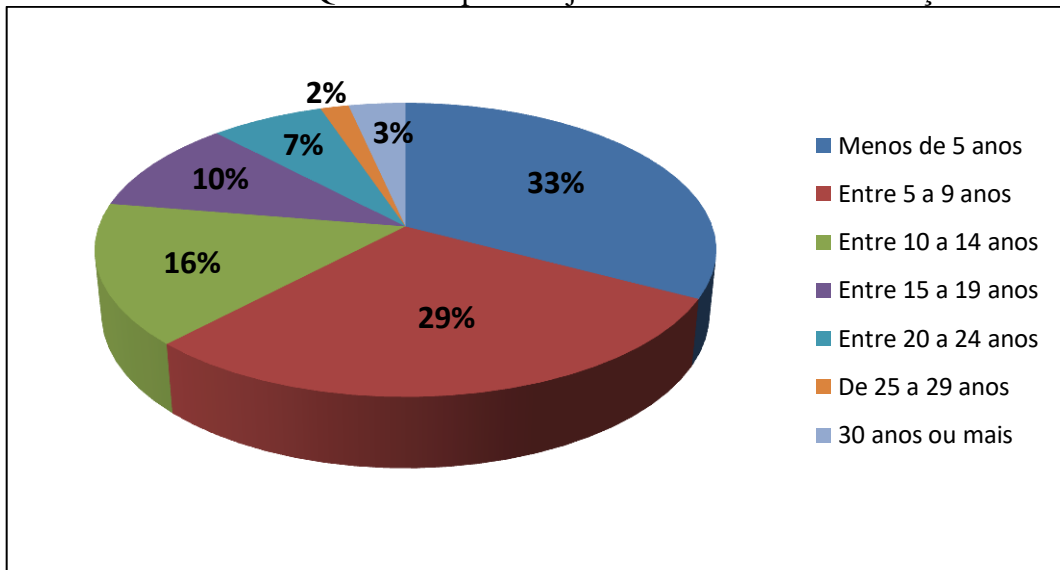
**GRÁFICO 4** – Faixa etária dos sujeitos da pesquisa.



Fonte: Dados da pesquisa, 2016.

A segunda questão da Seção Identificação buscou identificar há quanto tempo os pesquisadores atuam na pós-graduação, em que: 19 pesquisadores atuam há menos de 5 anos; 17 pesquisadores têm entre 5 a 9 anos de atuação; 09 atuam na pós cerca de 10 a 14 anos; 06 pesquisadores têm entre 15 a 19 anos de atuação; 04 atuam há 20 a 24 anos; 01 atua acerca de 25 a 29 anos; e 02 pesquisadores atuam há 30 anos ou mais na pós-graduação. Como representado, em porcentagem, no Gráfico 5.

**GRÁFICO 5** – Quanto tempo os sujeitos atuam na Pós-Graduação.

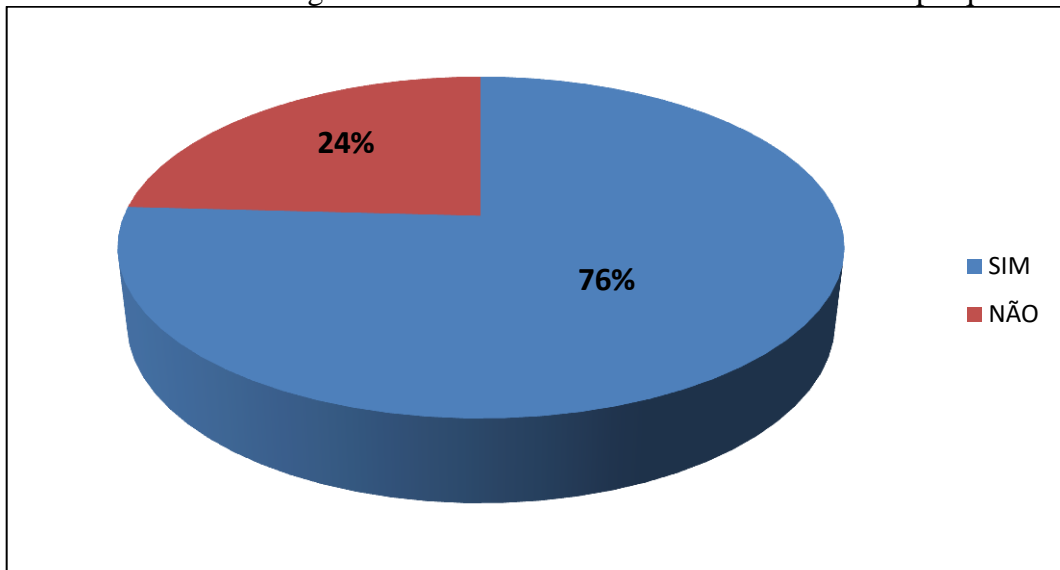


**Fonte:** Dados da pesquisa, 2016.

De acordo com as duas questões da seção introdutória, indicamos que, em sua maioria, 43% dos pesquisadores estão na faixa etária de 35 a 44 anos, e que 33%, também em sua maioria, tem pouco tempo de vínculo, no caso, menos de 5 anos de atuação em pós-graduação.

Dando continuidade à pesquisa, as próximas questões vão tratar sobre os dados de pesquisa, no geral. Como, por exemplo, a gestão dos dados científicos de pesquisa, a natureza dos dados, a origem dos dados, entre outros.

A quarta questão, indagou se os pesquisadores fazem a gestão dos dados científicos coletados nas suas pesquisas, 44 pesquisadores afirmaram fazer a gestão dos dados, enquanto que, de maneira oposta, 14 não fazem a gestão dos dados científicos coletados nas suas pesquisas, como apresentado no Gráfico 6.

**GRÁFICO 6** – Faz a gestão dos dados científicos coletados nas suas pesquisas?

**Fonte:** Dados da pesquisa, 2016.

Como visto, 24% dos sujeitos, o que equivale a 14 pesquisadores, responderam que não faziam gestão dos dados científicos. A próxima questão tinha como pré-requisito a negação à questão anterior, em que foi perguntado qual o motivo desses pesquisadores não fazerem a gestão dos dados científicos coletados nas suas pesquisas. E obtivemos as seguintes respostas, relatadas no Quadro 2, o ID é referente à planilha resultante da pesquisa, e o mesmo utilizado em quadros subsequentes, ou seja, cada pesquisador, tem um ID.

**QUADRO 2** – Por que não faz a gestão dos dados científicos?

<b>Por que não faz a gestão dos dados científicos coletados nas suas pesquisas?</b>	
<b>ID</b>	<b>Justificativa</b>
02	“Na minha área de atuação nos preocupamos mais com o produto final dos experimentos.”
10	“Falta de sistema de gestão.”
11	“Ainda com muitos dados novos para processar.”
13	“Publico em revistas apenas.”
15	“Falta de conhecimento de ferramentas adequadas.”
20	“Não tive essa ideia e não pesquisei formas de fazer isso.”
23	“Muitas vezes, por se tratar de projetos multidisciplinares, apenas parte dos dados ficam em minha posse e organizados de forma a serem reutilizados.”
26	“Ainda não incorporei esta prática na minha produção.”
28	“Não tinha atentado para este aspecto.”
32	“Desconhecimento.”
33	“Não tenho sistema de gestão.”
38	“Porquê todos os dados obtidos nas pesquisas coordenadas ou realizadas por mim se encontram publicadas, ficando comigo os artigos publicados.”
40	“Não tenho tantas informações assim.”
71	“A pergunta não está clara. Me senti mais à vontade em discordar.”

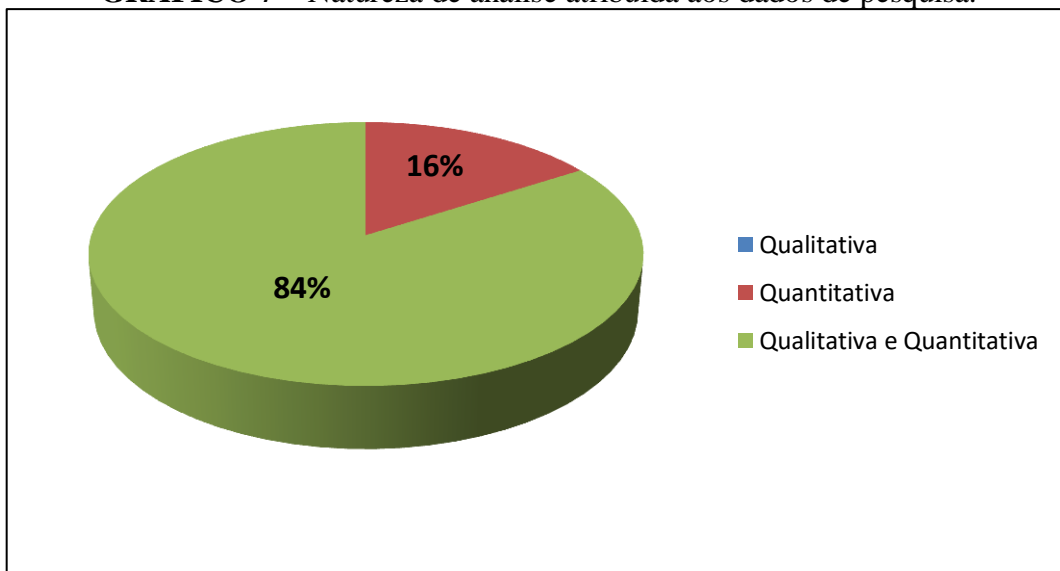
**Fonte:** Dados da pesquisa, 2016.

Percebemos que entre os 14 pesquisadores que não fazem a gestão dos dados científicos, a maioria deles alegam desconhecer o processo de gestão de dados, e afirmam que todos os seus dados estão publicados.

Obedecendo as condições impostas na criação do questionário, os pesquisadores que respondessem “NÃO” na questão 04 e justificassem na questão 05, encerrariam o questionário aqui. Dito isso, a pesquisa seguiu-se com 44 pesquisadores participantes.

A próxima questão investigou como o pesquisador classificaria a natureza da análise atribuída aos dados de pesquisa, se são de natureza qualitativa, quantitativa ou ambas. Como resultado obtivemos que, nenhum pesquisador atribuiu análise qualitativa aos seus dados, 07 pesquisadores atribuíram análise apenas quantitativa, e 37 pesquisadores atribuíram análise qualitativa e quantitativa, conforme exposto no Gráfico 7. Observamos que 84% dos pesquisadores atribuem análise mista (qualitativa e quantitativa) aos seus dados de pesquisa.

