

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO**

MARINA RODRIGUES MARTINS

**PERSONAS: OS PERFIS DO PÚBLICO USUÁRIO DAS CIÊNCIAS BIOLÓGICAS -
ECOLOGIA PARA INICIATIVAS EM DEFESA DO MOVIMENTO BRASILEIRO
PARA O ACESSO ABERTO AOS DADOS DE PESQUISA**

PORTO ALEGRE
2023

MARINA RODRIGUES MARTINS

**PERSONAS: OS PERFIS DO PÚBLICO USUÁRIO DAS CIÊNCIAS BIOLÓGICAS -
ECOLOGIA PARA INICIATIVAS EM DEFESA DO MOVIMENTO BRASILEIRO
PARA O ACESSO ABERTO AOS DADOS DE PESQUISA**

Tese apresentada como requisito parcial para
obtenção do título de Doutora em Comunicação
pela Universidade Federal do Rio Grande do Sul.

Orientadora: Profa. Dra. Samile Andréa de
Souza Vanz.

PORTO ALEGRE
2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Andre Bulhões Mendes

Vice-Reitora: Profa. Dra. Patricia Pranke

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Profa. Dra. Ana Maria Mielniczuk de Moura

Vice-Diretora: Profa. Dra. Vera Regina Schmitz

PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO

Coordenadora: Profa. Dra. Elisa Reinhardt Piedras

Coordenadora substituta: Profa. Dra. Samile Andrea de Souza Vanz

CIP - Catalogação na Publicação

Martins, Marina Rodrigues
Personas: os perfis do público usuário das Ciências
Biológicas - ecologia para iniciativas em defesa do
movimento brasileiro para o acesso aberto aos dados de
pesquisa / Marina Rodrigues Martins. -- 2023.
216 f.
Orientador: Samile Andréa de Souza Vanz.

Tese (Doutorado) -- Universidade Federal do Rio
Grande do Sul, Faculdade de Biblioteconomia e
Comunicação, Programa de Pós-Graduação em Comunicação,
Porto Alegre, BR-RS, 2023.

1. Compartilhamento de dados. 2. Dados abertos de
pesquisa. 3. Ecologia. 4. Personas. 5. Relações
Públicas. I. Vanz, Samile Andréa de Souza, orient.
II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os
dados fornecidos pelo(a) autor(a).

Programa de Pós-Graduação em Comunicação

Rua Ramiro Barcelos, 2705

Bairro: Santana – Porto Alegre, RS

CEP: 90.035-007

Campus Saúde

Universidade Federal do Rio Grande do Sul

Telefone: (51) 3308-5116

E-mail: ppgcom@ufrgs.br

MARINA RODRIGUES MARTINS

PERSONAS: OS PERFIS DO PÚBLICO USUÁRIO DAS CIÊNCIAS BIOLÓGICAS -
ECOLOGIA PARA INICIATIVAS EM DEFESA DO MOVIMENTO BRASILEIRO
PARA O ACESSO ABERTO AOS DADOS DE PESQUISA

Tese apresentada como requisito parcial para
obtenção do título de Doutora em Comunicação
pela Universidade Federal do Rio Grande do Sul.

Aprovada em: 23 de janeiro de 2023

BANCA EXAMINADORA

Profa. Dra. Samile Andréa de Souza Vanz (Orientadora)
Universidade Federal do Rio Grande do Sul (UFRGS)

Profa. Dra. Carolina Felicíssimo (Examinadora)
Rede Nacional de Ensino e Pesquisa (RNP)

Profa. Dra. Taís Flores da Motta (Examinadora)
Universidade do Vale do Rio dos Sinos (UNISINOS)

Profa. Dra. Ana Karin Nunes (Examinadora)
Universidade Federal do Rio Grande do Sul (UFRGS)

Profa. Dra. Sônia Caregnato (Examinadora)
Universidade Federal do Rio Grande do Sul (UFRGS)

Profa. Dra. Caterina Marta Groposo Pavão (Suplente)
Universidade Federal do Rio Grande do Sul (UFRGS)

AGRADECIMENTOS

Esta tese é dedicada a todas as mulheres que fizeram parte de minha trajetória. Ao lembrar-me de como cheguei até aqui e tentar listar cada pessoa que contribuiu nesse caminho, pude perceber como o círculo feminino que me rodeia influenciou de modo positivo. Eu poderia elencar uma série de nomes, mas não é esse o caso. Afinal, temos uma tese para ler.

O que eu posso afirmar é que sem vocês eu não estaria vivendo essa conquista. Cada uma, com sua pequena semente fez florescer uma floresta. Essa floresta é cheia de perguntas, anseios, lutas, curiosidades, percepções e uma vontade tamanha de evoluir. A pesquisa é o caminho que encontrei para isso. Vocês me ensinaram que por meio dela eu posso aprender, ensinar, compartilhar, provocar a mente de outros indivíduos e isso me encantou.

E, na busca de finalizar mais um percurso de minha vida, eu, mais uma vez, conto com uma banca examinadora formada por mulheres. Pesquisadoras, professoras, doutoras, profissionais que lutam diariamente pelo ensino, pesquisa e extensão. Que privilégio eu tenho, muito obrigada!

“A utopia está lá no horizonte. Me aproximo dois passos, ela se afasta dois passos. Caminho dez passos e o horizonte corre dez passos. Por mais que eu caminhe, jamais alcançarei. Para que serve a utopia? Serve para isso: para que eu não deixe de caminhar.”
Fernando Birri, citado por Eduardo Galeano in *Las palabras andantes*

RESUMO

Este estudo teve como objetivo compreender quem são as personas que representam o público usuário das Ciências Biológicas - ecologia, para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa. Os procedimentos metodológicos uniram técnicas da metodologia de construção de personas e da atividade de identificação e mapeamento de públicos em relações públicas (RP), fundamentando-se em pesquisa bibliográfica, reuso dos dados da *survey* realizada, em 2018, pelo Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira (GT – RDP Brasil) e entrevistas semiestruturadas. O corpus entrevistado contemplou indivíduos provenientes da *survey* e pessoas ligadas a grupos ou instituições que promovem iniciativas em prol do acesso aberto. A análise de conteúdo possibilitou identificar dois arquétipos. O primeiro é a Persona Ecologia (PE), quem produz, compartilha e reusa dados de pesquisa ecológicos. O segundo é a Persona Suporte (PS), quem visa auxiliar e garantir o processo de compartilhamento e reuso. As personas são comunicadas por infográficos que unem elementos visuais a textos reduzidos para transmitir as mensagens. Os resultados apontam que a Persona Ecologia é um público de *opinião*, *consciente*, de *lealdade* e de *voz*. Ela possui interesses nos serviços prestados pelas iniciativas, é afetada por eles e é ativa na prática do compartilhamento e reuso de dados ecológicos. A Persona Suporte se desenhou como um público *ativo*, de *voz*, de *decisão*, de *comportamento* e de *consulta*. Ela possui poder de determinar algo sobre a existência das mesmas já que executa ações que cumprem com suas finalidades, influenciando políticas e resultados. Conclui-se que é essencial compreender a PE como um público de *consulta*, para que suas opiniões sejam consideradas na tomada de decisão sobre ações que visam o acesso aberto aos dados de pesquisa ecológicos no Brasil. E que, a PS varia conforme potencial de poder que pode exercer, a fim de estimular ou desestimular ações em prol do acesso aberto. Por fim, constata-se que existe uma trajetória de ações concretas das personas em prol das pautas da ciência aberta, da difusão e da disseminação científica. No entanto, as iniciativas e os meios utilizados para o processo de compartilhamento e reuso priorizam um público de especialistas, o que limita a divulgação científica.

Palavras-chave: compartilhamento de dados; dados abertos de pesquisa; ecologia; mapeamento de públicos; personas; relações públicas.

ABSTRACT

The present study aimed to broaden the understanding of who the personas that represent the public user of Biological sciences - ecology are, for initiatives in defense of the Brazilian movement for open access to research data. The methodological procedures combined techniques of the persona method and the identification and mapping activity of public relations (PR), based on bibliographic research, reuse of the data from the survey conducted in 2018 by the Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira (GT – RDP Brasil) and semi-structured interviews. The corpus derived from two interviewed groups: individuals from the survey and people linked to groups or institutions that promote initiatives for open access. Content analysis made it possible to identify two archetypes. The first is the Persona Ecologia (PE), who produces, shares and reuses ecological research data. The second is the Persona Suporte (PS), which aims to assist and ensure the process of sharing and reusing. The personas are communicated by infographics that connect visual elements to reduced texts to convey messages. The results indicate that the Persona Ecologia is a public of opinion, conscious, loyalty and voice. It has interests in the services provided by the initiatives, is affected by them and is active in the practice of sharing and reusing ecological data. The Persona Suporte has presented itself as an active public, of voice, decision, behavior, and consulting. It has the power to determine something about their existence since it performs actions that fulfill its purposes, influencing policies and results. It is concluded that it is essential to understand PE as a public of consultation, so that its opinions are considered in decision-making on actions aimed at open access to ecological research data in Brazil. And that, the PS varies according to the potential of power that it can exert, in order to stimulate or discourage actions in favor of open access. Finally, it is observed that there is a trajectory of concrete actions of the personas in favor of the agendas of open science and scientific diffusion and dissemination. However, the initiatives and means used for the process of sharing and reusing prioritize a public of experts, which limits scientific dissemination.

Keywords: *data sharing; open research data; ecology; mapping publics; personas; public relations.*

LISTA DE FIGURAS

Figura 1 – Guarda chuva da ciência aberta	28
Figura 2 – Coletivos envolvidos na ciência aberta	30
Figura 3 – Ciclo de Vida de Dados	45
Figura 4 – Índice para o documento básico da persona	66
Figura 5 – As cinco fases do Ciclo de Vida das Personas	68
Figura 6 – Estrutura de dados para a comunicação dos esqueletos	69
Figura 7 – Esqueleto concluído de Oden, o "usuário ocasional" da <i>Microsoft Corporation</i>	69
Figura 8 – Exemplo de comunicação da persona construída para a Biblioteca Carnegie apresentada através de um Quadro de História e Narrativa	72
Figura 9 – Persona criada para a organização G4K	72
Figura 10 – Persona criada para um portal africano	73
Figura 11 – Formato de apresentação das personas	77
Figura 12 – Estrutura do Mapa do Cenário Organizacional	81
Figura 13 – Estrutura do Desenho da Persona	82
Figura 14 – Estrutura do Mapa de Relacionamento	83
Figura 15 – Resumo do processo de análise de conteúdo e suas fases	109
Figura 16 – Exemplo de interfaces de dois mecanismos utilizados pela PE	127
Figura 17 – Infográfico perfil parte 1 – Persona Ecologia	161
Figura 18 – Infográfico perfil parte 2 – Persona Ecologia	162
Figura 19 – Infográfico prática da pesquisa – Persona Ecologia	163
Figura 20 – Infográfico prática do compartilhamento e reuso – Persona Ecologia	164
Figura 21 – Infográfico meios de compartilhamento e reuso – Persona Ecologia	165
Figura 22 – Infográfico tipo de dado – Persona Ecologia	166
Figura 23 – Infográfico motivações para compartilhamento e reuso – Persona Ecologia	167
Figura 24 – Infográfico medos/insatisfações – Persona Ecologia	168
Figura 25 – Infográfico expectativas/necessidades – Persona Ecologia	169
Figura 26 – Infográfico perfil – Persona Suporte	170
Figura 27 – Infográfico meios de compartilhamento e reuso – Persona Suporte	171
Figura 28 – Infográfico comparativo: medos/insatisfações + expectativas/necessidades – Persona Suporte	172
Figura 29 – Infográfico visão PS sobre PE	173
Figura 30 – Infográfico ambiente relacional Persona Ecologia	174
Figura 31 – Quadro comparativo da Persona Ecologia e da Persona Suporte – enquanto público usuário de iniciativas brasileiras de acesso aberto aos dados de pesquisa	175

LISTA DE QUADROS

Quadro 1 – Cinco escolas de pensamento da Ciência Aberta	29
Quadro 2 – Síntese do corpus investigado	104
Quadro 3 – Esqueleto Persona Ecologia	106
Quadro 4 – Exemplo de planilha utilizada na fase de pré-análise para a estruturação dos dados coletados nas entrevistas semiestruturadas	110
Quadro 5 – Exemplo da planilha para exploração do material	111
Quadro 6 – Categorias finais de análise de conteúdo	112
Quadro 7 – Exemplo de planilha organizada para a análise de conteúdo	113

LISTA DE ABREVIATURAS E SIGLAS

AC – Análise de Conteúdo

ALA – *Atlas of Living Australia*

BDC – *Biodiversity Data Cleaning*

BDiG-PELD – Biblioteca Digital Georreferenciada reconhecida pela sigla

BIREME – Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde

BON – *Biodiversity Observation Network*

Brapci – Base de Dados em Ciência da Informação

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CPD – Centro de Processamento de Dados

CEDAP – Centro de Documentação e Acervo Digital da Pesquisa

CIDACS – Centro de Integração de Dados e Conhecimentos para Saúde

CIGA – Comitê Interministerial Governo Aberto

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CSV – *Comma-separated values*

CRIA – Centro de Referência em Informação Ambiental

DataONE – *Data Observation Network for Earth*

DOI – *Digital Object Identifier*

EML – *Ecological Metadata Language*

FAIR – *Findable, Accessible, Interoperable, Reusable*

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

FBDS – Fundação Brasileira de Desenvolvimento Sustentável

FIOCRUZ – Fundação Oswaldo Cruz

FURG – Universidade Federal do Rio Grande

GBIF – *Biodiversity Information Facility*

GT – RDP Brasil – Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia

IBP – *International Biological Program*

ICMBio – Instituto Chico Mendes de Conservação da Biodiversidade

ICSU – *Interdisciplinary Body of The International Council for Science*

ILTER – *International Long Term Ecological Research*

INCT-CENBAM – Centro de Estudos Integrados da Biodiversidade Amazônica

INPA – Instituto Nacional de Pesquisas da Amazônia

IODE – *International Oceanographic Data and Information Exchange*

JSON – *Java script Object Notation*

KNB – *Knowledge Network for Biocomplexity*

LBA – Programa de Grande Escala da Biosfera-Atmosfera na Amazônia

LD – *Linked Data*

LISA – *Library Information Science Abstracts*

LTER – *U.S. Long-Term Ecological Research*

MapBiomass – Projeto de Mapeamento Anual do Uso e Cobertura da Terra no Brasil

NCEAS – *National Center for Ecological Analysis and Synthesis*

NEON – *National Ecological Observatory Network*

NSF – *U. S. National Science Foundation*

OAI – *Open Archives Initiative*

OAI-PMH – *Open Archives Initiative Protocol for Metadata Harvesting*

OBIS – *Ocean Biogeographic Information System*

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

OF – *Open Format*

OGP – *Open Government Partnership*

OOI – *Ocean Observatories Initiative*

PE – Persona Ecologia

PELD – Programa Pesquisas Ecológicas de Longa Duração

PPBio – Programa de Pesquisa em Biodiversidade

PS – Persona Suporte

RD – Repositórios Digitais

RDA – *Research Data Alliance*

RNP – Rede Nacional de Ensino e Pesquisa

RP – Relações Públicas

SEEG/OC – Sistema de Estimativas de Emissões de Gases de Efeito Estufa do Observatório do Clima

SiBBR – Sistema de Informação sobre a Biodiversidade Brasileira

SI – Sistema de Informação

SIGAA – Sistema Integrado de Gestão de Atividades Acadêmicas

STM – *Text and Data Mining for non-commercial scientific research*

TCLE – Termos de Consentimento Livre e Esclarecido

TI – Tecnologia da Informação

UCD – *User-centered Design*

UFRGS – Universidade Federal do Rio Grande do Sul

UNESCO – Organização das Nações Unidas para a Educação, a Ciência e a Cultura

UC – Unidade de Contexto

UR – Unidade de Registro

URI – *Uniform Resource Identifier*

XLS – *Extensible Style Language*

WDS – *World Data System*

WorldClim – *Global Climate Data*

WoS – *Web of Science*

SUMÁRIO

1 INTRODUÇÃO	15
1.1 JUSTIFICATIVA	21
1.2 OBJETIVOS	26
1.2.1 Objetivo Geral	26
1.2.2 Objetivos Específicos	26
2 REVISÃO DE LITERATURA	27
2.1 CIÊNCIA ABERTA (<i>OPEN SCIENCE</i>) E DADOS DE PESQUISA	27
2.1.1 Dados de pesquisa na ecologia	39
2.1.2 Mecanismos de compartilhamento	43
2.2 USUÁRIOS COMO PÚBLICO E SEU ENTENDIMENTO COMO PERSONA	54
2.3 PERSONAS	57
2.3.1 Construção das Personas: mapeamento de estudos e métodos	64
2.4 PÚBLICOS EM RELAÇÕES PÚBLICAS: CONCEITOS, CLASSIFICAÇÕES E TÉCNICAS DE PESQUISA	84
2.5 PERFIL DOS USUÁRIOS DE SERVIÇOS DE ACESSO ABERTO AOS DADOS DE PESQUISA NO QUE TANGE A ECOLOGIA: A PERSPECTIVA TEÓRICA BRASILEIRA	95
3 PROCEDIMENTOS METODOLÓGICOS	101
3.1 PESQUISA BIBLIOGRÁFICA	101
3.2 CORPUS ENTREVISTADO, TÉCNICA E INSTRUMENTO DE COLETA DE DADOS	103
3.3 ANÁLISE DE CONTEÚDO E MOVIMENTO DE CATEGORIZAÇÃO	108
4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS: A PERSONA ECOLOGIA E A PERSONA SUPORTE	115
4.1 PERSONA ECOLOGIA (PE)	115
4.1.1 Perfil PE	116
4.1.2 Prática: da Pesquisa PE	119
4.1.3 Prática: de Compartilhamento e Reuso PE	121
4.1.4 Prática: Meios de Compartilhamento e Reuso PE	125
4.1.5 Tipo de dado produzido, compartilhado e reusado PE	130
4.1.6 Competência para compartilhamento e reuso PE	133
4.1.7 Motivações para compartilhamento e reuso PE	135

4.1.8 Medos/Insatisfações PE	138
4.1.9 Expectativas/Necessidades PE	141
4.2 PERSONA SUPORTE (PS)	145
4.2.1 Perfil PS	147
4.2.2 Prática: meios de compartilhamento PS	151
4.2.3 Medos/Insatisfações PS	153
4.2.4 Expectativas/Necessidades PS	155
4.2.5 Persona Ecologia na visão da Persona Suporte	157
5 COMUNICAÇÃO DA PERSONA ECOLOGIA E DA PERSONA SUPORTE	160
5.1 PERSONA ECOLOGIA (PE)	161
5.2 PERSONA SUPORTE (PS)	170
5.3 PERSONA ECOLOGIA E PERSONA SUPORTE ENQUANTO PÚBLICO	174
6 CONCLUSÕES	177
REFERÊNCIAS	187
APÊNDICE A – Roteiro de Entrevista Semiestruturada Partes Interessadas	204
APÊNDICE B – Termo de Consentimento Livre e Esclarecido Partes Interessadas	208
APÊNDICE C – Roteiro de Entrevista Semiestruturada Usuários	211
APÊNDICE D – Termo de Consentimento Livre e Esclarecido Usuários	214

1 INTRODUÇÃO

Esta pesquisa se incorpora à Linha 1: Redes Sociais, Interações e Sociabilidades, do Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul (UFRGS) e relaciona as temáticas: acesso aberto aos dados de pesquisa, dados abertos de pesquisa, dados ecológicos, estudo de usuários, públicos em relações públicas e personas. Como estímulo central tem o mapeamento, a construção e a compreensão dos perfis que compõem um dos públicos para iniciativas brasileiras que visam o acesso aberto aos dados de pesquisa. Esse público é o grupo de usuários – pesquisadores brasileiros da área das Ciências Biológicas – ecologia. Para tanto, adota técnicas investigativas da metodologia de construção de personas e da atividade de identificação e mapeamento de públicos em relações públicas, como meios para atingir a resposta do problema e dos objetivos da investigação.