**GRÁFICO 7** – Natureza de análise atribuída aos dados de pesquisa.



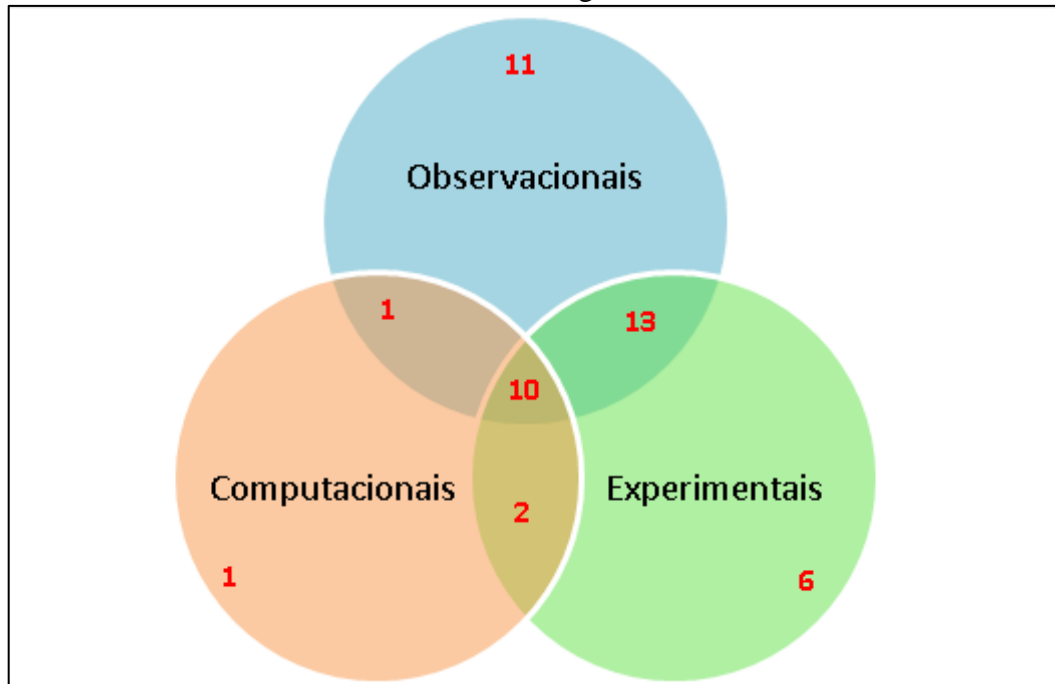
**Fonte:** Dados da pesquisa, 2016.

Prosseguindo a pesquisa, a sétima questão trouxe uma breve explanação sobre a origem dos dados de pesquisa, se são observacionais, computacionais e experimentais, e investigou qual a origem dos dados de pesquisa dos sujeitos. Como nesta questão os pesquisadores podiam assinalar mais de uma alternativa, optamos reproduzir esta questão por meio de conjuntos e para melhor entendimento descreveremos a figura 7, em que, 11 pesquisadores tem dados exclusivamente observacionais; 01 pesquisador tem dados apenas computacionais; e 06 pesquisadores com dados apenas experimentais.

Como dito, os pesquisadores podiam assinalar mais de uma alternativa, na qual, com dados computacionais e observacionais temos 01 pesquisador; dados computacionais e

experimentais 02 pesquisadores; e com dados experimentais e observacionais 13 pesquisadores. Percebeu-se ainda que 10 pesquisadores têm dados de origem observacional, computacional, e experimental, conforme apresentado na Figura 7.

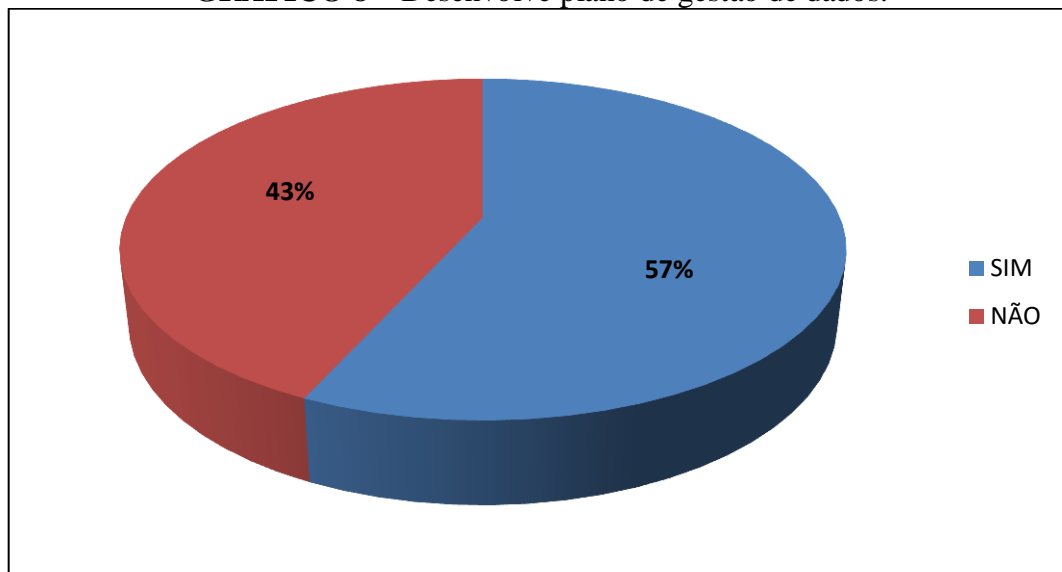
**FIGURA 7** – Origem dos dados.



**Fonte:** Dados da pesquisa, 2016.

Percebemos que os dados dos pesquisadores tem origem observacional e/ou experimental, em maioria.

A oitava questão, desvendou se os pesquisadores desenvolvem algum plano de gestão de dados ao longo da condução de suas pesquisas, as respostas estão reproduzidas no Gráfico 8, em qual 57%, 25 pesquisadores afirmaram desenvolver plano de gestão, enquanto que 19 pesquisadores negaram desenvolver algum plano de gestão.

**GRÁFICO 8** – Desenvolve plano de gestão de dados.

Fonte: Dados da pesquisa, 2016.

A nona questão tem como pré-requisito a resposta de forma negativa à questão anterior. 43% responderam de forma negativa a questão anterior, equivalente a 19 pesquisadores, que foram indagados sobre o motivo de não desenvolverem plano de gestão de dados ao longo da condução de suas pesquisas. As justificativas estão dispostas no Quadro 3.

**QUADRO 3** – Por que não desenvolve plano de gestão de dados ao longo da condução das pesquisas?

Por que não desenvolve plano de gestão de dados ao longo da condução das pesquisas?	
ID	Justificativa
04	“Desconhecimento.”
06	“Pois apenas foram dados coletados por minha pessoa.”
07	“Deixo a cargo dos orientandos.”
09	“Falta de tempo.”
21	“Não existe um planejamento específico, o que é feito é a tabulação dos dados em planilhas e o armazenamento.”
24	“Não tenho familiaridade.”
29	“Ainda é uma base de dados pequena temporalmente para responder a pergunta de cunho ecológico.”
34	“Falta de tempo.”
35	“Não faço.”
41	“Até o momento, a finalidade dos dados é a geração de gráficos e tabelas, os quais irão dar suporte a expressar os resultados. Após a publicação dos resultados, os dados são abandonados. Então, não existe nenhuma gestão específica.”
50	“Falta de tempo para planejar adequadamente as atividades de pesquisa.”
52	“Porque não atentei para essa questão até recentemente, quando procurei reutilizar os dados já produzidos.”
59	“Prefiro a análise no final da pesquisa e não durante.”
62	“Ainda não consegui aplicar as ferramentas necessárias. Mas tenho interesse.”
63	“Não tenho tempo, nem o perfil adequado para isto.”
64	“Falta de formação e de organização para este tipo de gestão.”
66	“Não está claro o que significa ‘plano de gestão’.”

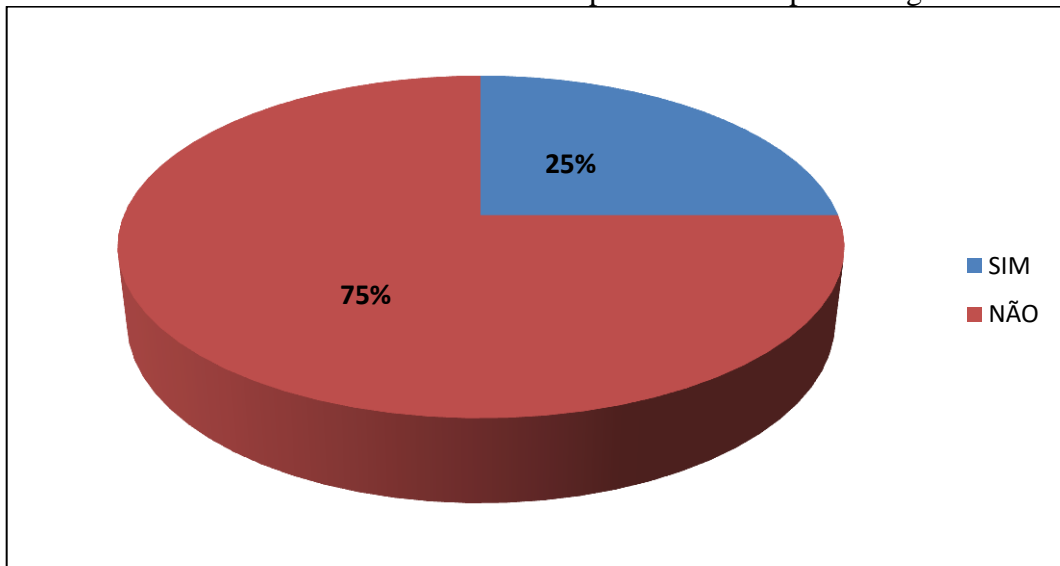
67	“É importante o desenvolvimento prático e descrito nas pesquisas, considerando também o uso de bases digitais.”
72	“Porque já o realizei antes de iniciá-la, durante a etapa de planejamento amostral e análise de dados.”

Fonte: Dados da pesquisa, 2016.

Nota-se que dentre os 19 pesquisadores que não desenvolvem plano de gestão de dados ao longo da condução de suas pesquisas, alguns pesquisadores alegam não ter tempo para tal, outros apresentam desconhecimento ou não atentaram para o assunto, e os demais afirmam fazer o uso de planilhas para tal função. Observou-se que alguns pesquisadores afirmaram realizar gestão de dados de pesquisa na questão 4, conforme pode ser visto no Gráfico 6, e nesta resposta aberta afirmou que não desenvolvem plano de gestão de dados, sendo possível notar que alguns pesquisadores apresentam desconhecimento acerca desta temática.

A décima questão interrogou se os pesquisadores fazem o uso de alguma ferramenta automatizada para auxiliar a criação do plano de gestão de dados. Dos 44 pesquisadores ativos na pesquisa, 33 não utilizam nenhuma ferramenta, representando 75%, e 11 afirmaram fazer o uso de ferramentas no auxílio de tal função, representando 25%, como exposto no Gráfico 9.

**GRÁFICO 9** – Uso de ferramenta automatizada para auxílio no plano de gestão de dados.



Fonte: Dados da pesquisa, 2016.

Como apresentado, 25%, o que equivale a 11 pesquisadores, afirmaram fazer uso de ferramenta automatizada para auxílio na criação do plano de gestão de dados. A questão 11 solicitou a estes pesquisadores que indicassem quais são essas ferramentas. As ferramentas indicadas estão relacionadas no Quadro 4.



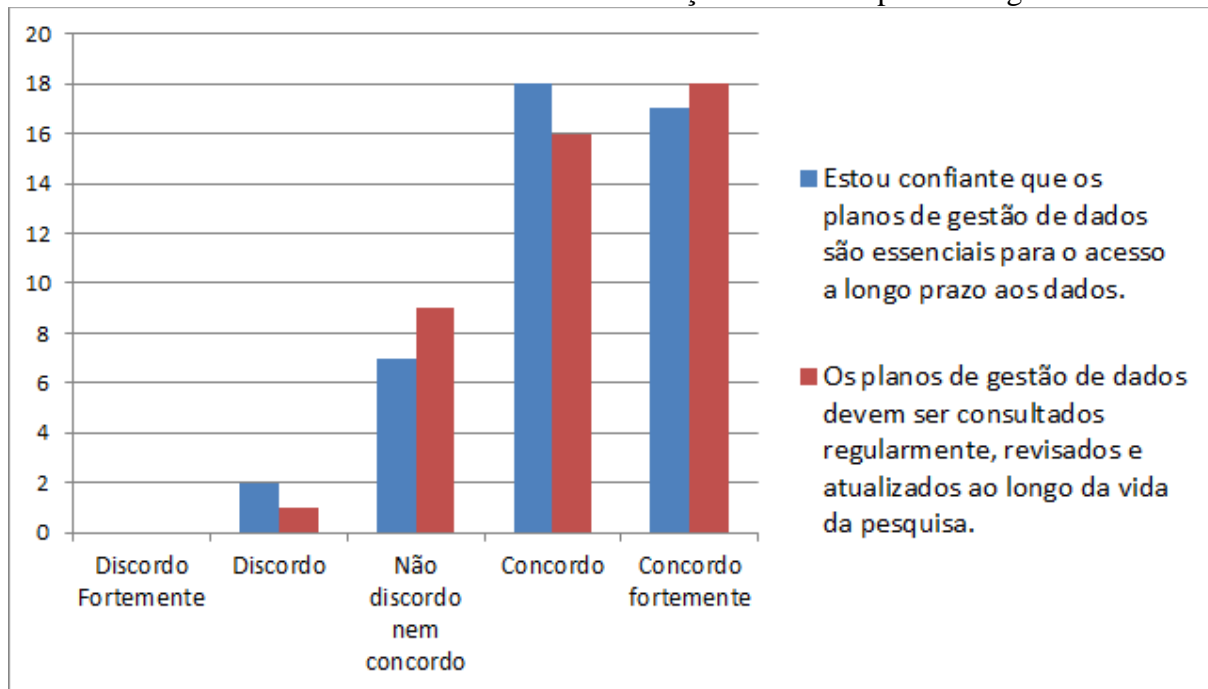
**QUADRO 4** – Ferramentas utilizadas na criação do plano de gestão de dados.

Indique quais as ferramentas utilizadas na criação do plano de gestão de dados.	
ID	Justificativa
01	“Excel, GDrive, Mendeley.”
03	“ECOCEM, Access e OpenOffice.”
05	“Excel, R.”
08	“Access, Excel.”
17	“Planilhas Excel.”
18	“Banco de dados, planilha Excel.”
19	“Banco de dados.”
45	“Planilha e tabela com banco de dados obtidos, e metodologia utilizada.”
53	“Excel, R, Trello.”
67	“Capes, Science Direct, Bireme.”
70	“Planilha de Excel.”

Fonte: Dados da pesquisa, 2016.

Dentre as ferramentas mais utilizadas na criação do plano de gestão de dados por parte dos pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro, a planilha Excel é utilizada por 07 dos 11 pesquisadores que afirmaram fazer o uso de ferramentas automatizadas para auxílio na criação do PGD.

A questão doze era composta por duas declarações acerca dos planos de gestão de dados, as quais os pesquisadores tinham que indicar o nível de concordância (discordo fortemente até concordo fortemente) com cada uma delas. O Gráfico 10 apresenta o nível de concordância para cada declaração.

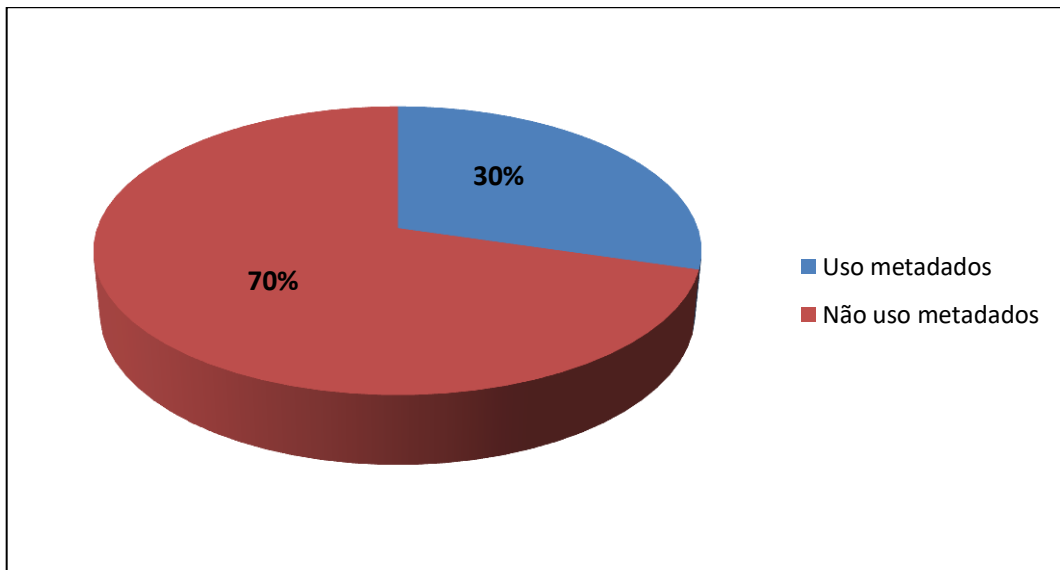
**GRÁFICO 10** – Nível de concordância com declarações acerca de planos de gestão de dados.

Fonte: Dados da pesquisa, 2016.

Foi possível observar que os pesquisadores tem ampla convicção da importância dos planos de gestão de dados para o acesso em longo prazo aos dados, e que para isso, os mesmos devem ser consultados, revisados e atualizados ao longo da pesquisa.

A décima terceira questão investigou se os pesquisadores fazem o uso de metadados, e em caso afirmativo, que indicassem quais os padrões de metadados eles utilizam para a descrição dos dados. Dos 44 pesquisadores, 31 não fazem o uso de metadados, e 13 fazem o uso de metadados. Conforme o Gráfico 11.

**GRÁFICO 11 – Uso de metadados.**



**Fonte:** Dados da pesquisa, 2016.

Ao afirmar que faziam uso de metadados, perguntamos quais os padrões de metadados eles utilizavam para descrever os seus dados. Dos 13 pesquisadores que afirmaram usar metadados, apenas 05 indicaram os padrões utilizados, as respostas estão descritas no Quadro 5, a seguir.

**QUADRO 5 – Padrões de metadados utilizados.**

<b>Indique quais as ferramentas utilizadas na criação do plano de gestão de dados.</b>	
<b>ID</b>	<b>Justificativa</b>
01	“Data, localidades, comportamentos, dados climáticos, equipamento utilizado no registro.”
05	“Utilizo dados de outros estudos ou pesquisadores em minhas pesquisas.”
49	“Visto que Metadados, são dados sobre outros dados é comum nas minhas pesquisas consultar a literatura sobre as pesquisas correlacionadas com o tema de modo a fazer a discussão dos resultados com base em dados já existente na literatura.”
51	“Localização geográfica, unidades de medidas, data e horário, métodos utilizados.”
64	“Localização (geográfica) de amostras, descrição do tipo de amostragem realizada.”