De modo generalista o termo usuário se refere a pessoas e/ou a grupos que utilizam determinado serviço; consomem ou usam algum produto; são beneficiadas por algum tipo de política pública e, portanto buscam as mais diversas informações sobre esses (LIMA 1989; FIGUEIREDO, 1994; HERNÁNDEZ SALAZAR, 1997; LARA; CONTI, 2003; MONFASANI; CURZEL, 2006; ROZADOS; PIFFER, 2009; RODRIGUES, 2017). Nesse estudo, compreende-se que o “usuário” é um dos diversos públicos que compõem o ambiente relacional de iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa. Essas iniciativas são promovidas e coordenadas por um sistema formado por instituições como a Rede Nacional de Ensino e Pesquisa (RNP), o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), o Grupo de Trabalho da Rede de Dados de Pesquisa Brasileira (GT – RDP Brasil) e outras formas de coletividades. Públicos são agentes de influência, uma vez que seus interesses são afetados e afetam as ações de uma organização na consecução de sua missão. Esses públicos podem se caracterizar como parceiros, pessoas ou conjunto de pessoas, grupos ou demais organizações que influenciam ou detêm o controle das decisões da organização e suas ações (SIMÕES, 2001; MARTINS, 2019). Conforme França (2012, p. 45), públicos são “[...] grupos organizados [...] que podem em determinadas condições prestar a efetiva colaboração às organizações, autorizando a sua constituição ou lhes oferecendo o suporte de que necessitam para o desenvolvimento de seus negócios”.

Personas são representações dos diversos perfis que formam o público estratégico de uma iniciativa, serviço, programa, etc.. São modelos baseados em informações coletadas com usuários reais ou em potencial. A construção das personas visa apresentar um ou mais paradigmas de quem são os usuários, como se comportam, pensam, o quê e o porquê desejam/necessitam algo ou alguma coisa. Os arquétipos podem definir tanto perfis ideais, como também não ideais, a fim de estreitar a compreensão de quem e como é esse público estratégico. As personas facilitam o processo de desenvolvimento de serviços, produtos e estratégias de comunicação a fim de estimular o uso ou consumo de algo, otimizando resultados de organizações e/ou instituições privadas ou públicas, promotoras das iniciativas. Cooper, Reimann e Cronin (2007) definem personas como modelos que descrevem os usuários, fornecendo informações sobre como eles se comportam, como pensam, o que desejam realizar e por quais motivos.

Personas são arquétipos compostos por dados comportamentais e motivacionais coletados com pessoas reais. Grudin e Pruitt (2002) afirmam que personas são pessoas fictícias baseadas em dados factuais, as quais têm nomes, roupas, ocupações, famílias, animais de estimação, posses, etc. Pruitt e Adlin (2006) consideram as personas como pessoas imaginárias descritas de modo detalhado e construídas a partir de dados sobre pessoas reais. Essas auxiliam na criação de produtos e serviços que pessoas reais têm interesse de usar. Sturmer *et al.* (2019a) apresentam a compreensão de personas em duas perspectivas. A primeira é a da tecnologia, focada no usuário e baseada na concepção de Alan Cooper (2004). Divide-se em três categorias: marketing personas, proto-personas e design personas. A segunda perspectiva é a do *marketing*, por meio dos conceitos de *buyer persona* e *inbound marketing*. Essas concepções serão mais bem explicadas nos capítulos subsequentes.

O acesso aberto aos dados de pesquisa almeja a cooperação e o avanço da ciência no âmbito global, promovendo a ideia e a prática do uso e do compartilhamento de dados de pesquisa dentro de repositórios digitais (KREBS; CAREGNATO, 2017). Trata-se de disponibilizar literaturas de caráter acadêmico ou científico na internet, permitindo a qualquer usuário ler, baixar, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral dos documentos (FIOCRUZ, 2022). Esse movimento faz parte do guarda-chuva da ciência aberta, que abrange diferentes práticas, desde o acesso aberto aos recursos educacionais e à literatura científica, até um modelo metodológico que visa à colaboração e ao compartilhamento de dados e conteúdos científicos entre pesquisadores e público interessado. Nessa perspectiva os dados de

pesquisa podem ser acessados livremente por qualquer indivíduo que tenha interesse em utilizá-los, reutilizá-los e redistribuí-los, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença (FIOCRUZ, 2022). O movimento da ciência aberta “promove o aumento dos estoques de conhecimento público, propiciando [...] a ampliação dos índices gerais de produtividade científica e de inovação, [...] das taxas de retornos sociais dos investimentos em ciência e tecnologia” (ALBAGLI, 2015, p. 14).

A ecologia é uma especialidade das Ciências Biológicas que estuda a interação entre os seres vivos e o ambiente em que vivem. Conforme Reichman, Jones e Schildhauer (2011), ecologia é uma disciplina sintética que se beneficia do acesso aberto a dados da terra, da vida e das ciências sociais. Esses dados são de natureza dispersa e heterogênea, características que ampliam os desafios tecnológicos para o compartilhamento dos mesmos. A variedade de tipos de dados pode incluir aspectos físicos, químicos e biológicos de diferentes tipos de organismos, provenientes de diversos locais/sensores/ambientes monitorados e documentados numa gama de instrumentos como cadernos de campo e de laboratório, planilhas, entre outros. Conforme a classificação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) a ecologia é uma das disciplinas clássicas da área da Biodiversidade e faz parte do Colégio de Ciências da Vida – Ciências Biológicas. As demais disciplinas consideradas clássicas são: botânica, zoologia, oceanografia biológica e biologia da conservação.

O *Global Research Report Brazil* (ADAMS; KING, 2009), informa que entre 2003 e 2007, o Brasil publicou mundialmente cerca de 85.000 artigos. Entre as dez principais categorias classificadas pela participação do país nas publicações está "Meio Ambiente/Ecologia", na quarta posição, antecipada por Ciência Vegetal e Animal, Ciências Agrárias e Microbiologia; seguida por Farmacologia e Toxicologia, Neurociência e Comportamento, Física Imunologia, Ciência Espacial, Biologia e Bioquímica. Tal resultado demonstra o crescimento das Ciências da Vida, particularmente relacionada aos recursos naturais. Os autores citam “É realmente a economia do conhecimento natural” (ADAMS; KING, 2009, p. 6). Os dados publicados por Haeffener, Zanoto e Guimarães (2021) indicam que a ecologia está na oitava posição entre as áreas mais destacadas no Brasil em produção científica e colaboração internacional, antecedida por oncologia, astronomia e astrofísica, física de campo e de partículas, física nuclear, sistema cardíaco e cardiovascular, biologia celular e neurologia clínica.

De acordo com a publicação de Santos, Freitas e Freitas (2016) o país formou aproximadamente 500 doutores por ano nos diversos temas ligados à biodiversidade, no período de 2013 a 2014. Porém, apesar da qualificação de recursos humanos e de capacitação profissional, alguns problemas básicos se mantiveram. Entre eles: a pouca comunicação entre a ciência produzida pela área no Brasil e a tomada de decisão em diferentes níveis da administração pública e da decisão empresarial; e a carência de profissionais nas áreas de gestão de biodiversidade e biologia da conservação, limitando a incorporação dos avanços, formulações e achados científicos e tecnológicos por parte dos atores sociais (SANTOS; FREITAS; FREITAS, 2016). Os autores observaram que o Brasil possuía a maior biodiversidade do planeta, abrigando cerca de 20% das espécies existentes e era provedor de importantes serviços ambientais para o mundo. Na última década o país teve considerável avanço na geração de conhecimento e formação de recursos humanos ligados à biodiversidade. Em 2019, a área de biodiversidade era constituída por 149 programas de pós-graduação, que se dividem em 100 programas com mestrado e doutorado acadêmicos, 43 cursos apenas de mestrado acadêmico, um programa profissional com mestrado e doutorado e cinco mestrados profissionais, abrangendo as disciplinas clássicas citadas (SANTOS; FREITAS; MELO; 2019).

Ao considerar sobre o futuro da área da biodiversidade, que inclui a ecologia, no que tange à inovação e à transformação, Santos, Freitas e Melo (2019) apontam a necessidade de estabelecer uma agenda de pesquisa e de formação de recursos humanos, uma vez que questões relacionadas à área têm um papel decisivo na competitividade econômica dos países tropicais e na qualidade de vida das futuras gerações. Esse contexto impõe novas articulações para a produção e disseminação de conhecimento. Os autores afirmam que o Brasil possui um sistema amplo e competitivo de produção de ciência e formação de jovens na área. Portanto, tem a vocação para liderar o empenho global no sentido de harmonizar o uso e a conservação dos recursos naturais com o bem-estar humano.

Desde 2017, a RNP possui um acordo de cooperação com o IBICT para atividades de Pesquisa e Desenvolvimento (P&D). Nos últimos anos, trabalham conjuntamente nos Planos de Ação Nacional no âmbito da Parceria para Governo Aberto (*Open Government Partnership – OGP*). A RNP é uma rede nacional para educação superior, pesquisa e inovação. Em 1992, auxiliaram no processo de implantação da internet no Brasil e continuam a promover o uso inovador de Tecnologias da Informação e Comunicação para impulsionar ciência e educação para

todos. O IBICT tem como missão promover a competência, o desenvolvimento de recursos e a infraestrutura de informação em ciência e tecnologia para a produção, socialização e integração do conhecimento científico e tecnológico (IBICT [*on-line*]; RNP, 2021).

Em 2018, o GT – RDP Brasil coordenou o projeto de mapeamento do uso de dados abertos de pesquisa no Brasil, a fim de identificar as práticas de acesso aberto aos dados de pesquisa em instituições brasileiras. Projeto esse iniciado pela Rede Nacional de Ensino e Pesquisa (RNP), coordenado pela Universidade do Rio Grande do Sul (UFRGS) e apoiado pela Universidade Federal do Rio Grande (FURG), até 2020. As equipes de pesquisadores estão ligadas ao Centro de Documentação e Acervo Digital da Pesquisa (CEDAP) da Faculdade de Biblioteconomia e Comunicação e ao Centro de Processamento de Dados (CPD/Lume) da UFRGS, em parceria com o IBICT (REDE NACIONAL DE ENSINO E PESQUISA, 2018).

Uma das atividades do GT – RDP Brasil foi realizar uma *survey* em 2018, que teve como objetivo mapear os potenciais usuários nacionais de serviços de acesso aberto aos dados de pesquisa. O questionário *on-line* foi composto por oito questões demográficas, 13 questões relativas às práticas dos pesquisadores no que tange dados de pesquisa e seis que buscaram identificar as percepções em relação ao compartilhamento, uso e reuso de dados. Todas as 27 questões foram objetivas. O instrumento foi enviado para um corpus formado por pesquisadores de todo o Brasil, “contatados por meio de correio eletrônico de instituições de pesquisa, coordenadores dos programas de pós-graduação, líderes de grupo de pesquisa cadastrados no Diretório do CNPq” (VANZ *et al.*, 2018, p. 3). Como resultado a *survey* apresentou dados importantes que podem subsidiar as primeiras fases para a construção das personas. Algumas das principais conclusões em relação aos usuários de acesso aberto aos dados de pesquisa no país são que 56,78% se identificam como sendo do gênero masculino, com média de idade de 48 anos, 45,20% são doutores e 36,95% pós-doutores que produzem dados dos tipos observacionais, experimentais e documentais. A produção dos dados de pesquisa é desenvolvida em sua maioria por grupos de até cinco pessoas (55,97%) e muitos possuem o conceito equivocado sobre o que é um repositório de dados, indicando que esse é um tema que precisa ser explorado.

Quando analisado com foco nos usuários de acesso aberto aos dados de pesquisa das Ciências Biológicas – ecologia, observa-se que esses são em maioria do sexo masculino, doutores e pós-doutores pertencentes a diferentes instituições, principalmente do setor acadêmico público. Essas pessoas costumam compartilhar parte dos dados de pesquisa por meio de repositórios de

revista ou temáticos, estimulados pela publicação em periódicos ou em anais de eventos, como também por iniciativa própria (VANZ *et al.*, 2018). Os resultados obtidos na *survey*, no que tange aos respondentes da área da ecologia serão mais bem explicados na seção 2.5 dessa pesquisa.

Em 2021, a RNP, o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e o IBICT celebraram acordo tripartite para ciência aberta. A iniciativa visa promover ações relacionadas à construção de uma federação nacional de repositórios de dados de pesquisa, uma vez que a demanda de abertura e de compartilhamento de dados cresce exponencialmente no país. O acordo tem apoio consultivo da Faculdade de Biblioteconomia e Comunicação da UFRGS, por meio da equipe do GT – RDP Brasil. Esse contexto torna fundamental que universidades e institutos de pesquisa implantem repositórios institucionais de dados de pesquisa, com o intuito de armazenar, organizar e disponibilizar objetos digitais (RNP, 2021b).

Ações para o acesso aberto aos dados de pesquisa estão em contínuo desenvolvimento no Brasil e demandam a conexão entre investigadores e o conhecimento produzido. Nesse cenário o GT – RDP Brasil considera importante a articulação estratégica entre os atores dessa realidade, a fim de estabelecê-las e fortalecê-las. No entanto, quem e como são esses atores? Pesquisadores, professores, doutores, acadêmicos, instituições de pesquisa? De que área e subáreas do conhecimento? Como os usuários brasileiros de serviços de acesso aberto aos dados de pesquisa compreendem e praticam o acesso aos dados de pesquisa produzidos por si e seus pares? Como são caracterizados e representados? Sem saber sobre esses perfis de forma mais detalhada, temos como promover e incentivar de modo eficiente, eficaz e efetivo o compartilhamento de dados de pesquisa, visando o crescimento da produção de conhecimento científico no Brasil?

Nessa lógica, a presente tese se desdobra numa sequência de capítulos que embasa a investigação. Inicia expondo as razões para a pesquisa ser, trazendo argumentos sobre sua importância para a sociedade, categoria profissional e comunidade acadêmica. Justificando então, o problema central que motiva a investigação, como também os objetivos – geral e específicos. Logo, discorre sobre a ciência aberta (*Open Science*) e seu contexto, trazendo conceitos sobre dados de pesquisa, tipos de dados produzidos pelos ecologistas e mecanismos de compartilhamento.

Em seguida, explica o usuário como público e como ele é entendido como persona. Na sequência, apresenta um compilado de referências sobre a metodologia de construção de personas e suas diferentes concepções sobre o tema. Posteriormente, os conceitos de públicos em relações

públicas, a partir de autores que estudam o objeto e propõem classificações e técnicas de pesquisa para identificá-los e mapeá-los. A última subseção descreve o perfil dos usuários brasileiros de acesso aberto aos dados de pesquisa no que tange às Ciências Biológicas – ecologia a partir de uma perspectiva teórica nacional. Esses dados preliminares foram fundamentais para conhecer algumas informações importantes para a construção das personas dessa pesquisa.

Os capítulos subsequentes apresentam a metodologia, a análise dos dados obtidos na investigação, discussão e apresentação dos resultados, finalizando com as conclusões.