**Fonte:** Dados da pesquisa, 2016.

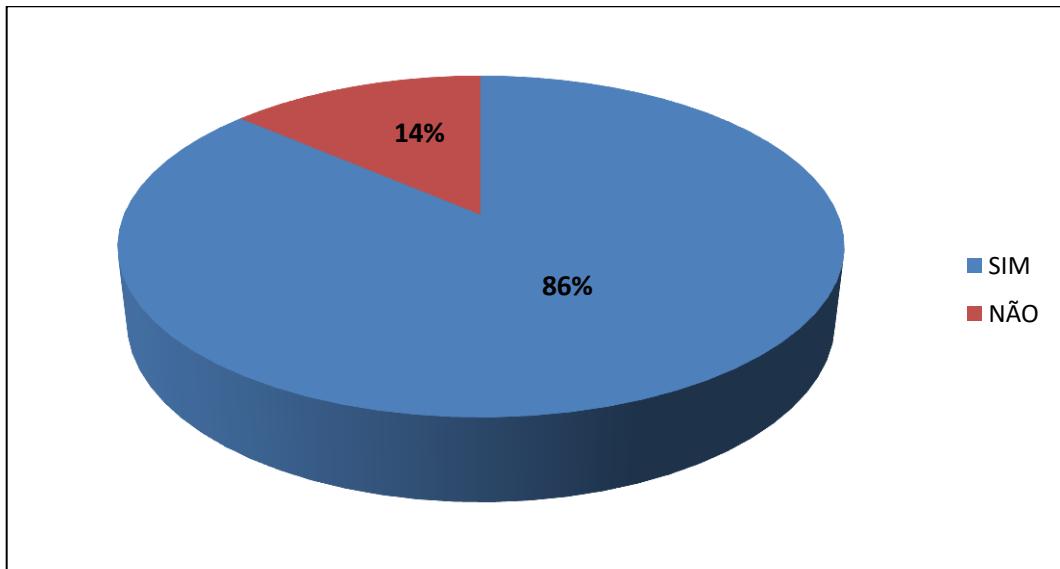
Verificou-se que 03 deles, realizam a descrição dos dados oriundos de suas investigações científicas, mas não souberam informar os padrões de metadados utilizados, já

que informaram apenas elementos, campos, que compõe um padrão de metadados. Outros 02 pesquisadores, não indicaram o padrão de metadados utilizado por fazerem reuso de dados de outros pesquisadores.

As questões, 14 e 15, identificou o uso de ferramentas de software que possibilite a exploração, visualização ou análise de dados, e quais seriam estas ferramentas.

Dos 44 pesquisadores, 38 utilizam ferramentas de software para tais finalidades representando 86%, enquanto que 06 pesquisadores, representando 14% não fazem uso de ferramenta de software para exploração, visualização ou análise de dados, conforme exposto no Gráfico 12.

**GRÁFICO 12** – Uso de ferramenta de software para exploração, visualização ou análise de dados.



Fonte: Dados da pesquisa, 2016

Em seguida foi perguntado aos 38 pesquisadores que utilizam ferramentas de software para exploração, visualização ou análise de dados, quais seriam essas ferramentas. O Quadro 06 relata as ferramentas citadas pelos pesquisadores.

**QUADRO 6** – Ferramentas de software utilizadas para exploração, visualização ou análise de dados.

Indique quais são as ferramentas de <i>software</i> utilizadas para exploração, visualização ou análise de dados.	
ID	Justificativa
01	“Excel.”
03	“R.”
04	“R, Statistica, Sigma-Stat 3.1, Past (Paleontological Statistics 2.15),Stimates 9.1.”
05	“R.”
06	“R.”
07	“Mata Nativa e softwares estatísticos.”

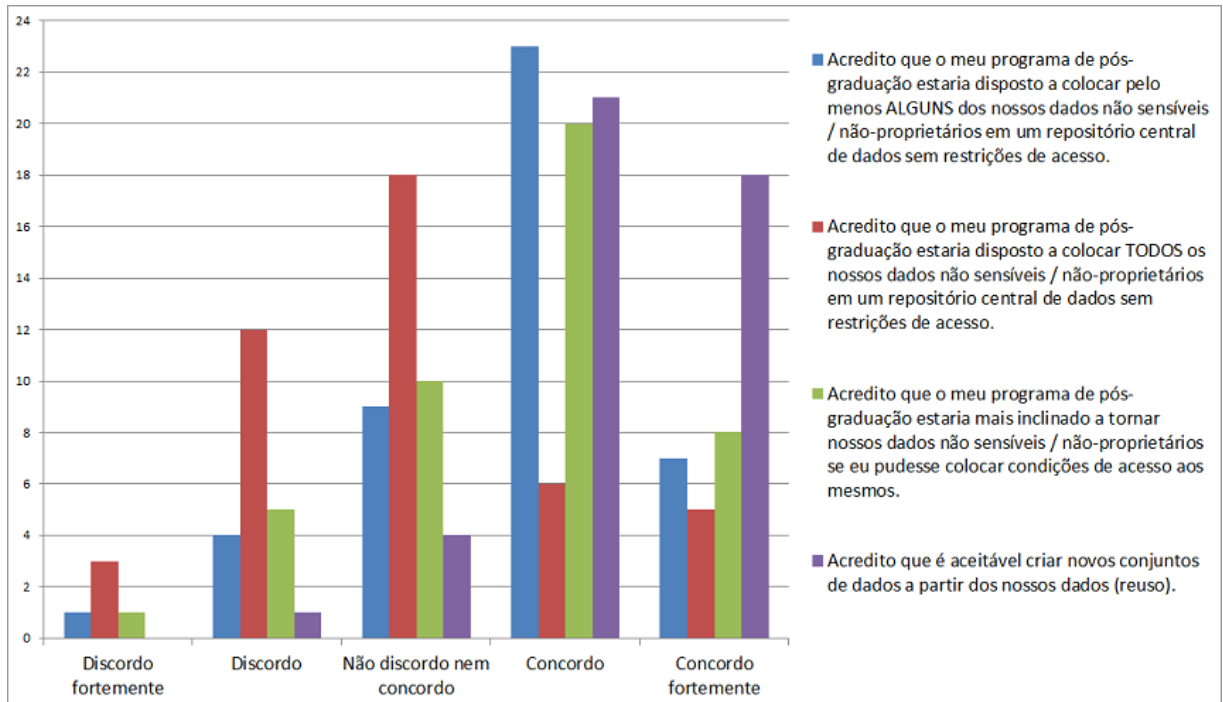
08	“R.”
09	“Surfer, Origin, Past, e outros softwares.”
14	“EstimateS.”
16	“R.”
17	“Excel.”
19	“Excel.”
21	“R, diva-gis, permanova, circuitscape, maxent.”
22	“Free commander XE.”
24	“Estatísticos.”
25	“R.”
29	“PRIMER, CANOCO e Statistica.”
31	“Excel.”
34	“Excel, R, Statistica.”
35	“R.”
37	“Programa Estatístico R.”
41	“Planilha eletrônica.”
45	“Image J e GraphPad Prism.”
48	“Software R.”
49	“Web of Science (base de dados), Origin (ferramenta de análise de dados gráficos).”
50	“SPSS, Statistica, BioEstat, R, Excel, Fitopac”.
52	“R, Past, Quantum Gis, Corel Draw, Excel.”
53	“R.”
56	“Excel, R.”
59	“Primer, R, Statistica.”
61	“R.”
62	“R, Bioestat.”
64	“SIG (QGIS).”
67	“Tratamento estatístico de dados.”
68	“Excel e programas de dinâmica de populações.”
70	“Excel, Systat, SPSS, Statistica, R.”
72	“R.”
73	“Access.”

**Fonte:** Dados da pesquisa, 2016.

Verificou-se que mais da metade dos sujeitos da pesquisa utilizam ferramentas de software para exploração, visualização ou análise de dados, e que dentre elas, as mais utilizadas pelos pesquisadores foram: *software* R, com 21 ocorrências; Planilha Excel com 11 ocorrências; e o *software* Statistica com 06 ocorrências.

A décima sexta questão tinha 04 declarações que abordavam sobre o compartilhamento de dados, as quais os pesquisadores tinham que indicar seu nível de concordância (discordo fortemente até concordo fortemente) com cada uma delas. Como apresentado no Gráfico 13.

**GRÁFICO 13** – Nível de concordância com declarações acerca do compartilhamento de dados.



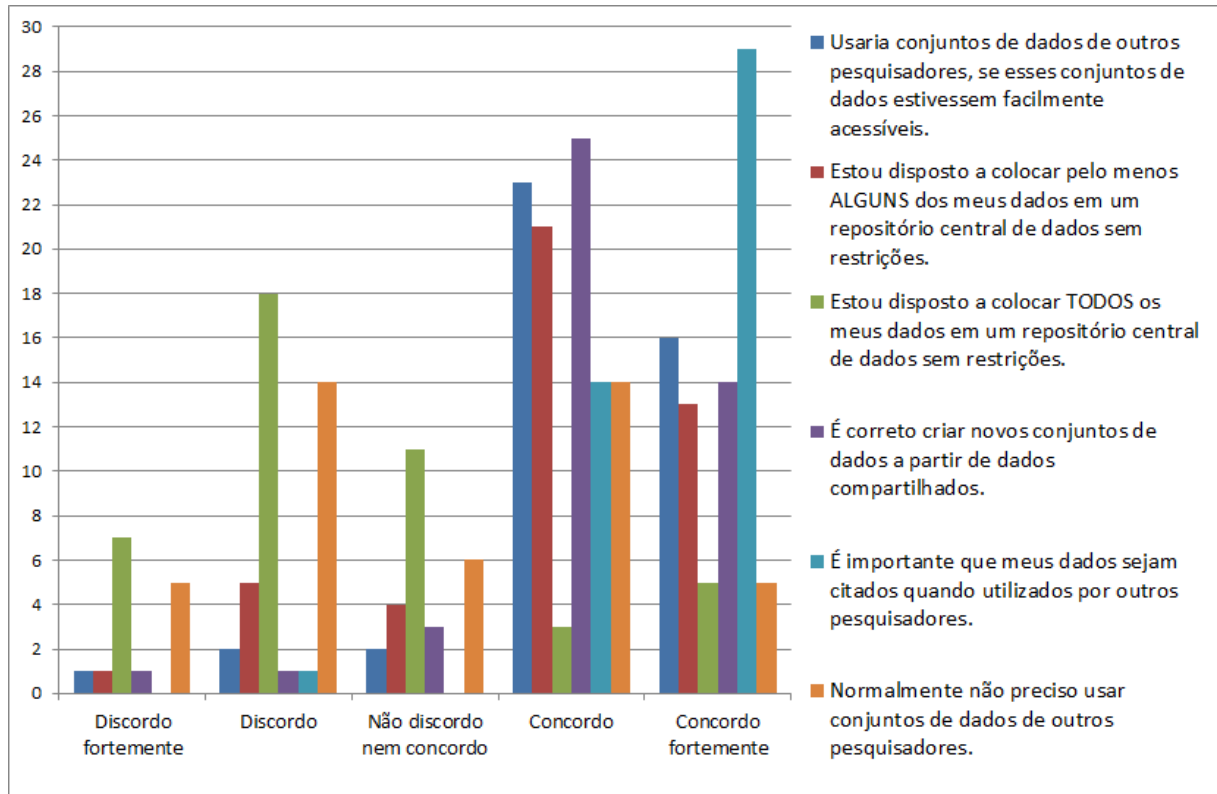
Fonte: Dados da pesquisa, 2016

De acordo com Gráfico 13 exposto, percebeu-se 23 pesquisadores concordam com a primeira declaração; 18 pesquisadores não discordam nem concordam com a segunda declaração, mostrando indiferença; 20 pesquisadores concordam com a terceira declaração; e 21 pesquisadores concordam com a última declaração.

Analisando as declarações com maior nível de concordância, foi possível identificar que os pesquisadores, em sua maioria, são a favor do compartilhamento de dados, visto que acreditam que o programa de pós-graduação estaria disposto a compartilhar os dados do corpo docente em um repositório de dados sem restrições de acesso, e acreditam ainda que é aceitável criar novos conjuntos de dados a partir dos seus dados, ou seja, concordam com a ideia de reuso.

A próxima questão, também era composta por 06 declarações referentes ao compartilhamento de dados, as quais os pesquisadores tinham que estabelecer o nível de concordância com cada uma delas, estes níveis de concordância, são os mesmos das questões anteriores.

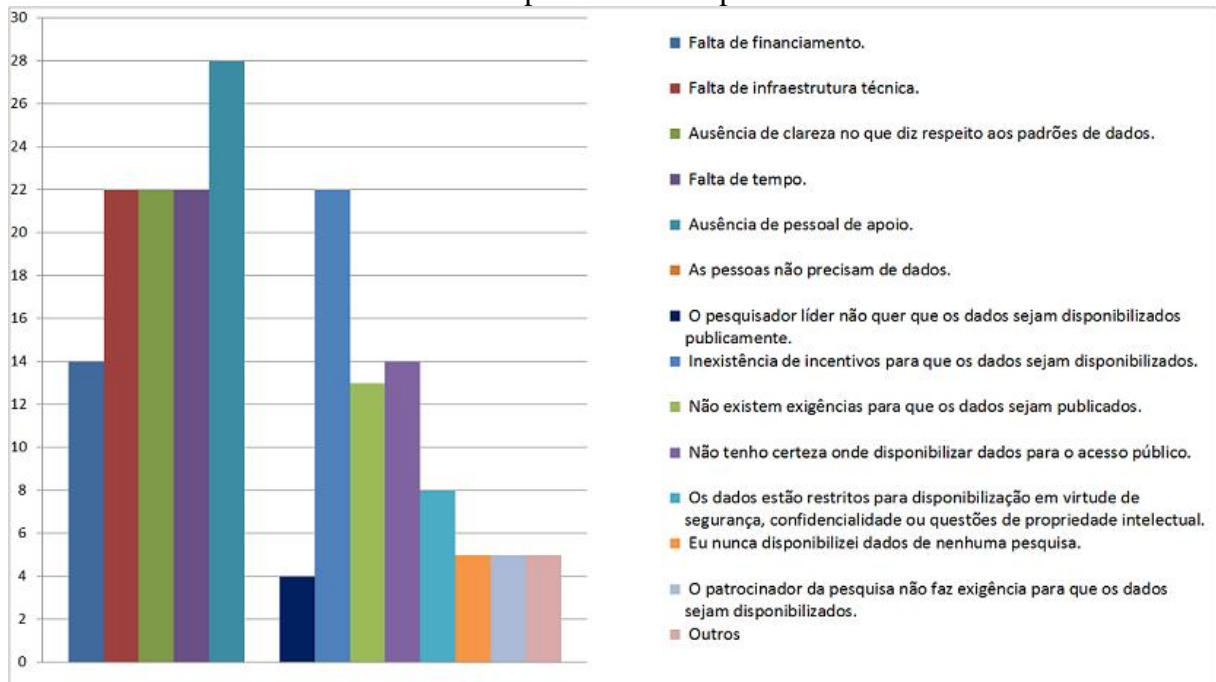
**GRÁFICO 14** – Nível de concordância com declarações referentes ao compartilhamento de dados.



Fonte: Dados da pesquisa, 2016.

Podemos analisar que dentre os 44 pesquisadores, notou-se que 23 pesquisadores concordaram com a primeira declaração afirmando que realizariam a prática de reuso de dados. Em relação à segunda declaração, 21 pesquisadores concordaram com a disponibilização de alguns dos seus dados em um repositório de dados sem restrições, sendo favoráveis ao compartilhamento de dados. Em contrapartida, 18 pesquisadores discordaram da possibilidade de disponibilizar todos os seus dados em um repositório de dados sem restrições. A quarta declaração julgou ser correto à criação de novos conjuntos de dados a partir de outros dados, com 25 pesquisadores concordando, reafirmando serem favoráveis ao compartilhamento de dados. A quinta declaração, teve o maior nível de concordância quando 29 pesquisadores julgaram ser importante que os dados sejam referenciados quando reutilizados. E a última declaração, tiveram 14 concordâncias, e 14 discordâncias, na qual os pesquisadores julgaram não utilizar conjuntos de dados de outros pesquisadores em suas investigações científicas.

A décima oitava questão investigou quais as razões para o não compartilhamento de dados coletados pelos pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro. Foram relacionadas 14 razões para tal fato, as quais os pesquisadores podiam selecionar quantas se aplicavam. Conforme o Gráfico 15.

**GRÁFICO 15** – Razões para o não compartilhamento de dados.

Fonte: Dados da pesquisa, 2016.

De acordo com o Gráfico 15, percebeu-se que a falta de infraestrutura técnica; ausência de clareza dos padrões de metadados; falta de tempo; ausência de pessoal de apoio; e a inexistência de incentivo para disponibilização dos dados, foram às alternativas mais selecionadas pelos pesquisadores como justificativa para o não compartilhamento de dados coletados por eles. Através do Gráfico 15, percebe-se que 05 pesquisadores selecionaram a opção ‘Outros’, e ao selecionar esta opção, os pesquisadores tinham que escrever quais seriam estas razões, conforme descritas no Quadro 7.

**QUADRO 7** – Outras razões para o não compartilhamento de dados.

Razões para o não compartilhamento de dados.	
ID	Justificativa
06	“Estou trabalhando em cima desses dados ainda, e caso os tornassem disponíveis, pesquisadores podem realizar análises similares das quais estou trabalhando, o que inviabilizaria a publicação de artigos futuros.”
22	“Ausência de clareza no que diz respeito aos deveres do usuário dos dados.”
25	“Compartilho meus dados.”
61	“Dados são compartilhados.”
73	“Não há clareza de como os dados serão usados.”

Fonte: Dados da pesquisa, 2016.

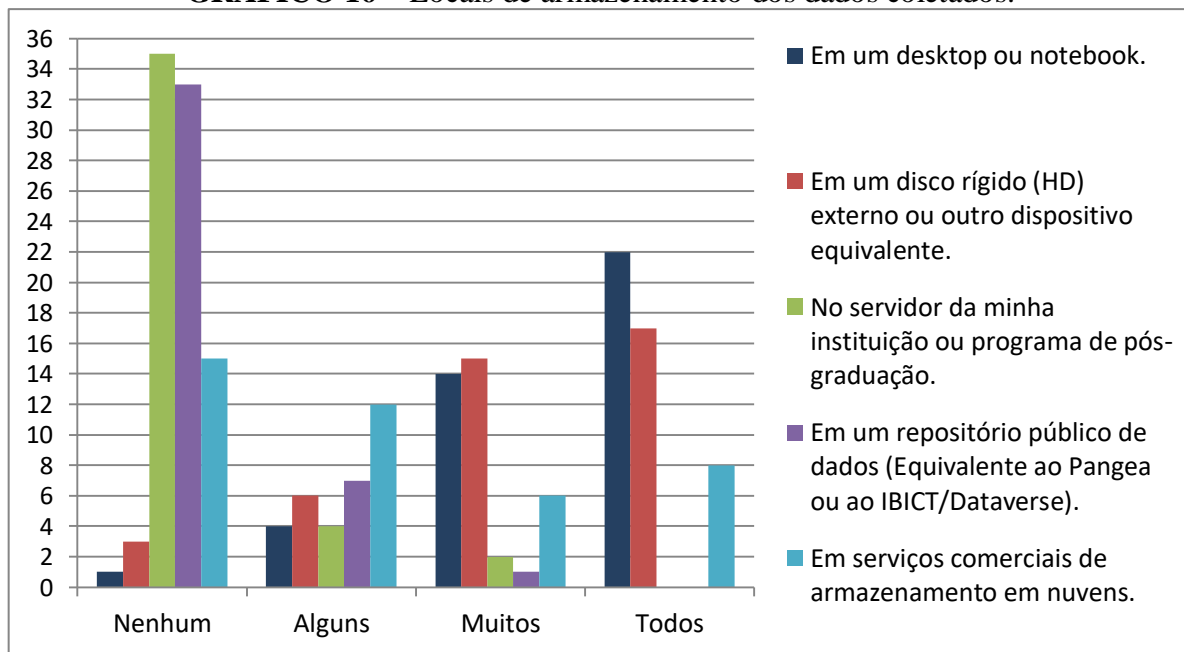
Pôde-se perceber, mediante as questões anteriores sobre o compartilhamento de dados, e esta, que os pesquisadores são a favor e acreditam no compartilhamento de dados para reuso por outros pesquisadores em outras investigações científicas, mas que a falta de incentivo, sobrecarga de trabalho, entre outros, acaba fazendo com que os dados fiquem armazenados

em um *desktop* ou *notebook*, ou ainda sejam perdidos. Além disso, 05 pesquisadores selecionaram a opção ‘outros’, e como respostas, 02 elencaram que fazem o compartilhamento de dados e por isso não selecionaram razões, outros 02 que não compartilham seus dados por não haver uma política de como se dar o reuso dos dados, e outro, argumentou que ainda está trabalhando com os dados, e caso fossem disponibilizados, impossibilitaria a publicação de artigos futuros.

Após a questão 18, 03 pesquisadores finalizaram o questionário. A partir da questão 19, até o fim do questionário, temos o total de 41 pesquisadores.

A questão 19 procurou identificar os locais de armazenamento, onde ficam os dados coletados nas pesquisas dos sujeitos desta investigação. Foram elencadas 05 opções de armazenamento, nas quais os pesquisadores deveriam informar a quantidade (nenhum, alguns, muitos, todos) de dados que são armazenados nestas. Como observado no Gráfico 16.