1.1 JUSTIFICATIVA

A defesa de causa dessa pesquisa é orientada por três balizadores. São eles: benefícios para sociedade, para categoria profissional de relações públicas e para a comunidade acadêmica.

O primeiro balizador é o de sua importância para a sociedade, uma vez que tem como proposta a construção de personas para iniciativas em prol do acesso aberto aos dados de pesquisa, uma das pautas do movimento da ciência aberta no Brasil. Essa ciência estimula a reflexão sobre atividades, processos, métodos, avaliações e produções científicas; incentiva a transparência da pesquisa científica e propicia a ampliação dos estoques de conhecimento público – promovendo aumento dos índices gerais de produtividade científica e de inovação; e das taxas de retornos sociais dos investimentos em ciência e tecnologia (ALBAGLI, 2015; SILVA; SILVEIRA, 2019; FIOCRUZ, 2022). Portanto, identificar e mapear os perfis de usuários reais ou em potencial para iniciativas nacionais pode colaborar para a efetividade dessa pauta da ciência aberta na sociedade brasileira.

O segundo balizador é o benefício para a área de relações públicas. Confia-se que a compreensão da metodologia de construção de personas agregará valor às atividades desempenhadas pelos profissionais no que tange à identificação, ao mapeamento e à classificação de públicos organizacionais. Atividade essa considerada básica no planejamento de RP, segundo o Conselho Regional de Relações Públicas da 2ª Região (2016). Fortes (2003) afirma que o conhecimento dos públicos é crucial. O conhecimento dos grupos que estão direta ou indiretamente interessados na organização e a identificação desses como público é uma das fases do processo de RP. No entanto, o autor observa que o problema principal é definir por onde começar a pesquisa. Ainda não existem parâmetros ou requisitos para metodizar a classificação

dos públicos na área. De Avellar Barbosa (2006) também considera que a análise dos públicos envolvidos com a organização e o reconhecimento da situação desses relacionamentos são etapas fundamentais e estratégicas no planejamento de RP. Como afirma França (2012, p. 8), baseado no Documento Final do Parlamento Nacional de Relações Públicas, apresentado no XV Congresso Brasileiro de RP em 1998 “[...] cabe à atividade de relações públicas “diagnosticar os relacionamentos das entidades com seus públicos””. Segundo o autor, o gerenciamento de modo adequado dos relacionamentos entre a organização e seus públicos é o objeto de preocupação dos profissionais da área, uma vez que os relacionamentos podem ser múltiplos, interatuantes, coletivos, efêmeros, duráveis, permanentes, internos ou externos. Portanto, conhecer as características do público é essencial para “particularizar mensagens, responder a uma necessidade percebida e oferecer argumentação lógica” (FRANÇA, 2012, p. XIII). Assim sendo, este debate estimula ver e posicionar o profissional da área como especialista em levantamento de perfis estratégicos para ações organizacionais, visando resultados assertivos, tanto no âmbito institucional quanto mercadológico do negócio.

O terceiro balizador abrange os benefícios para a comunidade acadêmica a partir de quatro perspectivas. Primeiramente, observa-se que metodologia de construção de personas é utilizada por diversas áreas do conhecimento. Mesmo sendo similares, não há consenso sobre o processo percorrido para a construção dos arquétipos. Desse modo, debater sobre técnicas qualitativas e quantitativas para o levantamento dos dados que formam os perfis enriquecerá a discussão sobre o tema. A segunda perspectiva abrange os benefícios para a comunidade acadêmica no panorama de pesquisa em RP, já que estimula refletir sobre técnicas que beneficiam e renovam as concepções e as práticas metodológicas de identificação, mapeamento e classificação de públicos. “É complexa a relação das organizações com seus públicos. Para compreendê-la, é necessária uma pesquisa que possa analisar as diferentes formas dessa interação e que forneça dados suficientes para o conhecimento mais profundo dos diferentes públicos” (FRANÇA, 2012, p. 48). Junto a isso, possibilita ampliar a literatura sobre o assunto no Brasil, já que são poucos os (as) autores (as) que escrevem sobre a metodologia da atividade na área. A terceira perspectiva abrange a área de Ciência da Informação e se divide em dois vieses: 1) Aprofunda o conhecimento sobre estudo de usuários por meio da metodologia de construção de personas; 2) Propõe a construção dos arquétipos pertencentes à primeira das áreas do

conhecimento mais citadas na *survey* do GT – RDP Brasil (2018). Esses conhecimentos visam agregar valor às pesquisas já realizadas pelo GT, a fim de otimizar resultados e investimentos.

A quarta perspectiva de benefícios para a comunidade acadêmica favorece os pesquisadores das Ciências Biológicas de instituições brasileiras, especificamente da ecologia, já que busca compreender quem são esses indivíduos e o contexto do compartilhamento e reuso de dados entre esses, seus pares e a sociedade. Barros (2005) afirma que os pesquisadores da ecologia têm realizado pesquisas de longa duração, envolvendo dados com características geográficas, ecológicas e históricas, o que demonstra a diversidade de tipos de informação produzida.

Frehner e Brändli (2006) afirmam que novos desafios e requisitos para o gerenciamento de dados ambientais exigem interoperabilidade, integração e compartilhamento de dados aprimorados. A interoperabilidade é de crescente importância para o gerenciamento de dados ecológicos espacialmente relacionados, pois o gerenciamento de dados em ecologia consiste em várias tarefas, incluindo aquisição, controle e garantia de qualidade, manipulação, análise, interpretação, arquivamento e publicação. Michener (2006; 2015) destaca que os bancos de dados ecológicos continuam crescendo em volume, amplitude e complexidade, apresentando nível superior nas descrições de dados. Porém, reforça que o compartilhamento de dados tem uma longa história em outras áreas de conhecimento, como oceanografia e ciências da biodiversidade, mas surgiu recentemente nas ciências ecológicas. E destaca a ausência de políticas eficazes para o surgimento de comunidades e políticas de compartilhamento de dados, juntamente com o desenvolvimento e adoção de dados e padrões de metadados e ferramentas de habilitação na área. Para Reichman, Jones e Schildhauer (2011) a parcela de dados legíveis e acessíveis no contexto da ecologia é pequena, devido aos desafios tecnológicos e sociais que dificultam o acesso aberto a eles. A padronização de métodos e o desenvolvimento de metadados robustos podem aumentar o acesso aos dados, mas esse esforço ainda não é o suficiente. A reprodutibilidade das análises também é aspecto importante na ecologia. No que tange aos desafios sociais, os autores observam que as recompensas para compartilhar dados ainda são inadequadas, tanto para pesquisadores quanto para instituições de pesquisa. O estabelecimento de repositórios de dados federados com boa curadoria fornecerá um meio de preservar os dados enquanto promove a atribuição e o reconhecimento de seu uso. Esses aspectos sugerem a complexidade dos processos da área que

podem influenciar na elaboração e implantação de um repositório de dados abertos e compartilhados para a ecologia.

Michener (2015) observa que o compartilhamento de dados ecológicos tem evoluído lentamente desde 1950, no entanto, a cada dia se torna um pré-requisito para financiamento de pesquisas. A introdução de repositórios de dados baseados na comunidade, padrões de dados e metadados, *softwares*, identificadores persistentes e pesquisa e descoberta federadas ajudam a promover o compartilhamento de dados. Anjos (2016) considera que a área da ecologia está em constante desenvolvimento no Brasil, principalmente no nordeste, necessitando do uso e da preservação de dados, já que esses são fontes primárias, podendo ser utilizados e reutilizados em pesquisas futuras, a fim de compreender e/ou certificar fenômenos. Isto posto, observa-se a considerável demanda da ecologia no desenvolvimento de ações que visam o acesso aberto aos dados de pesquisa produzidos pela área.

Corroborando, os pesquisadores das Ciências Biológicas – ecologia – questionados na *survey* do GT – RDP Brasil (2018) – foram os que mais demonstraram interesse em participar de uma entrevista em profundidade para explorar as questões abordadas no questionário. Dos 4.703 respondentes, 1.061 se dispuseram a participar de entrevista para aprofundar o conhecimento sobre os potenciais usuários nacionais de serviços de acesso aberto aos dados de pesquisa no Brasil (PAVÃO; GABRIEL JÚNIOR; ROCHA, 2018; VANZ *et al.*, 2018; CAREGNATO *et al.*, 2019). No entanto, segundo informações coletadas diretamente com os integrantes do GT – RDP Brasil, a etapa de entrevista não ocorreu como previsto. Esse contexto oferece mais oportunidade para o presente estudo, posto que busca levantar qualitativamente percepções, aspirações e receios quanto ao processo de compartilhamento de dados abertos dos participantes.

Portanto, partindo do pressuposto de que as práticas de compartilhar e reutilizar dados são amplamente variáveis entre os cientistas; que para garantir o sucesso de projetos e iniciativas se torna ponto *sinequa non* conhecer o comportamento, anseios e inquietações de seus públicos estratégicos; e que esses usuários têm vários perfis e características, não é indicado englobá-los numa mesma esfera. Esses usuários podem ter posicionamentos divergentes, caracterizando perfis diversos em relação à iniciativa (LISBOA, 2017). Junto disso, o acesso, a recuperação e a disseminação de informação só se dá a partir de um sistema desenvolvido conforme os interesses e necessidades desses atores (LIMA, 1989), mas para isso é preciso utilizar técnicas e métodos procurando saber Quem? Qual? Quando? Quanto? Como? “conhecer comportamentos, hábitos,

necessidades e demandas de clientes e/ou usuários, tendo como fim [...] qualificar [...] criar [...] serviços e produtos adequados às demandas e expectativas percebidas (ROZADOS; PIFFER, 2009, p. 179). Ao valorizar apenas as tecnologias utilizadas como pontos essenciais das estratégias as iniciativas possuem “visão turva” sobre determinada realidade (MONFASANI; CRUZEL, 2006; ROSSI, 2012). Como afirmam Pavão, Gabriel Júnior e Rocha (2018), é necessário buscar informações diretamente com o público interessado. E, como constata Baynes (2017), pesquisadores têm tempo limitado, eles precisam de informação clara, políticas simples, orientação objetiva, ferramentas que facilitem o compartilhamento e o gerenciamento de dados válidos e por isso é preciso compreender suas realidades, aspirações e inquietações.

A metodologia de construção de personas vem a beneficiar os objetivos traçados pelo GT – RDP Brasil e pelo movimento brasileiro que visa ao acesso aberto aos dados de pesquisa, já que ao categorizar e identificar esses perfis qualitativamente será possível aperfeiçoar os resultados da *survey* e a continuidade das ações do GT – RDP Brasil. Entende-se que a não distinção ou entendimento das especificidades desses atores pode influenciar na eficiência, eficácia e efetividade das soluções a serem propostas (MONFASANI; CRUZEL, 2006).

A partir do exposto chega-se a algumas questões que permeiam o tema: Quais as práticas e as diferenças disciplinares de produção, compartilhamento e reuso de dados entre os pesquisadores das Ciências Biológicas – ecologia no Brasil? Quais os estímulos e motivos os fazem compartilhar todos os dados de pesquisa? E porque outros compartilham parcialmente ou nem compartilham? Carvalho (2018, p. 163) também sugere:

Analisar diferenças disciplinares nas práticas de produção/obtenção, compartilhamento e (re) uso em uma mesma área do conhecimento.

Analisar as práticas de produção/obtenção, compartilhamento e (re) uso em projetos de pesquisas interdisciplinares.

Analisar diferenças disciplinares em relação ao tratamento de dados de pesquisa em repositórios digitais.

Analisar como as diferenças disciplinares interferem na gestão de dados de pesquisa.

Investigar como os periódicos científicos têm abordado a questão dos dados de pesquisa em diferentes disciplinas.

Portanto, a pergunta central dessa pesquisa busca saber **Quem são as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa?**

1.2 OBJETIVOS

Os objetivos dessa pesquisa se dividem em geral e específicos, a fim de definir o que se pretende atingir com a investigação e quais as etapas a serem realizadas para alcançar a resposta do problema.

1.2.1 Objetivo Geral

Compreender as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa.

1.2.2 Objetivos Específicos

Para atingir o objetivo geral foram pensados os seguintes objetivos específicos:

- a) Investigar quem é o público usuário na visão dos atores que promovem e/ou coordenam ações e iniciativas em defesa do movimento brasileiro de acesso aberto aos dados de pesquisa;
- b) Mapear as principais características dos perfis que formam o público usuário das Ciências Biológicas - ecologia, identificados na *survey* do GT – RDP Brasil (2018), a fim de aprofundar o conhecimento sobre esses;
- c) Construir as personas que representam esse público;
- d) Formular a apresentação das personas construídas.

2 REVISÃO DE LITERATURA

Este capítulo apresenta alguns pontos de vista sobre os conceitos-chave pesquisados nessa investigação. O texto é baseado em referências obtidas como resultado da pesquisa bibliográfica realizada. Primeiramente contextualiza o movimento da ciência aberta e a definição de dados de pesquisa. Seguindo nesse fluxo, o texto apresenta dados ecológicos e mecanismos de compartilhamento utilizados pelos pesquisadores da área. Logo, elucida sobre a metodologia de construção de personas, trazendo algumas das propostas investigativas. Isso posto, o texto discorre sobre os conceitos e classificações de públicos em RP a partir de alguns autores que propõem técnicas de pesquisa para identificá-los e mapeá-los. Na sequência explana o conceito de usuário, como esse grupo é entendido como público, como ele se torna o centro da estratégia e seu entendimento como persona. Por último, discorre sobre as informações preliminares sobre os reais e potenciais usuários brasileiros de serviços de acesso aberto aos dados de pesquisa no que tange à área da ecologia, numa perspectiva teórica nacional.

2.1 CIÊNCIA ABERTA (*OPEN SCIENCE*) E DADOS DE PESQUISA

Neste subcapítulo são apresentadas algumas abordagens sobre ciência aberta e o contexto que essa envolve atualmente (Figura 1), principalmente, no que tange às pautas dados abertos e acesso aberto aos dados de pesquisa, trazendo referências nacionais e internacionais. A ciência aberta é um movimento que envolve o acesso aberto à literatura científica e aos dados de pesquisa, como também outras práticas do fazer científico. A proposta é alinhada à cultura digital, à colaboração e ao compartilhamento de conteúdos científicos (CIDACS, 2022).

Silva e Silveira (2019) definem a ciência aberta como um movimento de incentivo a transparência da pesquisa científica, desde sua concepção até o uso de *softwares* abertos para compartilhamento, uso e reuso de dados. O Glossário de Acesso Aberto da Fundação Oswaldo Cruz (FIOCRUZ, 2022) compreende que esse movimento estimula a reflexão sobre atividades, processos, métodos, avaliações e produções científicas, visando não apenas à transparência, como também à colaboração e à abertura de dados. Essa perspectiva da ciência busca promover a elaboração de metodologias e gestão de dados científicos, para que esses possam ser disseminados, reutilizados e acessíveis, sem custos. Ela promove a disseminação científica que

visa à comunicação da informação científica e tecnológica transcrita em códigos especializados, orientada para um público formado por especialistas (ALBAGLI, 1996).

Além de promover a comunicação científica, que tem como objetivo à disseminação de informações especializadas entre os pares (BUENO, 2010), a ciência aberta busca tornar a pesquisa científica mais acessível para toda sociedade, estimulando a livre circulação do conhecimento. Prevendo a colaboração de não cientistas na pesquisa, promove a ciência cidadã – que estimula a contribuição direta de cidadãos no fazer científico, com recursos, cognição, dados e como pesquisadores (FIOCRUZ, 2022). Nesse sentido, a ciência aberta encoraja práticas de divulgação científica, já que compreende meios para a veiculação de informações científicas, tecnológicas ou associadas a inovações ao público leigo (BUENO, 2010).

Para Santos *et al.* (2017) a ciência aberta “promove a abertura de todo o processo científico e a translação do conhecimento, amplia os impactos sociais e econômicos da ciência, reforçando o conceito de responsabilidade social científica” (SANTOS *et al.*, 2017, p. 12). A figura a seguir representa o contexto da ciência aberta, incluindo as pautas que dele fazem parte:

Figura 1: Guarda chuva da ciência aberta



Fonte: elaborado pela autora, baseado em FIOCRUZ, 2022 [on-line].

Outro ponto importante nesse contexto é a existência de cinco escolas de pensamento da ciência aberta. Entre elas a de infra-estrutura; a pública; a escola de medição; a democrática e a escola pragmática. O quadro a seguir sintetiza informações importantes sobre cada vertente.

Quadro 1: Cinco escolas de pensamento da Ciência Aberta

Escola de Pensamento	Suposição Central	Grupos Envolvidos	Objetivo Central	Ferramentas e Métodos
Democrática	O acesso ao conhecimento é distribuído de modo desigual	Cientistas, políticos e cidadãos	Tornar o conhecimento livremente disponível para todos	Acesso aberto, propriedade intelectual direitos, dados abertos, código aberto
Pragmática	A criação do conhecimento poderia ser mais eficiente se os cientistas trabalhassem juntos	Cientistas	Abrir o processo de criação de conhecimento	Sabedoria das multidões, efeitos do trabalho em rede, dados abertos, código aberto
Infra-estrutura	A pesquisa eficiente depende das ferramentas e aplicações disponíveis	Cientistas e fornecedores de plataformas	Criar abertamente plataformas, ferramentas e serviços para cientistas	Plataforma e ferramentas colaborativas
Pública	A ciência precisa ser acessível ao público	Cientistas e cidadãos	Tornar a ciência acessível para cidadãos	Cidadão, ciência, relações públicas da ciência, blogs científicos
Medição	As contribuições científicas hoje necessitam de mediações de impacto alternativas	Cientistas e políticos	Desenvolver um sistema métrico alternativo para impacto científico	Altimetrias, revisão por pares, citação, fatores de impacto

Fonte: FRIESIKE; FECHER, p. 20, 2014 – tradução nossa.