**GRÁFICO 16** – Locais de armazenamento dos dados coletados.



Fonte: Dados da pesquisa, 2016.

Como dito anteriormente, a maioria dos dados de pesquisa são perdidos após a publicação dos resultados em artigos científicos, ou ficam armazenados apenas nos computadores pessoais dos pesquisadores, impossibilitando o desenvolvimento de novas investigações. Analisando a questão 19, podemos perceber que tal fato é real, visto que 36 dos 41 pesquisadores investigados armazenam todos ou muitos dos seus dados em um *desktop* ou *notebook*, e que 32 armazenam todos ou muitos dos seus dados em um disco rígido externo (HD), ou dispositivo equivalente. Em contrapartida, 35 pesquisadores não disponibilizam

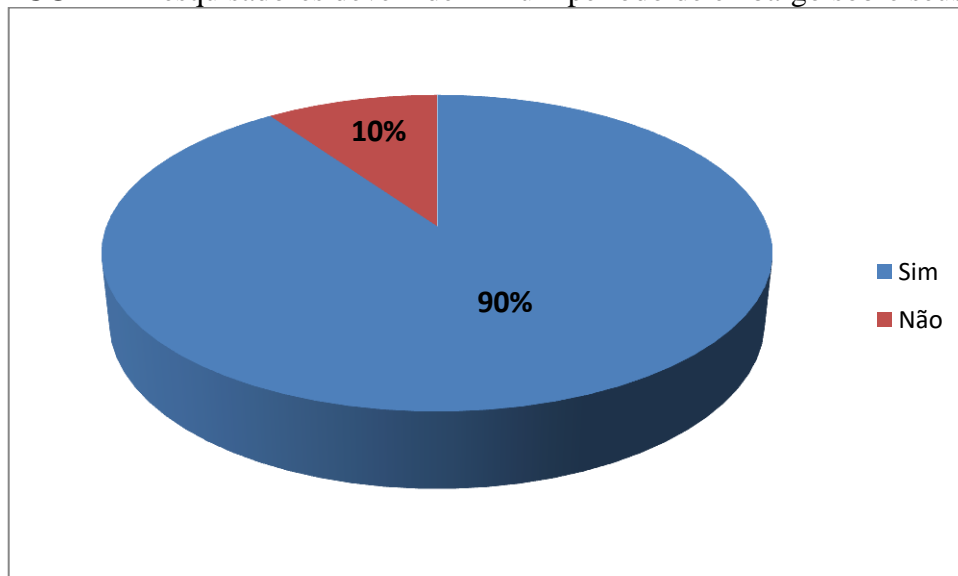


nenhum dos seus dados em servidores das suas instituições ou dos seus programas de pós-graduação.

As próximas questões investigaram se os pesquisadores devem ter autonomia para definir um período de embargo sobre os seus dados, especificando quando os mesmos poderão ser disponibilizados para o público, e qual seria o tempo necessário para isso. Visando evitar umas das razões que foram mencionadas anteriormente, a de ainda estar trabalhando em cima de dados que caso fossem logo disponibilizados, outros pesquisadores poderiam fazer análises similares, inviabilizando a publicação de artigos futuros, como o pesquisador de ID 06, colocou no Quadro 7. Além de estar em concordância com o pensamento de Frické (2015) exposto anteriormente.

Dos 41 pesquisadores, 37 concordam que o pesquisador deve poder definir um período de embargo sobre os seus dados, e de maneira oposta, 04 pesquisadores discordam da definição de um período de embargo sobre os dados por parte dos pesquisadores. Conforme apresentado no Gráfico 17.

**GRÁFICO 17** – Pesquisadores devem definir um período de embargo sobre seus dados?

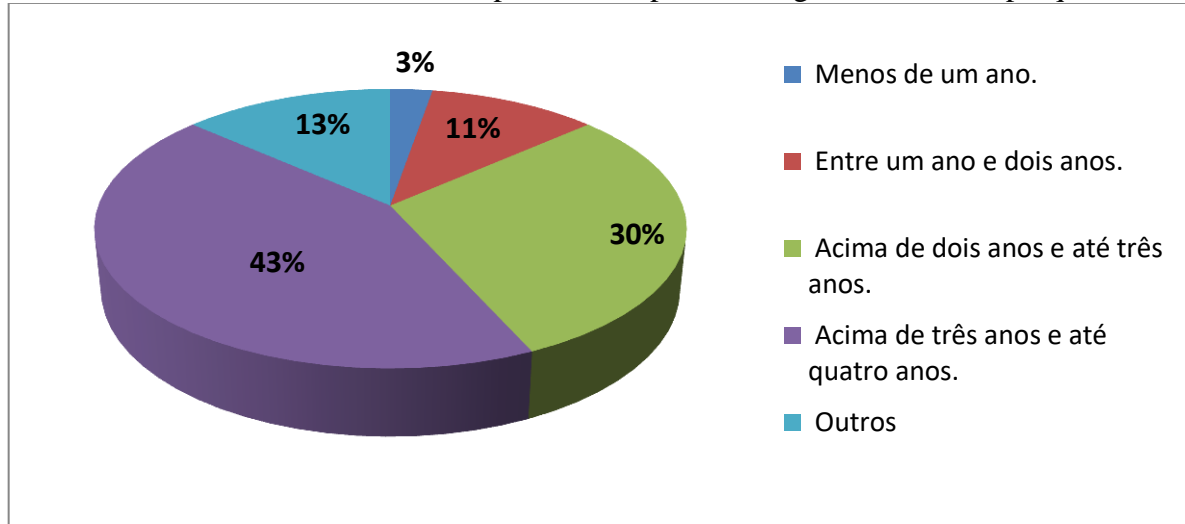


**Fonte:** Dados da pesquisa, 2016.

Em relação ao período de embargo, os 37 pesquisadores que concordaram com a definição do mesmo, foram questionados sobre qual seria um período razoável de tempo para a restrição do acesso público aos dados finais, fazendo com que o pesquisador líder da pesquisa possa elaborar os primeiros artigos científicos relacionados aos dados coletados. As opções foram selecionadas da seguinte forma: menos de um ano (01 pesquisador); entre um ano e dois anos (04 pesquisadores); acima de dois até três anos (11 pesquisadores); acima de

três até quatro anos (16 pesquisadores); e outros (05 pesquisadores). Conforme ilustrado no Gráfico 18.

**GRÁFICO 18** – Período de tempo razoável para embargo aos dados de pesquisa.



Fonte: Dados da pesquisa, 2016.

Os 05 pesquisadores que selecionaram a opção ‘outros’, conforme relacionado no Quadro 8, definiram que:

**QUADRO 8** – Período de tempo razoável para embargo aos dados de pesquisa.

Período de tempo razoável para embargo aos dados de pesquisa.	
ID	Justificativa
01	“Depende do quão sensível é o dado.”
22	“Cinco anos.”
53	“Depende de projetos individuais com os dados em questão.”
54	“A depender do dado.”
72	“Até o artigo que o utilizou for publicado.”

Fonte: Dados da pesquisa, 2016.

Com base nas duas questões anteriores, identificou-se que a 91% dos pesquisadores são a favor do embargo aos dados de pesquisa, e que este embargo deveria ser acima de três até quatro anos. Percebeu-se ainda que o embargo deve ser definido de acordo com o tipo de dado que foi coletado, ou até o artigo que o utilizou ser publicado.

A última questão do instrumento pesquisa foi de caráter aberto, na qual solicitamos aos pesquisadores que fizessem comentários pertinentes a temática, e importantes para o andamento desta pesquisa. Apenas 05 pesquisadores fizeram comentários, disponíveis no Quadro 9.

**QUADRO 9** – Comentários pertinentes à temática da pesquisa e para a pesquisa.

<b>Comentários pertinentes à temática da pesquisa e para a pesquisa.</b>	
<b>ID</b>	<b>Justificativa</b>
21	“Na minha já área existem várias revistas que exigem a publicação dos dados brutos quando da publicação de um artigo. Acredito que isso é um ótimo incentivo.”
31	“Acho que o compartilhamento de dados deveria ser fortemente incentivado pelas agências de fomento, para que os pesquisadores não se sintam donos dos dados e não deixem informações importantes perdidas em computadores pessoais.”
52	“O principal complicante para a manutenção de dados em repositórios é a dificuldade estrutural das instituições de pesquisa (material e de pessoal técnico), o que aumenta o tempo possível para que os dados sejam publicados, a partir do tempo em que foram coletados. Isso deve ser levado em conta, em minha opinião, como uma consideração inevitável e importante ao longo das análises que, creio, devam decorrer da presente pesquisa.”
63	“Duas coisas. Acho o questionário relevante, mas estranho que o mesmo seja relacionado à região Nordeste. Tal dado deveria ter um alcance nacional, qual a contribuição destes dados para uma política pública nacional?”
66	“Dar a opção "salvar e terminar agora"; incluir o número de perguntas e tempo médio para finalizar o questionário.”

**Fonte:** Dados da pesquisa, 2016.

Dentre os comentários, observamos que os pesquisadores investigados têm conhecimento da importância da temática desta pesquisa, e que a mesma deveria ser estendida ao país inteiro, e a outras áreas do conhecimento.

## **7 CONSIDERAÇÕES FINAIS**

O presente estudo cumpriu todos os objetivos específicos propostos anteriormente, para que o objetivo geral fosse alcançado.

O primeiro objetivo específico buscou traçar o perfil dos sujeitos desta pesquisa, em que se percebeu que 43% dos pesquisadores estão entre 35 e 44 anos, e que 33% possuem pouco tempo de vínculo aos programas de pós-graduação, com menos de 05 anos de atuação.

O segundo objetivo, almejou identificar as ferramentas adotadas pelos pesquisadores na realização do PGD, em que se verificou que 75% dos pesquisadores não fazem uso de ferramentas no plano de gestão de dados, e aqueles 25% que utilizam ferramenta, em sua maioria, fazem o uso da planilha Excel. Observou-se também que 86% dos pesquisadores fazem uso de ferramentas de software para a exploração, visualização ou análise de dados, e que entre as mais citadas, destacaram-se o *software* R; planilha Excel; e o *software* Statistica

No que tange a natureza da análise que os pesquisadores dão aos seus dados de pesquisa, 84% dão análise mista (qualitativa e quantitativa). E em sua maioria, os dados dos pesquisadores tem origem observacional e/ou experimental.

Em relação ao uso de Metadados, 70% dos pesquisadores não fazem uso de Metadados, e o restante, que afirmaram fazer o uso, não souberam indicar os padrões de metadados utilizados.

Constatou-se que os pesquisadores são a favor do compartilhamento de dados visto que 36 dos 44 pesquisadores investigados, afirmaram que disponibilizariam alguns dos seus dados em um repositório de dados sem restrições, mas que existe uma série de fatores que impedem tal fato, como falta de tempo, ausência de pessoal de apoio, infraestrutura técnica, entre outros.

No que corresponde às ações tomadas pelos pesquisadores no sentido de armazenamento e preservação dos dados, os pesquisadores ainda, em sua maioria, armazenam os dados oriundos de suas investigações científicas em *desktop* ou *notebook*. O que possibilita o armazenamento e preservação dos dados, mas que não incentiva o compartilhamento, uso e reuso de dados de pesquisa.

Com a realização desse estudo, concluímos que a realidade do compartilhamento de dados por meio da gestão dos mesmos é emergente. Devido ao fato de que alguns pesquisadores já utilizam ferramentas que auxiliam no plano de gestão de dados, e outros, tem imenso interesse em aumentar seus conhecimentos, em relação às práticas de gestão de dados científicos para posterior compartilhamento.

Percebeu-se que os pesquisadores dos Programas de Pós-Graduação em Ecologia do nordeste brasileiro tem conhecimento acerca do uso, da gestão, do compartilhamento de dados, e como esses são um incentivo para o crescimento da Comunicação Científica.

De acordo com o pesquisador ID 31 “o compartilhamento de dados deveria ser fortemente incentivado pelas agências de fomento, para que os pesquisadores não se sintam donos dos dados e não deixem informações importantes perdidas em computadores pessoais”, acreditamos junto a ele, que o desenvolvimento da comunicação científica está amplamente relacionado ao compartilhamento de dados, o que possibilita a geração de novas informações e conhecimentos.

## REFERÊNCIAS

- AGÊNCIA FAPESP. **Editor-chefe da Nature fala sobre a abertura da ciência**. 2013. Disponível em: <[http://agencia.fapesp.br/editorchefe\\_da\\_nature\\_fala\\_sobre\\_a\\_abertura\\_da\\_ciencia/16919/](http://agencia.fapesp.br/editorchefe_da_nature_fala_sobre_a_abertura_da_ciencia/16919/)>. Acesso em: 10 set. 2016.
- BORGMAN, C. L. **Big data, little data, no data: Scholarship in the Networked World**. London: MIT Press, 2015.
- CORRÊA, F. C. O papel dos bibliotecários na gestão de dados científicos. **Revista Digital de Biblioteconomia e Ciência da Informação**. V. 14, n. 3. Campinas. Set./Dez. 2016. p. 387-406. Disponível em: <[http://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8646333/pdf\\_1](http://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8646333/pdf_1)>. Acesso em: 07 set. 2016.
- COSTA, M. M.; CUNHA, M. B. O bibliotecário no tratamento de dados oriundos da e-science: considerações iniciais. **Perspectivas em Ciência da Informação**. V. 19, n. 3. Belo Horizonte. Jul./Set. 2014. p. 189-206. Disponível em: <<http://www.scielo.br/pdf/pci/v19n3/a10v19n3.pdf>>. Acesso em: 06 set. 2016.
- CUNHA, M. B. A biblioteca universitária na encruzilhada. **DataGramZero: Revista de Ciência da Informação**. V. 11, n.6. Rio de Janeiro. 2010. Disponível em: <[http://www.dgz.org.br/dez10/Art\\_07.htm](http://www.dgz.org.br/dez10/Art_07.htm)>. Acesso em: 06 set. 2016.
- DATAONE. **Primer on Data Management: what you always wanted know**. Disponível em: <[https://www.dataone.org/sites/all/documents/DataONE\\_BP\\_Primer\\_020212.pdf](https://www.dataone.org/sites/all/documents/DataONE_BP_Primer_020212.pdf)>. Acesso em: 09 set. 2016.
- DUDZIAK, Elisabeth. **Dados de Pesquisa agora devem ser armazenados e citados**. 2016. Disponível em: <<http://www.sibi.usp.br/?p=6189>> Acesso em: 08 set. 2016.
- FRICKÉ, M. Big Data and its Epistemology. **Journal of the Association for Information and Technology**. V.66, n. 4, abr. 2015.
- GORDON, B. Foreword. In: Hey, Anthony J. G., Stewart Tansley, and Kristin Michele Tolle. **The Fourth Paradigm: Data-intensive Scientific Discovery**. Redmond, Wash: Microsoft Research, 2009. Disponível em: <<http://digital.library.unt.edu/ark:/67531/metadc31516/&gt;>>. Acesso em: 06 set. 2016.
- Instituto Brasileiro De Informação Ciência E Tecnologia - IBICT. **Manifesto de acesso aberto a dados de pesquisa brasileira para ciência cidadã**. Disponível em:<<http://www.ibict.br/Sala-de-Imprensa/noticias/2016/ibict-lanca-manifesto-de-acesso-aberto-a-dados-da-pesquisa-brasileira-para-ciencia-cidada>>. Acesso em: 07 set. 2016.
- MEDEIROS, J. S. CAREGNATO, S. E. Compartilhamento de dados e e-Science: explorando um novo conceito para a comunicação científica. **Liinc em revista**. V. 8, n. 2, set. Rio de Janeiro. 2012, p. 311 – 322. Disponível em: <[https://chasqueweb.ufrgs.br/~jacksonmedeiros/pubs/2012\\_escience\\_liinc.pdf](https://chasqueweb.ufrgs.br/~jacksonmedeiros/pubs/2012_escience_liinc.pdf)>. Acesso em: 11 set. 2016.

MINISTÉRIO DA JUSTIÇA E CIDADANIA. **MJ apresenta nova versão do Anteprojeto de Lei de Proteção de Dados Pessoais**. Disponível em:

<<http://www.justica.gov.br/noticias/mj-apresenta-nova-versao-do-anteprojeto-de-lei-de-protecao-de-dados-pessoais>>. Acesso em: 08 set. 2016.

NATIONAL SCIENCE BOARD - NSB. **Long-lived digital data collections: enabling research and education in the 21st century**. National Science Foundation, Sept. 2005.

Disponível em: <<https://www.nsf.gov/geo/geo-data-policies/nsb-0540-1.pdf>>. Acesso em: 09 set. 2016.

OECD PRINCIPLES AND GUIDELINES FOR ACCESS TO RESEARCH DATA FROM PUBLIC FUNDING. **Paris**: Organization for Economic Co-operation and Development, 2007. Disponível em: <<http://www.oecd.org/sti/sci-tech/38500813.pdf>>. Acesso em: 09 set. 2016.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE – PPBIO. **Sobre**. Disponível em: <<https://ppbio.inpa.gov.br/en/About>>. Acesso em: 07 set. 2016.

SALES, L. F. et al. Publicações ampliadas: um novo modelo de publicação acadêmica para o ambiente de e-Science. In: XIV Encontro Nacional de Pesquisa em Ciência da Informação, 2013, Florianópolis. **Anais...** Florianópolis: XIV ENANCIB, 2013. Disponível em: <<http://eprints.rclis.org/20666/1/APRESENTA%C3%87%C3%83O%20-%20ENANCIB%202013.pdf>>. Acesso em: 05 set. 2016

SALES, Luana Farias. SAYÃO, Luís Fernando. Há futuro para as bibliotecas de pesquisa no ambiente de *e-Science*? **Informação & Tecnologia (ITEC)**. V.2,n.1, p. 30-52, jan./jul., 2015. Disponível em: <<http://periodicos.ufpb.br/ojs/index.php/itec/article/view/26029/14677>>. Acesso em: 05 set. 2016.

SANTANA, R. C. G. Ciclo de vida dos dados e o papel da Ciência da Informação, 2013. In: XIV Encontro Nacional de Pesquisa em Ciência da Informação, 2013, Florianópolis. **Anais...** Florianópolis: XIV ENANCIB, 2013. Disponível em: <<http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib/paper/viewFile/284/319>>. Acesso em: 12 set. 2016.

SAYÃO, L. F. SALES, L. F. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**. V. 8, n. 2. Rio de Janeiro . 2014. p. 76-92. Disponível em: <<http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/611/1252>>. Acesso em: 05 set. 2016.

\_\_\_\_\_. Curadoria geral: um novo patamar para a preservação de dados digitais de pesquisa. **Informação & Sociedade**. V. 22, n. 3. João Pessoa. Set./Dez. 2012. p. 179-191. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/viewFile/12224/8586>>. Acesso em: 10 set. 2016.

\_\_\_\_\_. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**.

Comissão Nacional de Energia Nuclear: Rio de Janeiro. 2015. Disponível em: <[http://carpedien.ien.gov.br/bitstream/ien/1624/1/GUIA\\_DE\\_DADOS\\_DE\\_PESQUISA.pdf](http://carpedien.ien.gov.br/bitstream/ien/1624/1/GUIA_DE_DADOS_DE_PESQUISA.pdf)>. Acesso em: 11 set. 2016.

SETZER, V.W. **Dado, informação, conhecimento e competência**. 2015. Disponível em: <<https://www.ime.usp.br/~vwsetzer/dado-info.html>>. Acesso em: 08 set. 2016.

STRASSER, C. et al. **Primer on Data Management: What you always wanted to know**. California: CDL, 2012. Disponível em: <<http://escholarship.org/uc/item/7tf5q7n3#page-1>>. Acesso em: 12 set. 2016.

TENOPIR, C.; BIRCH, B.; ALLARD, S. **Academic libraries and research data services: current practices and plans for the future**. Association of College & Research Libraries: Chicago, 2012. Disponível em: <[http://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/whitepapers/Tenopir\\_Birch\\_Allard.pdf](http://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf)>. Acesso em: 13 set. 2016.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. São Paulo: Atlas, 2006.



## APÊNDICES

### APÊNDICE A

Prezado(a) Pesquisador(a), gostaríamos de convidá-lo(a) para participar da pesquisa intitulada **Desvendando as percepções e práticas adotadas pelos pesquisadores dos Programas de Pós-Graduação em Ecologia do Nordeste Brasileiro na gestão dos dados científicos**. Esta investigação é conduzida pelos pesquisadores do Grupo de Pesquisa Web Representação do Conhecimento e Ontologias – WRCO, cadastrado no diretório de grupos de pesquisas do CNPq.