Observando o quadro acima, percebem-se os diferentes coletivos envolvidos no contexto da ciência aberta. Albagli (2019) considera que a adoção e a difusão dessa ciência compreendem distintos atores e níveis de intervenção. Alguns dos atores citados pela autora são: o pesquisador individual, instituições de ensino e pesquisa, editoras públicas e privadas de publicações acadêmicas (periódicos científicos), agências de financiamento, avaliação e certificação.

A figura 2 representa uma síntese dos grupos citados por Friesike e Fecher (2014) e Albagli (2019).

Figura 2: Coletivos envolvidos na ciência aberta



Fonte: elaborado pela autora.

O movimento pela ciência aberta foi fomentado pelas ações em favor do acesso aberto às publicações científicas nas últimas décadas do Século XX. Esse período foi marcado pelos altos preços praticados pelas editoras comerciais, pelo endurecimento dos regimes de proteção da propriedade intelectual e pela privatização do conhecimento (ALBAGLI, 2019). As ações tiveram início nos anos 70, quando em 4 de julho de 1971, foi lançado, por Michael Hart, o Projeto Gutenberg, uma biblioteca digital que disponibiliza acesso livre ao conhecimento. Ao longo desse tempo diversas iniciativas internacionais foram desenvolvidas a fim de impulsionar o acesso aberto aos dados de pesquisa, tendo realizações importantes em todas as décadas. O decênio de 80 foi marcado pelo lançamento da revista *on-line The Public-Access Computer Systems Review*, em 1989, por Charles W. Bailey Jr. A década de 90 foi a que mais despontou o movimento, desde o lançamento da proposta de Tim Berners-Lee do primeiro servidor de *web*, em março de 1991; em 1997, o despertar da SciELO (*Scientific Electronic Library On-line*), pelo projeto da Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP), em parceria com o Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (BIREME), que veio a operar em 1998; e, em 1999, a Declaração sobre a Ciência e Uso do Conhecimento Científico (*Declaration on Science and the Use of Scientific Knowledge*), concebida pela Organização das

Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) na *World Conference On Science*, em Budapeste, Hungria (FAUSTO, 2013).

Nas décadas seguintes o movimento pelo acesso aberto aos dados de pesquisa se manteve ativo. Lançada em 2007, a *Brussels Declaration* foi assinada por 35 editoras científicas e estabeleceu como um de seus princípios o acesso público aos dados; em 2008, o *Interdisciplinary Body of The International Council for Science (ICSU)* criou o *World Data System (WDS)*, em Moçambique, África. O sistema mundial *WDS* fomenta o acesso universal e igualitário às informações científicas, aos serviços de dados, produtos e informações, tanto das ciências naturais quanto das ciências sociais e das humanidades. Em 2013, a organização internacional *Research Data Alliance (RDA)* foi estabelecida no intuito de promover o desenvolvimento de infraestrutura e de atividades para redução das barreiras de troca e compartilhamento de dados (*RDA, 2022*). No mesmo ano foi lançada a orientação para a mineração de dados/textos e concessão de direitos autorais para uso científico na União Européia, chamada de *STM – Text and Data Mining for non-commercial scientific research* (em português Texto e Dados de Mineração para Investigação Científica Não Comercial). E, em 2014, foi assinada a *Joint Declaration of Data Citation Principles* (Declaração Conjunta de Princípios de Citação de Dados), que garante o crédito para o pesquisador e estabelece os princípios da citação de dados, entre eles dados de fácil acesso, devidamente preservados, citados e referenciados (DUDZIAK, 2016).

No Brasil, as ações iniciais foram incentivadas pelo setor público. Em 20 de setembro de 2011 foi assinada a Declaração do Governo Aberto, a partir da colaboração entre os governos do Brasil, da Indonésia, do México, da Noruega, das Filipinas, da África do Sul, do Reino Unido e dos Estados Unidos. A cooperação gerou a *Open Government Partnership (OGP)*, uma iniciativa internacional que tem como objetivos difundir e incentivar globalmente práticas governamentais relacionadas à transparência dos governos, ao acesso à informação pública e à participação social. Também em setembro de 2011, foi divulgado o 1º Plano de Ação Brasileiro no âmbito da parceria para Governo Aberto. O documento foi composto por 32 compromissos de órgãos do Governo Federal com ações que pretendiam refletir as diretrizes, desafios e princípios que orientam o conceito de governo aberto (BASE DE CONHECIMENTO DA CONTROLADORIA-GERAL DA UNIÃO, 2011). Nos anos subsequentes, outros planos de ação nacional para o governo aberto foram lançados, como por exemplo, o 2º plano em 2013, o 3º em 2016 e o 4º em 2018. A cada ano novos compromissos foram agregados para fortalecer os

princípios de governo aberto, uma vez que a prática é um recurso fundamental para a consolidação da democracia e a base de uma administração governamental profissional e consciente. Outra ação importante do governo brasileiro, ainda em 2011, foi a publicação da Lei Nº 12.527, de 18 de novembro, que visa regular e assegurar o direito fundamental de acesso à informação por parte dos cidadãos e cidadãs nacionais (SILVA; SILVEIRA, 2019; CONTROLADORIA-GERAL DA UNIÃO, 2022).

Em 2016, o IBICT lançou o Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã, que tem como objetivo:

[...] demonstrar o seu valor estratégico e informacional e 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, 2016, [*on-line*]).

A ação é continuidade do Manifesto de 2005 que expressou a adesão do país ao movimento Europeu. O documento de 2016 é direcionado às universidades e institutos de pesquisa brasileiros; às sociedades científicas e academias de ciência; aos órgãos de fomento à pesquisa, aos editores de revistas ou periódicos científicos, aos cursos de pós-graduação e graduação nas áreas de informação, em especial Ciência da Informação; aos gestores e executores de programas e projetos de dados de pesquisa, especialistas em Ciência da Informação e Ciência da Computação e profissionais de informação em geral e a pesquisadores. Em 2021, iniciou-se a construção do 5º Plano de Ação Nacional de Governo Aberto brasileiro. O documento prevê compromissos para desenvolver a transparência pública, a participação social, o combate à corrupção, a responsabilização e a inovação por parte do governo federal e terá validade até 2023 (ATOJI, 2021). O processo de elaboração e de execução desse Plano é um trabalho coparticipativo do governo e da sociedade civil e está dividido em três etapas. Primeiramente, a definição dos temas a serem desenvolvidos por meio de políticas de Governo Aberto considerará a consulta pública realizada com sociedade civil, os indicados como importantes para o Governo e também pelos poderes, Legislativo e Judiciário. No segundo momento, são previstas oficinas de cocriação para indicação de atividades, responsáveis e prazos para concretização do Plano. A terceira e última etapa é a aprovação do plano por parte do Grupo de Trabalho da Sociedade Civil e do Comitê Interministerial Governo Aberto (CIGA), ambos responsáveis pela metodologia

aplicada para Assessoramento em Governo Aberto (CONTROLADORIA-GERAL DA UNIÃO, 2022).

Percebe-se então, que o à proporção que o movimento pela ciência aberta se expande, novas pautas são agregadas em seu escopo, ampliando significados e debates. Conforme cita Albagli (2019, p. 16) “Ciência aberta pode ser então entendida como um “movimento de movimentos”. Para a autora, o termo *ciência aberta* integra várias frentes de discussão, que estimulam novas formas colaborativas de produção e circulação da informação e do conhecimento científico, visando também sua apropriação social. À vista disso, o debate e as ações que visam o acesso aberto aos dados de pesquisa são cada vez mais incontestáveis no âmbito científico. Junto a isso, a abertura de dados potencializa a citação das pesquisas e demais publicações (BAYNES, 2017).

Logo, faz-se necessário compreender o que são dados de pesquisa e como são entendidos no contexto da ciência aberta. Para tanto, são apresentados alguns conceitos sobre o termo *dados* e suas diferentes classificações, como por exemplo, dados científicos, dados de pesquisa, dados abertos, entre outros. O referencial tem como principais contribuições publicações de pesquisadores e instituições de pesquisa que atuam na perspectiva da ciência aberta, devidamente citados no decorrer do texto.

De acordo com o Glossário da FIOCRUZ (2022) o termo *dados* se refere a uma sequência de símbolos ou valores representados formalmente em algum meio que servem de base para o processo cognitivo. *Dados* são resultados de uma investigação, de um cálculo ou de uma pesquisa, desenvolvidos por processo natural ou artificial que observa algo de interesse. A maioria dos dados está em formato digital. Conforme o *National Science Board – NSB da U.S National Science Foundation*, dados digitais são os já gerados em formato digital, os analógicos que foram posteriormente digitalizados e os metadados digitais que podem ser associados aos dados de pesquisa digital. O *National Science and Technology Council Interagency Working Group on Digital Data* entende que *dados de pesquisa digital* são todos os tipos e formatos de dados importantes para a ciência e pesquisa e educação em contextos globais, nacionais, regionais e locais. E que, os *dados de pesquisa digital* devem ser válidos para uso e reuso em aplicações educacionais e pesquisas científicas ou técnicas, desde que respeitem condições de proteção e autorização conforme leis e regulamentações aplicáveis (NSB, 2011). De acordo com o relatório *Principles Guidelines for Access to Research Data from Public Funding*, da

Organization for Economic Co-operation and Development (OECD, 2007), os dados de pesquisa em formato digital legíveis por computador possuem maior potencial para a distribuição eficiente. Ainda assim, a publicação considera os princípios e diretrizes para acesso a dados de pesquisa de financiamento público podem ser aplicados também aos dados em formato analógico, desde que os custos de acesso sejam razoavelmente baixos.

Já, Silva (2019) apresenta conceitos para o termo *dados científicos*. Entende-se então, que são informações registradas ou produzidas por meio de pesquisa, independente da metodologia, instrumento ou técnica aplicada no processo de investigação. Como também, “todas as evidências que um investigador necessita para validar suas conclusões após uma pesquisa” (SILVA, 2019, p. 21). Esses dados podem ser numéricos, descritivos e/ou visuais armazenados em papel ou em formato digital. Incluem amostras biológicas e/ou ambientais, relatórios de campo ou de laboratório, caracteres, símbolos, sons, imagens (estáticas ou em movimento), ondas eletromagnéticas, *bits*, gráficos, tabelas, etc. Silva *et al.* (2020, p. 308) cita que *dados científicos* (*scientific data* – em inglês) são “todo e qualquer tipo de registro coletado, observado, gerado ou usado pela pesquisa científica, tratado e aceito pela comunidade científica como necessário para validar os resultados de pesquisa”.

Para o relatório *OECD* (2007), o termo *dados de pesquisa* é definido como “[...] registros factuais (escores numéricos, documentos textuais, imagens e sons) utilizados como fontes primárias para a pesquisa científica, e que são comumente aceitos nesta comunidade como necessários para validação de resultados [...]” (*OECD*, 2007, p. 13). Conforme o Glossário de Termos do Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS, 2022) *dados de pesquisa* são:

Unidades de informação criadas ou coletadas no curso da pesquisa científica, que são frequentemente formatadas de maneira a torná-las adequadas à comunicação, à interpretação e ao processamento por computador. São exemplos de dados de pesquisa: planilhas de estatísticas, dados retirados de bases de dados públicas ou privadas, um registro sonoro de uma entrevista, um registro descritivo de um espécime de rocha, uma coleção de imagens digitais [*on-line*].

O Glossário de Acesso Aberto da FIOCRUZ (2022) apresenta definição similar para o termo *dados para pesquisa*, considerando-os objetos informacionais digitais aceitos pela comunidade científica para validar resultados de pesquisa. Incluem uma coleção de objetos digitais adquiridos e gerados durante o processo das investigações. Alguns exemplos são

documentos (Word), planilhas (Excel), cadernos de laboratório, diários de campo, questionários, transcrições, fitas de áudio e/ou vídeo, fotografias, sequência de proteínas ou genéticos, slides, artefatos, amostras, algoritmos, *scripts*, arquivos de log, *software* de simulação, metodologias, fluxos de trabalho, procedimentos operacionais, padrões e protocolos.

Sayão e Sales (2014) utilizam o conceito da *OECD* (2007) para definir o termo *dados de pesquisa* e mencionam que o entendimento sobre o mesmo pode variar consideravelmente entre pesquisadores da mesma área, ou não, de conhecimento. Os autores afirmam que os *dados de pesquisa* possuem uma heterogeneidade intrínseca e que isso demanda políticas que identifiquem e sustentem essa diversidade. Esses dados são gerados para diversos propósitos, por diferentes áreas acadêmicas e podem ser expressos de distintas formas conforme sua natureza (números, imagens, textos, vídeos, áudio, *software*, algoritmos, equações, animações, modelos, simulações). Variam também no seu valor, alguns têm valor imediato e duradouro, já outros adquirem valor ao longo do tempo. Outros ainda são capturados num momento específico e irrecuperável, enquanto outros são passíveis de reprodução (SAYÃO; SALES, 2015).

Baseado na categorização proposta pelo *National Science Board*, Sayão e Sales (2014) apresentam três classificações de *dados de pesquisa*, agora distinguidos por sua origem. São eles.

Dados observacionais – são obtidos por meio de observações diretas, que podem ser associadas a lugares e tempo específicos, como por exemplo, a erupção de determinado vulcão numa data específica, a fotografia de uma super nova, o levantamento das atitudes de uma comunidade. Os dados observacionais – por sua natureza instantânea – guardam uma importância crítica que os qualifica como registros históricos que não podem ser coletados uma segunda vez e, portanto, devem ser submetidos a processos de curadoria que os preserve para sempre.

Dados computacionais – são resultados da execução de modelos computacionais ou de simulações, seja, por exemplo, no domínio da física ou para a criação de ambientes virtuais culturais ou educacionais. Para esta categoria de dados a preservação por longo prazo pode não ser necessária, posto que os dados possam ser replicados ao longo do tempo. Entretanto, replicar o modelo ou a simulação no futuro pode exigir um grande número de informações que incluem descrição das dependências de hardware, software e outras dependências técnicas, e ainda os dados de entrada. É preciso notar que algumas vezes é mais conveniente preservar somente os dados de saída.

Dados experimentais – são provenientes de situações controladas em bancadas de laboratórios, como por exemplo, medidas de uma reação química. Em tese, dados experimentais provenientes “de experimentos que podem ser precisamente reproduzidos não necessitam ser armazenados indefinidamente; porém, na prática, nem sempre é possível reproduzir precisamente todas as condições experimentais, particularmente onde algumas variáveis experimentais não podem ser conhecidas e quando os custos de reprodução do experimento são proibitivos (SAYÃO; SALES, 2014, p. 81).

Em 2015, Sayão e Sales reforçam que os *dados de pesquisa* podem ser caracterizados conforme sua origem ou natureza, reapresentando as classificações acima a partir das definições de Green *et al.* (2009). Os autores acrescentam que os dados também podem ser classificados de acordo com a fase da pesquisa, sendo divididos em: *dados brutos, crus ou preliminares* (RAW DATA em inglês) proveniente diretamente dos instrumentos científicos utilizados para coleta; *dados derivados*, que são resultados do processamento ou da combinação de dados brutos ou de outros dados; e *dados canônicos ou referenciais*, que são coleções de dados consolidados e arquivados, geralmente, em grandes centros de dados, como por exemplo, sequência genética, estrutura química, etc. (SAYÃO; SALES, 2015).

O Glossário do CIDACS (2022) entende que *dados brutos* são os provenientes diretos dos instrumentos científicos, inalterados. Ou seja, não sofreram qualquer manipulação ou processamento. São também chamados de dados crus ou dados primários. A publicação também contempla as definições de *dados confidenciais*: sendo os que não estão em domínio público, como por exemplo, informações sobre negócios, lucros, saúde, detalhes médicos e opiniões políticas, entregues em confiança ou quando duas partes concordam em mantê-los confidenciais (segretos). De *dados pessoais* que são informações relacionadas a indivíduos vivos, que podem ser identificados a partir deles ou a partir da combinação com outras informações. E, de *dados abertos* que “São dados publicados em um formato legível por máquina e sem restrição de licenças, patentes ou mecanismos de controle, de modo a estarem livremente disponíveis para serem utilizados e redistribuídos à vontade” (CIDACS, 2022 [on-line]). Conforme Parham, Bodnar e Fuchs (2012), *dados abertos de pesquisa* possuem vários formatos e permitem comunicação, interpretação ou processamento, em formato digital organizado de modo estruturado por procedimento formal que visa gerar novas pesquisas.

No Glossário FIOCRUZ (2022) o termo *dados abertos* define todo e qualquer dado que pode ser usado, reutilizado e redistribuído por qualquer pessoa livremente, estando essa sujeita, no máximo, a referenciar a autoria desses. O índice também apresenta as definições de *dado anonimizado* que não identifica o indivíduo, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento. *Dado pessoal sensível* que retrata informações sobre a origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou à organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural. *Dados*

administrativos que são gerados e/ou acumulados em atividades administrativas do governo; já os *dados governamentais* correspondem à universalidade dos dados gerados e acumulados pelo governo. E, os dados, *identificado ou pessoal*, que são relacionados à pessoa natural identificada ou identificável.