- Suas respostas às questões devem refletir a sua experiência na gestão de dados científicos.
- Solicitamos a sua colaboração para responder este questionário, como também sua autorização para apresentar os resultados deste estudo em eventos da área e divulgar os dados coletados. Sua identificação será mantida em sigilo em todas as etapas do processo investigativo.
- Esclarecemos que sua participação no estudo é voluntária e, portanto, o(a) senhor(a) não é obrigado(a) a fornecer as informações e/ou colaborar com as atividades solicitadas pelo pesquisador(a).

Agradecemos antecipadamente a sua preciosa colaboração.

Instrumento de pesquisa baseado no questionário intitulado “*Research data management assessment: Building understanding of your tools, data management, and training needs*” desenvolvido pela USGS.

#### **Seção Introdutória**

[01] O processo de geração do conhecimento científico e acadêmico utiliza o escopo da comunicação científica formal como um dos meios para a disseminação, uso e acesso à pesquisa científica. Este escopo está se reconfigurando, abrindo novos espaços de interação e conectividade mediados pelas Tecnologias Digitais de Informação e Comunicação (TDICs). Os dados de pesquisa estão assumindo um novo papel no processo de investigação científica. Tradicionalmente os dados gerados ao longo do processo de investigação científica são descartados ou esquecidos após a conclusão da pesquisa e da consequente publicação dos artigos científicos resultantes do trabalho. Este cenário está mudando, os dados demonstram ser uma fonte primária de recursos científicos que pode ser utilizada por outros pesquisadores na análise de fenômenos distintos e até para replicar um determinado experimento para atestar a sua validade. Desta forma, muitos pesquisadores das mais variadas áreas do conhecimento, em diversos países e instituições estão dando especial atenção para a preservação e reutilização dos dados de suas pesquisas. Ao invés de um simples descarte, estes dados estão sendo preservados e compartilhados através de repositórios especificamente construídos para este fim (BORGMAN, 2015).

**Com base nesta breve explanação, você gostaria de responder alguns questionamentos pertinentes à temática introduzida?**

- Sim  
 Não

BORGMAN, C. L. **Big data, little data, no data**; Scholarship in the Networked World. London: MIT Press, 2015. (Ebook).

### **Seção de Identificação**

**[02] Qual a sua faixa etária?**

- Menos de 25 anos  
 De 25 a 34 anos  
 De 35 a 44 anos  
 De 45 a 54 anos  
 De 55 a 64 anos  
 65 anos ou mais

**[03] Há quanto tempo você atua na pós-graduação?**

- Menos de 5 anos  
 Entre 5 e 9 anos  
 De 10 a 14 anos  
 De 15 a 19 anos  
 De 20 a 24 anos  
 De 25 a 29 anos  
 30 anos ou mais

### **Sobre dados**

**[04] Você faz a gestão dos dados científicos coletados nas suas pesquisas?**

- Sim  
 Não

**[05] Por quê?**

Só responder essa pergunta sob as seguintes condições:

A resposta foi 'Não' na questão '[04]' (Você faz a gestão dos dados científicos coletados nas suas pesquisas?)

Por favor, coloque sua resposta aqui:

**[06] Como você classificaria a natureza de sua análise atribuída aos dados de pesquisa?**

- Qualitativa  
 Quantitativa  
 Qualitativa e Quantitativa

**[07] Os dados de pesquisa observacionais são aqueles obtidos por meio de observações associadas a lugares e tempos específicos, configurando-se como registros históricos não podendo ser coletados uma segunda vez. Os dados de pesquisa computacionais são frutos da execução de modelos computacionais ou de simulações. Por sua vez, os dados de pesquisa**

experimentais, são oriundos de situações controladas, por meio de experimentos. (SALES e SAYÃO, 2013).

**Com relação à origem dos dados usados nas suas pesquisas, indique se os mesmos são *observacionais, computacionais, experimentais*, ou qualquer combinação entre eles.**

- Observacionais  
 Computacionais  
 Experimentais

SALES, L. F. SAYÃO, L. F. SOUZA, R. F. Publicações ampliadas: um novo modelo de publicação acadêmica para o ambiente de e-science., 2013 . In: **XIV Encontro Nacional de Pesquisa em Ciência da Informação** (ENANCIB 2013), Florianópolis, Brasil, 2013, oct 29 - 2013, nov 01. [Apresnetação]

**[08] Você desenvolve algum plano de gestão de dados ao longo da condução de sua pesquisa?**

- Sim  
 Não

**[09] Por quê?**

Só responder essa pergunta sob as seguintes condições:

A resposta foi 'Não' na questão '[08]' (Você desenvolve algum plano de gestão de dados ao longo da condução de sua pesquisa? )

Por favor, coloque sua resposta aqui:

**[10] Você faz uso de alguma ferramenta automatizada para auxiliar a criação do plano de gestão de dados?**

- Sim  
 Não

**[11] Indique qual(is) a(s) ferramenta(s) você faz uso para criação de planos de gestão de dados.**

Só responder essa pergunta sob as seguintes condições:

A resposta foi 'Sim' na questão '[10]' (Você faz uso de alguma ferramenta automatizada para auxiliar a criação do plano de gestão de dados?)

Por favor, coloque sua resposta aqui:

**[12] Por favor, indique o nível de concordância com cada declaração.**

	Discordo fortemente	Discordo	Não discordo nem concordo	Concordo	Concordo fortemente
Estou confiante que os planos de	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

gestão de dados são essenciais para o acesso a longo prazo aos dados.					
Os planos de gestão de dados devem ser consultados regularmente, revisados e atualizados ao longo da vida da pesquisa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**[13] Por favor, responda se faz o uso de metadados, em caso afirmativo, indique quais padrões de metadados você utiliza para descrever seus dados.**

- Uso metadados.  
 Não uso metadados.

Comente aqui sua escolha:

**[14] Você utiliza alguma ferramenta de *software* que possibilite a exploração, visualização ou análise de dados?**

- Sim  
 Não

**[15] Indique qual(is) é (são) essa(s) ferramenta(s) de *software*.**

Só responder essa pergunta sob as seguintes condições:

A resposta foi 'Sim' na questão '[14]' (Você utiliza alguma ferramenta de software que possibilite a exploração, visualização ou análise de dados?)

**[16] O seguinte grupo de declarações abordam o compartilhamento de dados. Por favor, indique o quanto concorda com cada declaração.**

	Discordo fortemente	Discordo	Não discordo nem concordo	Concordo	Concordo fortemente
Acredito que o meu programa de pós-graduação estaria disposto a colocar pelo menos <b>ALGUNS</b> dos nossos dados não	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

sensíveis / não-proprietários em um repositório central de dados sem restrições de acesso.					
Acredito que o meu programa de pós-graduação estaria disposto a colocar TODOS os nossos dados não sensíveis / não-proprietários em um repositório central de dados sem restrições de acesso.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acredito que o meu programa de pós-graduação estaria mais inclinado a tornar nossos dados não sensíveis / não-proprietários se eu pudesse colocar condições de acesso aos mesmos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acredito que é aceitável criar novos conjuntos de dados a partir dos nossos dados (reuso).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[17] Com relação ao compartilhamento de dados. Por favor, indique o quanto você concorda com cada declaração.

	Discordo fortemente	Discordo	Não discordo nem concordo	Concordo	Concordo fortemente
Usaria conjuntos de dados de outros pesquisadores, se esses conjuntos de dados estivessem facilmente acessíveis.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estou disposto a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

colocar pelo menos ALGUNS dos meus dados em um repositório central de dados sem restrições.					
Estou disposto a colocar TODOS os meus dados em um repositório central de dados sem restrições.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
É correto criar novos conjuntos de dados a partir de dados compartilhados.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
É importante que meus dados sejam citados quando utilizados por outros pesquisadores.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Normalmente não preciso usar conjuntos de dados de outros pesquisadores.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**[18] Se os dados que você coletou ou faz a gestão não foram compartilhados, indique qual a razão deste fato?**

- Falta de financiamento.
- Falta de infraestrutura técnica.
- Ausência de clareza no que diz respeito aos padrões de dados.
- Falta de tempo.
- Ausência de pessoal de apoio.
- As pessoas não precisam de dados.
- O pesquisador líder não quer que os dados sejam disponibilizados publicamente.
- Inexistência de incentivos para que os dados sejam disponibilizados.
- Não existem exigências para que os dados sejam publicados.
- Não tenho certeza onde disponibilizar dados para o acesso público.
- Os dados estão restritos para disponibilização em virtude de segurança, confidencialidade ou questões de propriedade intelectual.
- Eu nunca disponibilizei dados de nenhuma pesquisa.
- O patrocinador da pesquisa não faz exigência para que os dados sejam disponibilizados.
- Outros: \_\_\_\_\_ -

**[19] Com relação aos dados coletados em sua pesquisa, onde os mesmos são armazenados?**

	Nenhum	Alguns	Muitos	Todos
Em um <i>desktop</i> ou <i>notebook</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Em um disco rígido (HD) externo ou outro dispositivo equivalente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No servidor da minha instituição ou programa de pós-graduação.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Em um repositório público de dados (Equivalente ao Pangea ou ao IBICT/Dataverse).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Em serviços comerciais de armazenamento em nuvens.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**[20] Os pesquisadores devem ter autonomia para definir um período de embargo sobre seus dados especificando quando os mesmos poderão ser disponibilizados para o público?**

- Sim  
 Não

**[21] Na sua opinião, qual seria um período razoável de tempo para o embargo de dados (isto é, restringir o acesso público aos dados final) de forma que o pesquisador líder da pesquisa possa elaborar os primeiros artigos científicos relacionados com os dados finais coletados.**

Só responder essa pergunta sob as seguintes condições:

A resposta foi 'Sim' na questão '[20]' (Os pesquisadores devem ter autonomia para definir um período de embargo sobre seus dados especificando quando os mesmos poderão ser disponibilizados para o público?)

- Menos de um ano.  
 Entre um ano e dois anos.  
 Acima de dois anos e até três anos.  
 Acima de três anos e até quatro anos.  
 Outros: \_\_\_\_\_

**[22] Caso você queira fazer qualquer comentário que entenda ser importante para o andamento desta pesquisa, por favor, utilize o espaço a seguir.**