Retomando a definição de *dados abertos*, conforme o item III, do art. 2º do Decreto n.º 8.777 que institui a Política de Dados Abertos do Poder Executivo Federal, de 11.05.2016:

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

O item II, do art. 2º do Decreto n.º 8.777 define que *dados acessíveis ao público* são:

II – dado acessível ao público – qualquer dado gerado ou acumulado pelo Governo que não esteja sob sigilo ou sob restrição de acesso nos termos da Lei n.º 12.527, de 18 de novembro de 2011 que regula o acesso a informações.

Medeiros (2018) afirma que o termo *dados abertos* pode ser entendido como qualquer tipo de objeto digital armazenado que pode ser encontrado através de mecanismos de busca por qualquer indivíduo. Porém, observa que os dados brutos não são necessariamente públicos, devido a questões éticas ou de privacidade.

Rautenberg (2018, p. 111), apresenta as definições da 5-Star OPEN DATA para classificar *dados abertos*, de acordo com seu nível de abertura e sua conexão a outros dados:

1 Estrela – conjuntos de dados publicados sob uma licença aberta (*Open License – OL*), entretanto, em um formato proprietário. Os dados somente podem ser manipulados (lidos, visualizados ou impressos) por determinados softwares. Como exemplo, tem-se os dados publicados no formato *Portable Document Format (PDF)*;

2 Estrelas – conjuntos de dados estruturados legíveis por máquinas (*Readable Machine – RE*). A esse nível, os dados são processados por softwares proprietários e podem ser exportados em outros formatos, facilitando os procedimentos de manipulação. As planilhas eletrônicas no formato Excel *Spreadsheet (XLS* – formato de planilha eletrônica da Microsoft) são classificadas neste nível;

3 Estrelas – conjuntos de dados publicados em algum formato aberto (*Open Format – OF*), como por exemplo, *Comma Separated Value (CSV)* ou *Java script Object Notation (JSON)*. Salienta-se que a manipulação dos dados no formato aberto não necessita o uso de um software proprietário;

4 Estrelas – é designada à utilização dos Identificadores Uniforme de Recursos (*Uniform Resource Identifier – URI*) para rotular os dados, permitindo que os usuários criem ligações e façam o reuso dos dados disponibilizados;

5 Estrelas – conjuntos de dados que são conectados (*Linked Data – LD*) a outros dados em uma infraestrutura de rede. Isso permite a navegação entre dados e a descoberta de

informação. Dessa forma, acrescenta-se valor aos recursos de dados ao promover uma contextualização mais ampliada.

Independente da definição e classificação utilizadas para os *dados de pesquisa*, alguns aspectos são fundamentais para a compreensão do contexto do acesso aberto. Entre os tópicos citados no relatório da *OECD* (2007, p. 15 – 22) como princípios e diretrizes do acesso aos dados de pesquisa estão: a “Abertura” que significa acesso aos dados de pesquisa em igualdade de condições para a comunidade de pesquisa internacional com o menor custo possível, de preferência não mais do que o custo marginal de disseminação. O acesso aberto a dados de pesquisa deve ser fácil, oportuno, amigável e, de preferência, disponibilizado via Internet. A “Flexibilidade” que leva em consideração a rápida e muitas vezes imprevisível mudança nas tecnologias de informação, as características de cada pesquisa e a diversidade de sistemas de pesquisa, sistemas jurídicos e culturas. A “Transparência” que considera que organizações de pesquisa e agências governamentais devem disseminar de modo ativo informações sobre políticas de dados. A “Conformidade Legal” que menciona o respeito aos direitos legais e os interesses legítimos de todas as partes interessadas. Os princípios incluem também a proteção de propriedade intelectual, o profissionalismo ao disseminar e reusar os dados, a interoperabilidade tecnológica e semântica, a qualidade, a segurança, a eficiência, a prestação de contas, ou seja, avaliação constante desses dados e, a sustentabilidade, que significa assumir a responsabilidade administrativa por medidas que garantem o acesso permanente a dados que foram determinados para existir em longo prazo (*OECD*, 2007). Acrescenta-se também a noção de “legibilidade” como um dos aspectos do contexto do acesso aberto, uma vez que influencia no processo de compartilhamento, uso e reuso. A legibilidade visa à fácil compreensão dos dados pelos usuários interessados, de agora e do futuro. Dados fazem parte do processo de comunicação para um conhecimento contínuo e, junto com a Informação¹, configuram o conteúdo das mensagens, a fim de transmitir e gerar sabedoria (*CORREIA*, 2009).

¹ Informação “tudo que possa representar notícia, conhecimento ou comunicação. Em computação uma coleção de fatos ou outros dados fornecidos à máquina para posterior processamento” ou “todo ou qualquer elemento referencial contido num documento (*SANTOS; RIBEIRO*, 2003, p. 128). A Informação pode ser estruturada disponível em banco de dados ou não estruturada, presentes em documentos (*SANTOS; RIBEIRO*, 2003). “A informação é um produto social, com valor de uso e valor de troca, que se manifesta em um ambiente dentro de um processo de comunicação” (*LIMA*, 1989, p. 165).

A partir da compreensão sobre ciência aberta e do conceito de dados de pesquisa nesse contexto, apresenta-se a seguir o entendimento sobre os dados de pesquisa produzidos pela ecologia.

2.1.1 Dados de pesquisa na ecologia

A ecologia é uma ciência observacional, nascida do trabalho de campo de naturalistas amadores no final do Século XIX. No início do Século XX foi transformada em disciplina profissional em universidades e laboratórios privados. Portanto, os dados da ecologia, são originalmente observacionais (SAGARIN; PAUCHARD, 2010). Em âmbito internacional, o campo da ecologia amadureceu rapidamente ao longo dos anos, ou seja, partiu de estudos de pequena escala, com observações de curto prazo e experimentos conduzidos por indivíduos para o desenvolvimento de projetos multidisciplinares de larga escala e de longo prazo. Esses estudos produzem dados diversos a partir de abordagens analíticas sofisticadas. As investigações da ecologia, muitas vezes, também requerem interações com disciplinas adjacentes, como por exemplo, evolução, genômica, geologia, oceanografia e climatologia. Além de campos díspares como epidemiologia e economia. Essa diversidade de subdisciplinas é a responsável pela heterogeneidade dos dados (por exemplo, ecossistemas/ecologia marinha/de água doce/ecologia vegetal/animal/microbiana/métodos/unidades de medição/detalhes do projeto experimental) (REICHMAN; JONES; SCHILDHAUER, 2011).

Compreende-se que ampliar os conhecimentos sobre os dados produzidos pelos pesquisadores das Ciências Biológicas – ecologia é importante, uma vez que:

A Biologia é a ciência responsável pelo estudo da vida (em Grego: bios = vida; logos = estudo). Assim, desde uma única célula até grandes animais, ali está ela. A origem, evolução, constituição, funcionamento e as interações que existem e que envolvem os seres vivos são também aspectos que ela aborda. Mais que isso, ela também estuda os fatores abióticos, e que interferem na existência da vida. Moléculas, substâncias, eventos biológicos – e até mesmo ações causadas pelo homem – também contemplam esta ciência [...] Assim como a filosofia, a Biologia pode dar subsídios para que encontremos respostas para dúvidas que acompanham a humanidade desde sempre: Quem somos? De onde viemos? Para que viemos? Para aonde vamos? (BIOLOGIANET, 2020, [on-line]).

Essa ciência pode contemplar uma variedade de tipos de dados, uma vez que investiga aspectos físicos, químicos e biológicos de diferentes tipos de organismos; pesquisa sobre as leis

da hereditariedade; é responsável pelo estudo dos vegetais, entre outros. Ao encontro disso, esses pesquisadores podem estar em locais/ambientes diferenciados durante a coleta, tabulação e análise de dados, oferecendo uma visão mais ampla sobre contextos onde a pesquisa brasileira é desenvolvida, aplicada e disseminada.

Para tanto, esta subseção contempla uma sequência de autores que centralizam suas investigações em dados de pesquisa produzidos no contexto das Ciências Biológicas – ecologia. As referências tiveram como objetivo tentar caracterizar quais as especificidades dos dados da área, a fim de entender esse universo. O texto foca inicialmente nas publicações internacionais e logo, apresenta a perspectiva nacional.

Os estudos em âmbito internacional apontam a diversidade dos dados produzidos no ambiente de pesquisa em ecologia. Michener (2006) informa que os dados ecológicos abrangem as ciências biológicas, químicas, físicas, sociais e muitas outras subdisciplinas, uma vez que a ecologia investiga os organismos e as relações estabelecidas entre esses e seus ambientes. Por causa das complexas interações nesse contexto, a ecologia tem evoluído como ciência e cada vez mais aborda questões em escalas de organização biológica, espaciais e temporais mais amplas e múltiplas. Resultando em mudanças nos tipos, magnitude e complexidade dos dados que são adquiridos e analisados por ecologistas.

Em 2008, Zimmerman afirmou que os ecologistas estavam no ápice de mudanças potencialmente significativas em suas práticas sociais e científicas. Essas transformações foram impulsionadas pela coleta de novos tipos de dados, pela disponibilidade *on-line* de grandes volumes de dados de interesse para área, por tecnologias que tornam mais fácil gerenciar e integrar dados separados e pela mudança lenta de pontos de vista sobre o valor da análise secundária para abordar importantes questões ecológicas. A autora menciona que ecologistas conduzem experimentos em ambientes de campo e de laboratório; e observam fenômenos no mundo natural usando uma gama de ferramentas e abordagens que eles aprendem como parte de sua cultura para a disciplina, baseada nas considerações de Kohler (2002); Roth e Bowen (2001a, 2001b). Onde quer que trabalhem, os ecologistas são confrontados com a variação entre os indivíduos de uma população, com o caráter inesperado do ambiente natural e com sistemas complexos com grande número de variáveis e muitas interações sutis. Por exemplo, mesmo com um guia de campo e espécimes em mãos, pode ser difícil para um ecologista distinguir uma árvore de outra porque um guia não pode mostrar todas as variações possíveis nas folhas, casca

ou estrutura, que ocorrem ao longo do tempo. O resultado é que os dados ecológicos são extraordinariamente complexos e apresentam problemas na interpretação e análise para reuso (ZIMMERMAN, 2008). No levantamento realizado com 13 entrevistados entre 2001 e 2002, Zimmerman (2008), constatou que os ecologistas produzem e reutilizam uma grande variedade de dados. Entre eles: observações de espécies, laboratório experimental, dados de campo, plantas, solo, medidas químicas da água, instâncias de controle biológico, história, características da floresta, anéis de árvores, entre outros.

Os autores Hunt, Baldocchi e Van Ingen (2011) afirmam que é eminente o crescimento exponencial de dados de pesquisa na área das Ciências Biológicas, em especial no âmbito da ecologia. Os estudos de síntese desenvolvidos estão mais facilitados devido à convergência de sensores de baixo custo, sensoriamento remoto, conectividade pela internet e computação comoditizada. O trio menciona que os dados da ecologia provêm de uma ampla gama de metodologias de amostragem ou simulação, aplicadas por diversos grupos que utilizam uma variedade de padrões de dados. Boa parte desses dados é digital, gerados a partir de sensores, radares ou satélites. Porém, dados de campo e históricos são extremamente importantes para a ciência, por isso devem também ser digitalizados. Os dados podem incluir “a data do surgimento dos primeiros brotos de folhas ou imagens aéreas em diferentes comprimentos de onda e resoluções para estimar quantidades em toda a bacia hídrica, como umidade do solo, vegetação e uso da terra” (HUNT; BALDOCCHI; VAN INGEN, 2011, p. 49).

Michener e Jones (2012) comentam que os pesquisadores da área abordam questões complexas em escalas que vão do gene à biosfera. A ecologia está se tornando cada vez mais uma ciência com uso intensivo de dados contando com grandes quantidades de dados coletados por satélites, plataformas de sensoriamento remoto e redes de sensores embutidos no ambiente. Os autores afirmam que os dados ecológicos são coletados e organizados de muitas maneiras, incluindo registro manual de observações no laboratório e no campo por meio de blocos de anotações, gravadores e computadores portáteis; e a coleta automatizada de dados por meio de instrumentação de laboratório e de campo. Alguns outros exemplos de tipos de dados produzidos são: experimental, observacional, bruto ou derivado, coleções físicas, modelos, imagens, etc. (MICHENER; JONES, 2012).

Hampton *et al.* (2013) aponta a considerável heterogeneidade de tipos de dados ecológicos ocorre por causa das complexas e diversas áreas de campo (sítios) utilizadas como

fontes de coleta. Os dados são provenientes de exaustivas observações comportamentais realizadas em locais remotos, através de redes de sensores, que reúnem um fluxo de dados contínuos. Em 2015, Vanderbilt *et al.*, ao estudarem sobre o processo de compartilhamento de dados ecológicos na *International Long Term Ecological Research Network*, mencionaram que os cientistas, os quais buscam abordar questões de pesquisa em escala regional ou global, frequentemente utilizam dados coletados em mais de uma disciplina e de vários locais (VANDERBILT, 2015).

Em 2015, Michener afirma que os dados de pesquisa da ecologia podem ser derivados de observações humanas diretas em campo, registrados em cadernos e outras mídias; de observações em laboratório, sensores remotos e instrumentos que são empregados para medir atributos particulares, por exemplo, presença, temperatura, como também do ambiente físico, por exemplo, ar, solo, água – como chuva, radiação solar, umidade do solo e pH. Conforme a *Biodiversity Information Facility (GBIF, 2021)* os dados ecológicos são derivados de muitas fontes, desde espécimes de museus coletados nos Séculos XVIII e XIX a fotos de *smartphones* georreferenciadas que são compartilhadas por naturalistas amadores nos últimos dias e semanas.

No Brasil, Barros (2005) alertou sobre a diversidade dos dados da área no que tange a aspectos geográficos e históricos, como também sobre o grande volume de dados existente no país (BARROS, 2005). Os dados ecológicos podem vir de qualquer lugar, como também ser coletados por meio de observações diretas no campo e em estudos baseados em experimentos (SAGARIN; PAUCHARD, 2010; DRUCKER, 2011). É comum entre os ecologistas, que o armazenamento dos dados ocorra em planilhas digitais, simultaneamente com um bloco de anotações. A observação direta é uma prática que tem origem na ciência ecológica e que volta a ser valorizada pelos pesquisadores da área. As diferentes formas de coleta, aliadas às várias abordagens e tradições de pesquisa da ecologia, suas subdisciplinas e áreas afins, geram dados altamente heterogêneos. Alguns desses dados podem ser representados por contagens de indivíduos, medidas de variáveis ambientais (temperatura ou precipitação) ou representações de processos como competição ou herbívora, ou de seus efeitos (grau de abertura de dossel ou quantidade de dano em folhas). Os diferentes tipos de dados ecológicos tornam sua compreensão complexa, principalmente se reutilizados por outra pessoa, que não participou do levantamento original (DRUCKER, 2011). Na publicação de Tabarelli *et al.* (2013) também é possível verificar a diversidade de tipos de dados utilizados pelos pesquisadores brasileiros. A referência traz uma

série de artigos, com diversos autores, que relatam suas experiências e dados de longa duração produzidos em suas investigações. Entre eles estão informações sobre: altitude, longitude, temperatura, níveis de relevo, tipos de terreno, imagens geográficas, entre outros (LUIZÃO *et al.*, 2013).

Desse modo, percebe-se que os dados produzidos pela ecologia são heterogêneos e podem ser classificados como *observacionais*, *experimentais* e *computacionais*, indo ao encontro das definições de Sayão e Sales (2014). Isso devido às suas características, os *dados observacionais* são os coletados por observação direta, em lugares e tempos específicos e possuem natureza instantânea. Os *dados experimentais* são obtidos em laboratório. E os *dados computacionais* são resultados da execução de modelos de computador ou de simulações. Verifica-se também que os dados produzidos pelos ecologistas podem ser classificados conforme a fase de pesquisa, citadas por Sayão e Sales em 2015, uma vez que se enquadram nos conceitos de *dados brutos*, *crus ou preliminares* vindos diretamente dos instrumentos científicos utilizados para coleta; e *dados derivados*, que são resultados do processamento ou da combinação de dados brutos ou de outros dados (SAYÃO; SALES, 2015). Confia-se que essa realidade influencia diretamente o contexto para a prática de uma ciência aberta entre os pesquisadores da ecologia, como também o público interessado. A variedade dos tipos de dados coletados e as formas de armazenamento desses, afeta diretamente o processo de compartilhamento, uso e reuso de dados de maneira integrada, tornando a proposta do acesso aberto aos dados de pesquisa um desafio no caso das Ciências Biológicas – ecologia.

O subcapítulo seguinte aborda brevemente sobre o ciclo de vida de dados de pesquisa e os possíveis mecanismos de compartilhamento existentes.

2.1.2 Mecanismos de compartilhamento

A ecologia consiste em uma disciplina integrativa e colaborativa, que gera dados heterogêneos, com crescimento exponencial, o que amplia a necessidade de acesso aos dados. Esse amplo escopo cria importantes desafios para encontrar maneiras eficazes de descobrir, acessar, integrar, curar e analisar o intervalo e volume de informações relevantes. Avaliações realizadas no aspecto tecnológico ligadas ao compartilhamento de dados da subárea ecologia descreveram três grandes desafios: dispersão de dados, heterogeneidade e proveniência.

Ecossistemas e habitats variam em todo o globo e os dados são coletados em milhares de locais. Dados ecológicos são difíceis de descobrir e preservar porque estão contidos em, relativamente, pequenos conjuntos de dados dispersos entre dezenas de milhares de pesquisadores independentes. A heterogeneidade dos dados cria desafios devido à variedade de tópicos estudados por ecologistas e os diversos protocolos experimentais usados por esses (REICHMAN; JONES; SCHILDHAUER, 2011).

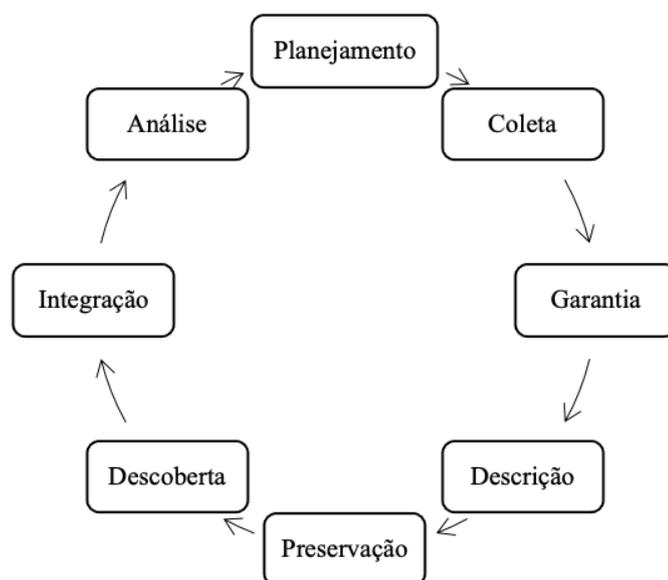
Hampton *et al.* (2013) comentam que os ecologistas produzem uma quantidade enorme de dados, mas ainda não existe a cultura de troca e agregação desses entre os estudiosos. Apoiados em Heidorn (2008), os autores afirmam que a ecologia é dominada pela ciência da “cauda longa”. Isto é, uma ciência conduzida por investigadores individuais, muitas vezes em escalas espaciais e temporais limitadas, sob modelos de financiamento que fornecem capacidade reduzida para curadoria ou compartilhamento de dados.

Michener (2015) destaca que o compartilhamento de dados tem uma longa história em outras áreas de conhecimento, como oceanografia e ciências da biodiversidade, mas evoluiu lentamente na ecologia, desde 1950. O autor comenta a ausência de políticas eficazes para o surgimento de comunidades e políticas de compartilhamento de dados juntamente com o desenvolvimento e adoção de padrões de metadados e ferramentas de habilitação na área. Antes e durante a primeira metade do Século XX, indivíduos ou um pequeno número de pesquisadores realizavam a maioria dos estudos de ecologia em um curto período de tempo e com financiamento limitado. Nessa época os dados eram publicados como tabelas em um manuscrito, o compartilhamento de dados não era a norma. Essa situação começou a mudar na década de 1960 em resposta ao surgimento da "grande ecologia". Desde então os programas variam de grandes esforços de pesquisa de ecossistemas baseados em locais a centros que oferecem suporte ecológico síntese e acesso a dados globais de biodiversidade. No decorrer dos anos os ecologistas passaram a usar dados compartilhados por outros cientistas para fins comparativos ou para aumentar suas coleções. Assim sendo, ações como a criação de repositórios de dados baseados na comunidade, padrões de dados e metadados, ferramentas de *software*, identificadores persistentes, pesquisa e descoberta federadas ajudam a promover o compartilhamento no campo da ecologia.

Para Michener e Jones (2012) o conhecimento é derivado da aquisição de dados e da transformação desses em informações que podem ser incorporadas ao corpus de fatos científicos,

princípios e teorias. A figura 3 ilustra os oito estágios pelos quais os dados ecológicos podem progredir para completar o processo de compartilhamento e de reuso.

Figura 3: Ciclo de Vida de Dados



Fonte: MICHENER; JONES, 2012, p. 86 – tradução nossa.

Os estudiosos consideram o planejamento de gerenciamento de dados uma etapa importante para economizar tempo, aumentar a eficiência da pesquisa, como também satisfazer os requisitos dos financiadores, cada vez mais exigentes em relação aos planos como parte das propostas das investigações. Para eles um planejamento abrangente deve conter informações sobre os dados (formato, quando, onde e como foram coletados); como foram processados para análise (*software*, fluxos de trabalho); a identificação de formatos de arquivo e outras fontes de dados utilizadas; conteúdo e formato de metadados; políticas de acesso, compartilhamento e reutilização; indicação do repositório que os dados serão armazenados e preservados; e orçamento (custos previstos para preparação de dados, documentação, *hardware*, *software*, pessoal e arquivamento, incluindo também fontes de financiamento).

A preservação engloba o armazenamento de dados e metadados em repositórios que oferecem suporte a diferentes níveis de verificação, replicação e curadoria de dados ao longo do tempo. Esses repositórios podem ser associados a instituições governamentais e organizações não

governamentais, universidades, bibliotecas, redes de observatórios ambientais e comerciais e empresas sem fins lucrativos (MICHENER; JONES, 2012).

Os repositórios digitais (RDs) são bases de dados *on-line* que armazenam arquivos de diversos formatos, beneficiando pesquisadores, instituições e sociedades científicas (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2018a). Os RDs podem ser institucionais, multidisciplinares, disciplinares, nacionais e são orientados por uma série de modelos de referência e princípios, como por exemplo, o princípio *FAIR* (*Findable, Accessible, Interoperable, Reusable*) (ROCHA *et al.*, 2018). No contexto da ciência aberta, os RDs intitulam os tipos de provedores de dados que gerenciam as informações da ciência e formam um tipo alternativo de comunicação científica. Esses repositórios de acesso aberto podem ser considerados um tipo de biblioteca digital (institucionais, temáticos, disciplinares, por área de estudo, de teses e dissertações) constituindo um sistema de informação científica, em ambiente digital, que reúne, armazena, organiza, preserva, recupera e dissemina a informação (LEITE, 2009). Configurando-os, então, como um mecanismo de compartilhamento.

Os RDs podem ser desenvolvidos a partir do conceito de base de dados relacional, que se trata de uma coleção de dados integrados num Sistema de Informação (SI) altamente estruturado, que permite várias solicitações por parte dos usuários (CALDEIRA, 2009). Os SIs são mecanismos facilitadores que – por meio de recursos humanos, *hardwares*, *softwares* e tecnologias de rede – recebem, processam, armazenam, controlam e convertem dados em informação a fim de disseminá-las em uma organização (O'BRIEN, 2004; DAMASCO, 2015). Conforme Tarapanoff (1995, p. 126). SIs são “conjunto de processos de coleta, processamento, armazenamento e recuperação de informação que têm como objetivo principal a organização de dados bibliográficos para posterior consulta e recuperação pelos usuários interessados”. Os SIs são compostos por dois subsistemas: um social, que compreende pessoas, processos, dados, informações e documentos; e um automatizado, que engloba máquinas, computadores e redes de comunicação que interligam os elementos do subsistema social. Existem vários tipos de SI, cada um com seus propósitos e objetivos (REISSWITZ, 2008).

Numa perspectiva brasileira no que tange a SIs para dados ecológicos, Barros (2005) entende que pesquisadores da área poderiam ter maior interação se possuíssem sistemas de informação que facilitassem a publicação e o acesso a dados científicos através de serviços especializados, atendendo a suas preferências e necessidades. O autor relata que os estudiosos da

biodiversidade brasileira passaram a interagir e promover o intercâmbio de conhecimento de maneira mais democrática e em maior escala nos últimos anos. No entanto, até aquele momento não existiam sistemas adequados para coletar, compartilhar e analisar esses dados. Os sistemas disponíveis eram limitados e não integravam dados como os de localização e de biodiversidade, o que tornava as análises demoradas e, por vezes, incompletas. De forma geral, as informações eram armazenadas em fontes textuais não estruturadas. Desse modo, uma infraestrutura computacional aberta que respeita as particularidades de cada fonte e de cada dado; e sistemas de informação que possibilitem a busca e a interpretação, tanto textual quanto geográfica, tornavam-se fundamentais para facilitar as atividades dos pesquisadores da ecologia. Os SIs em biodiversidade devem manipular dados heterogêneos e serem capazes de gerenciar grandes bancos de dados referentes às espécies, como por exemplo, informações de quando e onde elas foram observadas, por quem e como, incluindo dados que caracterizam o ecossistema onde cada espécie foi observada e a sua distribuição espacial.

Nessa perspectiva, Barros (2005) propôs uma Biblioteca Digital Georreferenciada reconhecida pela sigla BDiG-PELD para integração de dados dos sítios² do Programa Pesquisas Ecológicas de Longa Duração (PELD), uma solução prática e econômica para integração de dados gerados pelas investigações. Primeiramente, explica-se que o programa brasileiro PELD é financiado pelo CNPq e tem como objetivo “promover a organização e a consolidação do conhecimento existente sobre a composição e o funcionamento dos ecossistemas brasileiros, gerando informação e ferramentas para avaliar sua diversidade biológica” (BARROS *et al.*, 2005, p. 176). O PELD consolidou os estudos ecológicos no Brasil, impulsionando a revisão sistemática de sítios que representam a diversidade ecológica brasileira, a fim de identificar tendências naturais. O Programa é constantemente avaliado por meio de atividades periódicas que permitem o compartilhamento de dados e dos conhecimentos gerados em cada um dos sítios de pesquisa contemplados. Para compreender a dinâmica complexa das alterações ecológicas é fundamental um acompanhamento sistêmico e integrado de diferentes indicadores ecológicos e ambientais, em um curso temporal lento. O que demanda que pesquisadores da área se comuniquem, compartilhando dados e experiências (TABARELLI *et al.*, 2013).

Retomando a proposta de Barros (2005), o ambiente proposto para integração de dados dos vários sítios do PELD utiliza o protocolo *Open Archives Initiative Protocol for Metadata*

² Conjunto de locais de amostragem de um projeto. Área (ou sítio) de referência que a pesquisa é desenvolvida.

Harvesting (OAI-PMH) para coleta de dados e posterior inclusão na BDIG-PELD. A biblioteca digital possui interface de entrada de dados descentralizada, a partir dela os dados coletados em campo são publicados na *Ecological Metadata Language (EML)* – linguagem própria para descrever dados ecológicos. A interface oferece serviços de busca textual, navegação com facilidades de georreferenciamento e combina arquivos textuais, bancos de dados e repositórios XML no padrão *Open Archives Initiative (OAI)*. O padrão *EML* foi criado por pesquisadores da ecologia, do *National Center for Ecological Analysis and Synthesis (NCEAS)* e da rede *International Long Term Ecological Research (ILTER)* no início dos anos 1980. O *EML* visa à catalogação histórica de dados de natureza ecológica com ênfase nos aspectos: geográfico, data, hora, taxonomia, metodologia, dados gerais e específicos. O georreferenciamento relaciona documentos, conjuntos de dados, informações biológicas e espécies de seres vivos a locais geográficos através de nome de lugares ou de coordenadas de latitude e longitude.

Pavão *et al.* (2018a) identificaram alguns repositórios na área temática da ecologia mantidos ou nos quais havia participação de instituições brasileiras. Entre os citados estavam o *PPBio Data Repository*, promovido pelas instituições: Centro de Estudos Integrados da Biodiversidade Amazônica (INCT-CENBAM), *Data Observation Network for Earth (DataONE)* e *Long-Term Ecological Research Network (LTER)*. Ligado às temáticas: Ciências da Vida, Biologia Vegetal, Ecologia e Análise de Ecossistemas, Ecologia Animal, Biodiversidade e Pesquisa de Ecossistemas, Zootecnia, Botânica e Zoologia. O *PPBio Data Repository* suporta diferentes formatos de dados, como por exemplo, imagens, audiovisuais, texto estruturado, texto simples, formatos de dados científicos e estatísticos e dados brutos. O relatório também apontou o *Global Climate Data (WorldClim)* “um conjunto de dados climáticos globais (grades de clima) com uma resolução espacial de cerca de um quilômetro quadrado” (PAVÃO, 2018a, p. 11). Promovido pelas instituições: Centro de Referência em Informação Ambiental (CRIA), *University of California*, *Museum of Vertebrate Zoology*, *Ashoka Trust for Research in Ecology and the Environment*, *National Herbarium of the Netherlands* e *Rainforest CRS*. Focado nas temáticas Ciências Naturais Climatologia, Ecologia, Oceanografia, Geociências e Geografia, suporta os seguintes formatos de dados: gráficos estruturados, dados científicos e estatísticos e dados brutos. A referência também cita o repositório de dados PELD (PAVÃO, 2018a).

Num panorama internacional, Michener (2006) aponta o aumento da consciência sobre a importância dos metadados por pesquisadores da ecologia, com o objetivo de facilitar o

gerenciamento e a compreensão dos diversos e complexos dados coletados como parte das investigações ecológicas. Os metadados são definidos como "informações sobre os dados". Ou seja, as informações necessárias para compreender os dados, incluindo o conteúdo do conjunto de dados, contexto, qualidade, estrutura e acessibilidade. Resumindo, os metadados descrevem “quem, o quê, quando, onde e como” sobre cada aspecto dos dados (MICHENER *et al.*, 1997; MICHENER, 2000).

Conforme os autores, a abordagem mais indicada e útil para caracterizar os dados heterogêneos é o uso de metadados estruturados, como por exemplo, a *Ecological Metadata Language (EML)* e o *Biological Data Profile*. Esse último se trata de um conjunto comum de terminologias e definições para documentar dados da área. As especificações formais de metadados fornecem orientação para descrever consistentemente objetos de dados e tipos de dados (por exemplo, métodos, unidades de medição e detalhes do projeto experimental) (REICHMAN; JONES; SCHILDHAUER, 2011).

De volta à Michener (2006; 2015), o autor destaca que os bancos de dados ecológicos continuam crescendo em volume, amplitude e complexidade. Apresentam nível superior nas descrições de dados (metadados), nas informações derivadas de dados subsequentes, no processamento e análises (meta-informação no sentido mais amplo) e no movimento essencial para compreender e usar os dados cada vez mais complexos e volumosos. Esses dados podem ser usados para esforços de síntese, como por exemplo, verificar os resultados da pesquisa e viabilizar a reprodutibilidade de estudos. Os dados ecológicos também são empregados fora do contexto de pesquisa e são usados para a tomada de decisões, gestão de recursos naturais, educação e outros fins.

A seguir alguns exemplos citados por Coleman (2010) e Michener (2015), que demonstram como as práticas de compartilhamento de dados evoluíram por meio dos programas de pesquisa ecológica ao longo do tempo. Primeiramente menciona-se o *International Biological Program (IBP)*, inicialmente organizado em 1962 por Conrad Hal Waddington – biólogo desenvolvimentista britânico, oficialmente executado entre 1964 e 1974 para explorar "A Base Biológica da Produtividade e do Bem-Estar Humano". O *IBP* tinha como objetivo coordenar estudos ecológicos e ambientais em grande escala, com um escopo multidisciplinar que cobria uma ampla gama de biomas. Além disso, contemplou grandes esforços multinacionais para entender o ecossistema padrão e os processos da área (MICHENER, 2015).

Em 1980, foi estabelecida a *U.S. Long-Term Ecological Research (LTER)* a partir do financiamento da *U.S. National Science Foundation*. Nos dias de hoje, a *LTER* ainda mantém suas atividades de descoberta ecológica sobre a influência de fenômenos de longo prazo e em grande escala. A Rede consiste em 28 programas de pesquisa que se formam em torno do núcleo de pesquisas de longo prazo já existentes e podem contar com parcerias de universidades, organizações sem fins lucrativos ou outras agências federais. Esses programas atendem uma variedade de locais nos Estados Unidos, como também no Alasca e na Antártica. Os programas envolvem uma diversidade de pesquisadores, incluindo não só ecologistas microbianos, comunitários e paisagísticos, mas também hidrólogos, geoquímicos, cientistas sociais, economistas e até mesmo artistas, historiadores ou filósofos, ocasionalmente (*LTER*, 2021).

Em 1993, foi organizada a *International Long Term Ecological Research Network (ILTER)*, uma rede de trabalho que abrange centenas de sítios de pesquisa localizados em uma ampla gama de ecossistemas, que podem ajudar a compreender as mudanças ambientais em todo o mundo. Até hoje, a Rede contribui para a solução de problemas ecológicos e socioeconômicos internacionais por meio de pesquisas colaborativas para comparar dados de uma rede de locais e detectar tendências globais. A visão da *ILTER* é um mundo no qual a ciência ajude a prevenir e resolver problemas ambientais e socioecológicos. O seu principal propósito é fornecer uma rede globalmente distribuída e uma infraestrutura de locais de pesquisa de longo prazo para uso nas áreas de ecossistema, biodiversidade, zona crítica e pesquisa socioecológica, a fim de garantir serviços interoperáveis de alta qualidade em estreita interação com infra-estruturas e redes de investigação regionais e globais (*ILTER*, 2021, [on-line]).

Em 1995, foi criado o *National Center for Ecological Analysis and Synthesis (NCEAS)*, uma filial de pesquisa independente da Universidade da Califórnia, nos Estados Unidos. O *NCEAS* visa promover o conhecimento ecológico por meio de colaboração, síntese e compartilhamento de dados a partir de uma rede de colaboradores com impacto global. Desde então, a abordagem acelerada da descoberta científica aplicada pelo *NCEAS* o tornou uma das instituições científicas mais produtivas do mundo, com publicações influentes em uma diversidade de tópicos, como mudanças climáticas, doenças infecciosas, serviços ecossistêmicos, ecologia marinha e conservação. Como líder em ciência de dados ambientais, o *NCEAS* conduz o desenvolvimento de soluções tecnológicas para melhorar o acesso, o gerenciamento e a eficiência das análises de dados, ajudando a definir padrões científicos éticos para a comunidade. Esse

trabalho, também tornou o Centro pioneiro na promoção da ciência aberta na ecologia. Atualmente, o trabalho se estende além do conhecimento científico para a prática nas esferas de políticas públicas e gestão de recursos. Com um foco cada vez maior na aplicação da ciência a soluções eficientes na sociedade, o *NCEAS* estabelece parcerias com organizações para apoiar esforços locais a fim de construir um mundo no qual as pessoas e a natureza prosperem. O *NCEAS* não possui um corpo docente permanente, as práticas são promovidas por meio da natureza dinâmica dos cientistas ambientais e ecologistas residentes e visitantes que participam dos grupos de trabalho, junto à pós-doutores e bolsistas. Essa rede também ajuda a promover uma cultura de colaboração na comunidade científica em geral (*NCEAS*, 2021, [on-line]).

Em 2001, na Dinamarca, foi estabelecida a *Global Biodiversity Information Facility (GBIF)*, uma rede internacional de infraestrutura de dados financiada por diversos governos de todo o mundo, que visa fornecer a qualquer pessoa, em qualquer lugar, acesso aberto aos dados sobre todos os tipos de vida na terra. Em 2007, o portal de dados globais foi lançado a fim de facilitar o compartilhamento de dados e informações sobre biodiversidade. O *GBIF* surgiu de uma recomendação feita em 1999 no relatório do Subgrupo de Informática para a Biodiversidade do Fórum de Mega-ciência da Organização para Cooperação e Desenvolvimento Econômico. O documento concluiu que era “necessário um mecanismo internacional para tornar os dados e informações sobre biodiversidade acessíveis em todo o mundo”. Essa ação poderia trazer benefícios econômicos e sociais, além de permitir o desenvolvimento sustentável ao fornecer evidências científicas sólidas. Nos dias de hoje, o *GBIF* ainda mantém suas atividades oferecendo acesso livre e aberto a dados de biodiversidade em uma plataforma que une ferramentas que facilitam o processo de inclusão e busca. Junto a isso, disponibiliza artigos revisados e a lista de instituições publicadoras. Os publicadores fornecem acesso aberto aos seus conjuntos de dados usando definições de licença *Creative Commons* legíveis por máquinas, permitindo que cientistas, investigadores e outros apliquem os dados em centenas de publicações revistas por pares e documentos de apoio. Muitas das análises – que cobrem tópicos desde os impactos da mudança climática e a disseminação de pragas invasivas e exóticas até prioridades para conservação e áreas protegidas, segurança alimentar e saúde humana – não seriam possíveis sem esse mecanismo de compartilhamento (*GBIF*, 2021, [on-line]).

Também em 2001, nos Estados Unidos, foi formado o comitê internacional do *Ocean Biogeographic Information System (OBIS)* para fornecer acesso a dados e informações sobre

biodiversidade marinha em todo o mundo. O Sistema é uma base global de dados e informações de acesso aberto sobre a biodiversidade marinha para a ciência, conservação e desenvolvimento sustentável. Desde 2009, foi adotado como um projeto do programa *International Oceanographic Data and Information Exchange (IODE)* e da *Intergovernmental Oceanographic Commission (IOC)*. Em 2014, formou parceria com a *Global Biodiversity Information Facility (GBIF)* com o objetivo de melhorar o volume e a qualidade das informações sobre biodiversidade disponíveis aos formuladores de políticas para a conservação e uso sustentável dos recursos biológicos do oceano. Atualmente, o *OBIS* conta com mais de 20 núcleos ao redor do mundo, que conectam 500 instituições de 56 países. Coletivamente, eles forneceram mais de 45 milhões de observações de cerca de 120.000 espécies marinhas, de bactérias a baleias, da superfície a 10.900 metros de profundidade, e dos trópicos aos pólos. A base de dados oferece um conjunto de dados integrados para que seja possível pesquisar e mapear todos eles por nome de espécie, nível taxonômico superior, área geográfica, profundidade, tempo e parâmetros ambientais. Como visão o *OBIS* busca ser a porta de entrada mais abrangente para a biodiversidade dos oceanos do mundo e dados biogeográficos e informações necessárias para lidar com as preocupações costeiras e oceânicas mundiais. Sua missão é construir e manter uma aliança global que colabore com as comunidades científicas, para facilitar o acesso livre e aberto aos dados de biodiversidade, biogeográficos e às informações sobre a vida marinha (*OBIS*, 2021).

Em 2002, foi estabelecida a *speciesLink*, rede internacional que promove o acesso livre e aberto a dados, informações e ferramentas para qualquer indivíduo ou grupo, uma vez que os provedores de dados expressaram seu compromisso em compartilhar somente dados não confidenciais. O objetivo da rede é fomentar a pesquisa, a educação e a formulação de políticas para promover a conservação e o uso sustentável da biodiversidade. Os provedores de dados não garantem a precisão dos dados servidos através da rede e seus serviços. Assim sendo, o uso dos dados acessados por meio do portal e serviços *web* é de responsabilidade do usuário. As restrições indicadas por cada provedor de dados devem ser observadas e respeitadas, e suas fontes citadas de acordo com o texto de citação da rede *speciesLink* (*SPECIES LINK*, 2021).

Em 2009, surgiu a *Ocean Observatories Initiative (OOI)*, uma iniciativa de observação do oceano orientada para a ciência que fornece dados em tempo real, a fim de abordar questões científicas críticas sobre os oceanos do mundo. A *OOI* consiste em cinco matrizes marinhas que coletam continuamente dados sobre as condições do oceano. Essas matrizes são equipadas com

cerca de 800 instrumentos – de 36 tipos diferentes – medindo mais de 200 parâmetros. As matrizes reúnem dados físicos, químicos, geológicos, biológicos, da interface ar-mar com o fundo do mar, fornecendo uma riqueza de informações para pesquisa e educação. Os dados da *OOI* estão disponíveis *on-line* gratuitamente para qualquer pessoa com uma conexão à Internet (MICHENER, 2015; *OOI*, 2021).

Em 2019, foi concluído o observatório da *National Ecological Observatory Network (NEON)*. Projeto esse iniciado a partir do debate entre ecologistas e biólogos, em 1999, para o desenvolvimento da *Biodiversity Observation Network (BON)*. A partir dos anos 2000, a ideia gerou uma "rede de observatórios ecológicos" mais abrangente que abordava questões de escala continental. No presente, a *NEON* é operada por *Battelle* e responsável por coletar e armazenar dados ecológicos de longo prazo para acesso aberto, a fim de entender melhor como os ecossistemas dos Estados Unidos estão mudando. Hoje, engloba 81 locais de campo para observação, incluindo 47 locais terrestres e 34 de campo aquático de água doce. Junto a isso, uma plataforma de observação aerotransportada, bem como uma infraestrutura de informações para reunir dados e metadados de sensores e amostragem de campo, garantir a qualidade dos dados, processar as informações em produtos de dados e entregar esses produtos aos usuários por meio de um portal *on-line* (*NEON*, 2021). Ambos os observatórios ambientais, *NEON* e *OOI*, são financiados pela *U. S. National Science Foundation (NSF)* e foram projetados para serem totalmente operacionais.

Como visto, as ações que promovem o compartilhamento de dados ecológicos são variadas. Porém, como observa Michener (2015), essas práticas têm evoluído lentamente desde 1950. Segundo o autor, o compartilhamento de dados aumentou em resposta às mudanças socioculturais e à disponibilidade de tecnologias de informação. No entanto, apesar dessas melhorias, os desafios permanecem já que a maioria dos pesquisadores não fornece metadados suficientes para apoiar a interpretação e a reutilização dos dados. Junto a isso, freqüentemente cometem erros em relação à forma como os dados são organizados e documentados. Peters *et al.* (2014) observou que os ecologistas estão despreparados para lidar com questões de escala regional a continental devido à falta de cultura de compartilhamento de dados, métodos de coleta não padronizados e desatenção para documentar a proveniência. Sendo assim, ainda existem muitos desafios e oportunidades para o avanço de uma ciência mais aberta na ecologia. Os desafios da infraestrutura cibernética incluem uma escassez de sistemas de gerenciamento de

metadados fáceis de usar, dificuldades significativas na avaliação da qualidade e proveniência dos dados, além de uma ausência de abordagens analíticas e de visualização que facilitem a integração e harmonização dos dados (MICHENER, 2015).

Na sequência, explana-se o conceito de usuário, porque esse grupo é entendido como público, benefícios de torná-lo o centro das estratégias e seu entendimento como persona.

2.2 USUÁRIOS COMO PÚBLICO E SEU ENTENDIMENTO COMO PERSONA

Antes, é importante lembrar sobre os diferentes perfis dos grupos envolvidos no contexto da ciência aberta. Como apresentado no subcapítulo 2.1, essa ciência é um movimento que abrange uma série de pautas, como o acesso aberto e os dados abertos; e envolve diferentes coletivos, os quais podem exercer diversos níveis de intervenção no contexto. Alguns exemplos desses grupos são cientistas, políticos, cidadãos, fornecedores de plataforma, o pesquisador individual, instituições de ensino e pesquisa, editoras públicas e privadas de publicações acadêmicas (periódicos científicos), agências de financiamento, avaliação e certificação (FRIESIKE; FECHER, 2014; ALBAGLI, 2019).

Corroborando com as justificativas já apresentadas, essa pesquisa optou em focar no usuário para atingir um conhecimento mais profundo sobre esse, já que entende que esse coletivo ainda pode ser composto por diversos perfis. A decisão foi motivada pela questão norteadora dessa tese e pela compreensão de que as estratégias pensadas com foco no usuário estimulam a geração de valor para o movimento brasileiro de acesso aberto aos dados de pesquisa, tornando serviços e plataformas mais descomplicadas e acessíveis. Junto a isso, Sayão e Sales (2015) mencionam que é importante considerar num plano de gestão de dados quais são os usuários potenciais para os dados de pesquisa a serem compartilhados. Para os autores delinear e refletir sobre esses perfis "ajuda a determinar o repositório de dados mais apropriado para arquivar a sua coleção de dados." (SAYÃO; SALES, 2015). Desse modo, subentende-se que identificar e mapear esse público é essencial para estabelecer uma comunicação ativa entre os promotores das iniciativas e esse grupo, visando ao alcance de objetivos e metas em comum.

No campo de estudo da Ciência da Informação usuário é definido como "aquele que realiza uma pesquisa nos documentos para obter uma informação determinada" (LIMA, 1989, p. 165). Inicialmente, são divididos em dois grandes grupos: os reais e os potenciais. O primeiro

utiliza frequentemente a informação e o segundo é composto pelos perfis de usuários que demandam, mas não percebem a necessidade do serviço. Esses dois grandes grupos (real e potencial) são formados por uma variedade de perfis de usuários. Nesse universo, podem estar incluídos: uma única pessoa, grupos formados por categorias de profissionais ou uma organização inteira, por exemplo. Desse modo, as necessidades de cada perfil podem se desenhar dessemelhantes, conforme suas características, suas instituições e as comunidades em que estão inseridos (LARA, 2003; CONTI, 2003; MONFASANI; CURZEL, 2006; ROZADOS; PIFFER, 2009).

O estudo de usuários analisa aspectos e características da relação estabelecida entre informação-usuário, podendo ser centrada nos serviços oferecidos pela organização/instituição ou no próprio usuário. Essa última propõe conhecer hábitos, comportamentos, atitudes, opiniões, expectativas, desejos, necessidades, demandas e satisfação dos sujeitos. Como finalidade se pode ter: identificar as necessidades dos usuários para tomada de decisões; avaliar a efetividade de um serviço; analisar a interação do usuário com o serviço; identificar as características gerais desse usuário; prever o comportamento dele; apoiar estudos científicos e comparativos (TARAPANOFF, 1995; HERNÁNDEZ SALAZAR, 1997).

No campo de estudo de RP o grupo de usuários pode ser identificado como um dos diversos públicos que compõem o ambiente relacional de um sistema composto por diferentes organizações e coletividades que promovem iniciativas de acesso aberto aos dados de pesquisa. Esse grupo se caracteriza como público uma vez que os diferentes indivíduos que o formam: a) possuem interesses em comum em relação ao serviço prestado/produto ofertado; b) são afetados e afetam as ações das promotoras, na conquista de suas missões, objetivos e metas; c) variam conforme complexidade e potencial de poder que podem exercer ante os interesses das iniciativas, com variada probabilidade de promover ou dificultar as ações. Neste momento é importante explicar que esta tese entende organização como um agrupamento planejado de pessoas que desempenham funções e trabalham conjuntamente para atingir objetivos comuns (KUNSCH, 1986) e um conjunto de organizações que age para um determinado fim, forma um sistema (MATTOS, 1978). Portanto, o grupo de usuário é pensado como público em função desse sistema de organizações que promovem iniciativas em defesa do acesso aberto aos dados de pesquisa no Brasil. Como por exemplo, os esforços da RNP, do IBICT e do GT – RDP Brasil. Como afirmava Kunsch (1986), confia-se que ainda é crescente a criação das mais diversas e

inovadoras formas de organização para atender as demandas sociais e mercadológicas do mundo contemporâneo. Nessa perspectiva o relacionamento entre essas organizações e os públicos deve ocorrer dentro de um sistema social sustentado por processo de comunicação que possui a informação como matéria-prima e é baseado em sete dimensões: cultural; econômica; política; ideológica; histórica; jurídica e filosófica. O processo comunicativo se configura numa linha tênue de cooperação e conflito entre as partes, de modo não-linear (SIMÕES, 2001). Nesse sistema social, os coletivos estabelecem relações de influência e de poder sobre projetos, produtos e serviços desenvolvidos (MARTINS, 2020).

Assim sendo, compreende-se que o conjunto de iniciativas e agentes envolvidos na promoção do acesso aberto representa um sistema de organizações, o que torna o contexto dos relacionamentos mais desafiador. São coletividades, por vezes representadas por instituições, associações, organizações não-governamentais, etc., que possuem interesses em comum, mas a articulação em rede entre essas ainda parece incipiente. À vista disso, pode-se dizer que os diferentes agrupamentos integram um sistema social, mas esse ainda não é sustentado por um processo comunicativo planejado, o que cria barreiras para conexão e colaboração entre as partes.

Nessa conjuntura, considera-se que o grupo de usuários deve ser compreendido como um público estratégico, parte ativa durante o desenvolvimento das iniciativas de acesso aberto aos dados de pesquisa, influenciando políticas e resultados das organizações envolvidas na promoção. Os usuários formam um público estratégico de sistemas de informação, uma vez que esses devem adaptar constantemente seus arquivos e dinâmicas de acesso conforme as necessidades de uso do serviço. Essa adaptação é baseada em investigação e monitoramento dos usuários, a fim de garantir a satisfação dos diferentes perfis e a efetividade das propostas (MONFASANI; CURZEL, 2006). O termo *user-centered design (UCD)* foi criado por Donald Norman (1990)³ para definir uma filosofia baseada nos interesses e nas necessidades do usuário com o propósito de gerar mais valor aos produtos e serviços, tornando-os mais inteligíveis e usáveis. Em 1999 surgiu a primeira e única normatização sobre *UCD*, instituída pela ISO 13407 ("*Human-centred design processes for interactive systems*") (ISO, 1999). A norma orienta um modelo de processo que envolve o levantamento dos requisitos da organização e dos usuários visando à compreensão e a especificação do contexto de uso, para que o produto ou serviço seja bem sucedido. Entender e especificar o contexto se refere à atividade de coleta de dados sobre os usuários. Identificar os

³ Psicólogo cognitivo (SOUSA; BERTOMEU, 2015).

requisitos da organização e dos usuários possibilita a criação de premissas e metas para o trabalho. Esses levantamentos orientam a produção das soluções e viabilizam a avaliação dos projetos (VASARA, 2003; PINHEIRO; SZANIECKI; MONAT, 2016).

Buscando estudar o grupo de usuários é possível adotar as concepções de personas – registros que retratam perfis de pessoas fictícias, tendo como base características verdadeiras de um conjunto de indivíduos. Para tanto, o próximo subcapítulo discorre sobre o compilado de referências a respeito dos conceitos, estudos e métodos da construção de personas.

2.3 PERSONAS

A seguir são apresentadas as diversas concepções existentes sobre as personas. Inicialmente, o conceito do principal mentor do tema, Alan Cooper⁴. Logo, as demais visões num compilado de publicações dos últimos dezoito anos. A sequência do texto dar-se-á pela publicação mais antiga, 2002, até a mais recente, 2019. Os autores apresentados são resultados da pesquisa bibliográfica realizada para essa investigação, cujos procedimentos são detalhados na seção 3 Metodologia.

Sucintamente, pode-se dizer que personas são registros que retratam perfis de pessoas fictícias, tendo como base características verdadeiras de um conjunto de usuários. Alan Cooper utiliza a técnica desde 1983, quando pela primeira vez, optou por entrevistar o público-alvo do serviço em que estava trabalhando com grande atenção ao usuário principal em potencial (SIDORKO⁵, 2019). Tanto o termo quanto a metodologia só vieram a ser promovidos pelo autor – Cooper, em 1998 – por meio da primeira edição do livro *The Inmates are Running the Asylum*. Na época, uma técnica revolucionária para modelar e simular usuários na etapa de pré-desenvolvimento dos projetos de tecnologia. As personas ganharam uma rápida popularidade na indústria de *software* devido ao seu poder e eficácia incomuns (COOPER, 2008). Atualmente, o termo e a metodologia de construção de personas são aplicados em diferentes campos de estudo, como por exemplo, computação, *design*, negócios e mercado, tecnologia da informação, engenharia; e de modo incipiente nas relações públicas – conforme veremos na sequência do

⁴ Fundador da Cooper, consultoria líder em *design* de interação. Criador da metodologia de *design* direcionado a objetivos, pioneiro no uso de personas como ferramentas práticas de *design* de interação para criar produtos de alta tecnologia. Norte-Americano.

⁵ Mestre em Interação Humano-Tecnologia, Faculdade de Ciências da Comunicação Programa de Mestrado em Interação Tecnologia Humano, Universidade de Tampere, Finlândia.

texto. Observa-se que a construção e o uso exatos da metodologia variam em detalhes e descrição nas diferentes áreas que a aplicam. No entanto, quaisquer que sejam os esforços, todos têm o mesmo objetivo que é colocar as necessidades do usuário em primeiro lugar ao desenvolver qualquer tipo de iniciativa (GUÐJÓNSDÓTTIR⁶, 2010).

Retomando Alan Cooper (2008), a proposta de utilizar personas surgiu diante do desafio de demonstrar o uso de um *software* de negócios desenvolvido por ele em 1975. Como muitos *softwares* da época, era difícil de usar, a sequência de instruções a serem seguidas e/ou executadas não apresentava capacidade de ser utilizada pelos usuários. Porém, somente em 1983, após o desenvolvimento de 15 aplicativos comerciais, Cooper começou a desenvolver uma abordagem mais eficaz para tornar o *software* mais amigável para uso. Isso ocorreu quando ele estava desenvolvendo um programa de gerenciamento de projetos. No início, o estudioso entrevistou oito colegas e conhecidos que provavelmente eram candidatos a usar o *software* – usuários em potencial. Baseado no objetivo principal do programa a ser desenvolvido, Cooper aprofundou a entrevista com uma das colegas do grupo investigado. Os resultados desse processo serviram de base para a primeira persona desenvolvida por Cooper, chamada por ele de primitiva. Conforme Cooper (2008) a construção de personas pode ser compreendida em um instante, mas pode levar meses ou anos para ser dominada. Até mesmo os *designers* de interação da Cooper Consultoria passam semanas de estudo e meses de prática antes de serem considerados capazes de criar e usar personas em nível profissional.

Cooper, Reimann⁷ e Cronin⁸ (2007) definem personas como modelos descritivos dos usuários que fornecem informações sobre como eles se comportam, como pensam, o que desejam realizar e por quê. Personas não são pessoas reais, são arquétipos compostos e baseados em dados comportamentais e motivacionais coletados de/com usuários reais através de entrevistas etnográficas. Trazem aspectos de consciência social e política, devem ser típicas e críveis, mas não estereotipadas. Ao usar personas, é possível desenvolver uma compreensão dos usuários em contextos específicos.

Um dos principais elementos que permitem que as personas sejam bem-sucedidas como modelos de usuário é que elas são personificações. Embora sejam retratadas como indivíduos específicos, representam grupos de usuários. Uma persona encapsula um conjunto distinto de

⁶ Doutor em Interação Humano-Computador, Suécia.

⁷ Áreas de atuação: Experiência do usuário e design estratégico.

⁸ Área de atuação: experiência do usuário.

padrões de comportamento em relação ao uso de um produto específico (ou atividades análogas, se um produto ainda não existir), que são identificados através da análise dos dados da entrevista e apoiados por dados quantitativos. Esses padrões, juntamente com motivações específicas ou objetivos, definem as personas (COOPER; REIMANN; CRONIN, 2007).

As personas também são chamadas de arquétipos de usuário, porque são compostas por aspectos observados em indivíduos em funções semelhantes durante uma fase de pesquisa. Arquétipos de pessoas não são o mesmo que estereótipos. Os estereótipos representam vieses e suposições dos pesquisadores, em vez de dados factuais. É um “lugar comum”, personas são “o ideal”. Outro termo utilizado como sinônimo de persona é perfil do usuário. Não há problema desde que o perfil seja realmente gerado a partir de dados etnográficos. Os perfis de usuário geralmente consistem em um nome e uma imagem anexada a uma breve descrição demográfica, juntamente com um parágrafo curto e fictício, descrevendo o tipo de carro que a pessoa dirige, quantos filhos, onde mora e o que faz para viver. É provável que esse tipo de perfil de usuário seja baseado em um estereótipo (COOPER; REIMANN; CRONIN, 2007).

Existem até seis tipos de personas, são elas: primária, secundária, suplementar, cliente, servidor e negativo. As personas primárias representam o público principal para o *design* de uma interface. As secundárias estão satisfeitas com a interface criada para a persona principal, mesmo possuindo necessidades adicionais específicas, são bem servidas pelo produto não afetando em sua concepção. As personas suplementares são a combinação das primárias e secundárias, ou seja, suas necessidades estão completamente atendidas com a solução criada. As personas dos clientes atendem às necessidades dos clientes, não dos usuários finais. Normalmente, as personas dos clientes são tratadas como personas secundárias. No entanto, em alguns ambientes empresariais, algumas personas de clientes podem ser personas primárias para sua própria interface administrativa. As personas servidas são um pouco diferentes dos tipos de personas já discutidos. Eles não são usuários do produto; no entanto, são diretamente afetados pelo uso do produto. Por exemplo, um paciente que está sendo tratado por uma máquina de radioterapia não é usuário da interface da máquina, mas ele é afetado, uma vez que o equipamento oferece uma análise adequada do paciente. Personas negativas são tipos específicos de usuários para os quais o produto não está sendo construído. Não são usuários do produto. É importante ressaltar que os autores defendem que só é possível identificar uma persona para cada um dos seis tipos. Se forem

identificadas/construídas mais de uma, compreende-se que o escopo do projeto é amplo demais (COOPER; REIMANN; CRONIN, 2007).

Os profissionais de *marketing* também são familiarizados com um processo semelhante ao do *design* para o desenvolvimento de persona. A principal diferença entre os processos é que o *marketing* utiliza como base: dados demográficos, canais de distribuição e comportamento de compra. Essas personas são utilizadas para pensar e planejar processos de vendas. Já o *design* busca dados sobre o comportamento e as motivações de uso, utilizando as personas na definição e no processo de desenvolvimento de produtos. Observa-se que mesmo com as diferenças, dados de mercado não são descartados por Cooper, Reimann e Cronin (2007). Esse tipo de informação ajuda a determinar o intervalo demográfico dentro do qual se deve enquadrar a hipótese da persona, atua como um filtro inicial para limitar o escopo de entrevistas com pessoas dentro dos mercados-alvo.

Para Grudin⁹ e Pruitt¹⁰ (2002) personas são pessoas fictícias, as quais têm nomes, semelhanças, roupas, ocupações, famílias, amigos, animais de estimação, posses, e assim por diante. Elas têm idade, sexo, etnia, escolaridade e *status* socioeconômico. Possuem suas histórias de vida, objetivos e tarefas. Não são "agentes" ou "atores" em um *script*, são pessoas. Os autores indicam que as fotografias das personas sejam imagens de pessoas reais e que os voluntários "amadores" são melhores que os modelos profissionais para fotografar.

Conforme Vasara¹¹ (2003) personas são uma técnica de *design* popular, usada para descrever as características dos usuários e orientar os objetivos dos projetos. A técnica se dispõe a descrever usuários fictícios, que recebem nomes, rostos e são apresentados em detalhe com ênfase em características, objetivos e tarefas. É importante compreender que somente as identidades das personas são inventadas, suas características e seus objetivos são baseados na pesquisa do usuário. Elas possuem nomes e imagem para adicionar realismo. Ou seja, as personas trazem à vida os tipos-chave de usuários, especialmente para aqueles que não participaram do levantamento de dados (VASARA, 2003).

⁹ Principal pesquisador de design da Microsoft e professor afiliado da University of Washington – Information School. Áreas de interação homem-computador e trabalho cooperativo suportado por computador.

¹⁰ Gerente de Pesquisa de Usuários da Divisão de Tablet e Celular da Microsoft Corporation. Área: pesquisa de usuário.

¹¹ Mestre em Ciência da Computação, pela Helsinki University of Technology, Finlândia.

Nielsen¹² (2004) classifica personas como a descrição de um usuário fictício. Indo ao encontro do que mencionam Cooper, Reimann e Cronin (2007), a autora também concorda que o *marketing* usa usuários fictícios há mais tempo que o *design* e que os objetivos de uso entre as duas áreas são diferentes. No *marketing* o arquétipo serve para desenvolver uma abordagem ativa para atingir os grupos-alvo, o *design* usa a persona para desenvolver produtos que atendam às necessidades e objetivos dos usuários. Em ambos os casos, é possível compreender que o usuário fictício funciona como meio para criar empatia e identificação nas equipes de profissionais envolvidos.

Chapman¹³ e Milham¹⁴ (2006), mesmo apresentando as limitações da metodologia, afirmam que as personas são um componente padrão no trabalho da experiência do usuário em muitas empresas de *software*. O método defende uma abordagem psicológica convincente para comunicar e representar os usuários, reunindo informações sobre as necessidades, os comportamentos e as preferências das pessoas. Isso possibilita criar descrições vívidas sobre indivíduos ficcionais.

Pruitt e Adlin¹⁵ (2006) definem personas como pessoas imaginárias que auxiliam na criação de produtos que pessoas reais gostam de usar. São descrições detalhadas de pessoas imaginárias construídas a partir de dados bem compreendidos e altamente especificados sobre pessoas reais. Personas são representações memoráveis de usuários que permanecem visíveis na mente daqueles que projetam e constroem produtos e serviços. Elas beneficiam a usabilidade, utilidade e apelo geral de produtos e serviços; simplificam os processos internos da equipe e permitem mais êxito nas tomadas de decisões de negócios (organização – clientes) melhorando os resultados. Os autores distinguem as personas orientadas por dados e as orientadas por suposição. As personas orientadas por dados são baseadas na pesquisa do usuário e, como tal sua validade pode ser alta. No entanto, quando há pouco tempo para coletar e analisar dados, personas baseadas em suposições são frequentemente usadas para garantir que haja um entendimento compartilhado em toda equipe. As personas baseadas na suposição geralmente refletem “clichês

¹² A principal especialista da Dinamarca em personas. Ph.D. na área de personas e cenários, possui mais de cinco anos de experiência no uso prático de personas para órgãos da indústria e do governo.

¹³ Pesquisador quantitativo de experiência do usuário para Chrome OS e Chromebook. Áreas: interação e visualização humano-computador e marketing.

¹⁴ Especialista em *design* centrado no usuário e pesquisa de mercado.

¹⁵ Presidente da Adlin, inc., empresa de estratégia de experiência do usuário e pesquisadora da Microsoft. Área: experiência do usuário.

de *marketing*” em vez das necessidades e características da população real de usuários-alvo. As personas podem ser reutilizadas, reencarnadas ou aposentadas, dependendo do projeto.

Pruitt e Adlin (2006) diferenciam usuários e clientes. Às vezes, as pessoas que vão comprar seu produto não são as mesmas pessoas que o usarão diariamente. Por exemplo, gerentes e/ou profissionais da área de Tecnologia da Informação muitas vezes são os responsáveis pela tomada de decisões sobre a aquisição de um *software* que todos usarão na organização. Embora essas pessoas não sejam os usuários finais do produto, eles são essenciais para o sucesso do mesmo, uma vez que influenciam na decisão de compra. Da mesma forma, os influenciadores – incluindo analistas de mercado, especialistas e revisores – podem influenciar positiva ou negativamente num produto ou serviço, antes mesmo que ele chegue ao comprador ou cliente.

Aquino¹⁶ e Filgueiras¹⁷ (2008) observam que mesmo sendo fictícias as personas são representações precisas de usuários-alvo de um projeto. “Possibilitam construir suposições e conhecimento sobre os usuários, criando uma linguagem comum, tornando mais significativa a comunicação da equipe de projeto sobre o usuário” (AQUINO; FILGUEIRAS, 2008, p. 3).

Para Miaskiewicz, Summer e Kozar (2008)¹⁸ uma persona representa um grupo de usuários que compartilha características comportamentais comuns. Usando uma narrativa, imagem e nome, a persona fornece um *design* vívido e específico sobre o conjunto de pessoas. Portanto, as personas devem ser desenvolvidas com base nos resultados de pesquisas com os usuários.

LeRouge¹⁹ *et al.* (2013) também concordam que personas são modelos conceituais de grupos de usuários-alvo. Elas podem servir para promover o entendimento que sustenta o *user-centered design* (UCD) durante o processo de desenvolvimento de um projeto – envolvendo as etapas de análise, *design*, desenvolvimento e implementação do mesmo. Os perfis visam facilitar o entendimento compartilhado entre a equipe do projeto criando um "contexto mutuamente entendido".

¹⁶ Pesquisador do Centro Universitário da Faculdade de Engenharia Industrial - FEI, Universidade de São Paulo, Brasil. Área: ciência da computação.

¹⁷ Pesquisadora da Escola Politécnica, Universidade de São Paulo, Brasil. Áreas: interação humano computador e Engenharia Industrial.

¹⁸ Pesquisadores da Escola de Negócios e do Instituto de Ciência Cognitiva, da Universidade do Colorado, Estados Unidos da América. Áreas: teoria, conceitos e paradigmas do design de interação; interação humano computador; livrarias digitais e arquivos; design centrado no usuário e Sistemas de Informação.

¹⁹ Departamento de Ciências da Decisão/Gerenciamento de TI - John Cook School of Business, Nomeação Conjunta de Gerenciamento e Política de Cuidados de Saúde - Escola de Saúde Pública, Universidade de Saint Louis, Estados Unidos da América. Área: tecnologia da informação.

Para Marshall²⁰ *et al.* (2015) personas são modelos descritivos de usuários arquetípicos derivados da pesquisa do usuário. Trata-se de uma ferramenta poderosa para entender e visualizar a diversidade de estilos de vida presentes em uma população. Os perfis são uma síntese de várias pessoas que compartilham objetivos, motivações, relacionamentos e comportamentos semelhantes para com um produto ou serviço. Os objetivos de usar essas representações de usuário final é incentivar o realismo e aumentar o engajamento dentro das equipes de trabalho.

Madsen²¹ *et al.* (2017) citam que personas são arquetipos fictícios de usuários. Elas agrupam sucintamente as informações de um grupo de sujeitos na forma de uma pessoa que é facilmente compreendida e fundamentada. Nessa perspectiva a criação e o uso de personas é uma técnica de pesquisa concreta, na qual um conjunto de dados qualitativos – sobre as metas e experiências dos usuários – é sintetizado em arquetipos. Essas personas são generalizações de características-chave e objetivos de usuários em potencial.

Sturmer²² *et al.* (2019a) debate, de modo exploratório, o conceito de personas junto às abordagens sobre públicos de interesse no contexto da comunicação organizacional. A discussão inicial, na perspectiva de RP, propõe fomentar o debate sobre como as personas podem contribuir para encontrar "caminhos mais específicos para identificar as características dos públicos com os quais uma organização se relaciona, atrelando tais características à estratégia de comunicação" (STURMER *et al.*, 2019a, p. 25). Os autores apresentam a compreensão de personas em duas perspectivas. A primeira é a da tecnologia, focada no usuário e baseada na concepção de Alan Cooper (2004). Divide-se em três categorias:

Marketing personas explicam o comportamento do cliente, mas não chegam ao porquê por trás dele. As marketing personas são boas para determinar quais tipos de clientes serão receptivos a determinados produtos ou mensagens, ou para avaliar o ROI²³ potencial de um produto. Não são, no entanto, eficazes no processo de definição de um produto ou serviço – o que é, como ele funcionará e como ele será usado, ou mesmo para priorizar recursos em um produto ou serviço [...]
Proto-personas são utilizadas quando não há dinheiro ou tempo para criar personagens verdadeiros baseados em pesquisa – eles são baseados em pesquisa secundária e na percepção da equipe de desenvolvedores a respeito do usuário para o qual eles destinam seu projeto. De acordo com o Cooper (2014), usar uma proto-persona para dirigir as decisões de design ainda é melhor do que não fazer qualquer uso da abordagem de

²⁰ Pesquisador da Escola de *Design* da Universidade de Loughborough, Reino Unido.

²¹ Associação Americana de Professores de Física, Cleveland, Estados Unidos da América.

²² Doutora em Comunicação pela Universidade Federal de Santa Maria (UFSM/2018); Professora e coordenadora dos cursos de Relações Públicas e Jornalismo na Universidade Feevale.

²³ ROI é a sigla para *return investment*, ou “retorno sobre investimento”, e mede o ganho ou perda gerada em um investimento relativo à quantia de dinheiro investido (STURMER, 2019a, p. 35).

personas, embora o autor ressalte a importância destes personagens serem validados através de pesquisas.

Design personas se concentram nos objetivos do usuário, no comportamento atual e nas dores e necessidades do usuário em oposição às suas preferências de compra ou de mídia e comportamento. Elas são baseadas em pesquisa de campo e pessoas reais. Design personas contam uma história e descrevem por que as pessoas se comportam de determinada forma, na tentativa de ajudar todos os envolvidos na concepção e construção de um produto ou serviço a entender, se relacionar e lembrar do usuário final ao longo de todo o processo de desenvolvimento do produto. Design personas são eficazes para comunicar os insights da pesquisa, ou seja, os aprendizados e dados obtidos, os objetivos do usuário, auxiliar na compreensão e foco em certos tipos de usuários e definir um produto ou serviço, além de evitar o surgimento do “usuário elástico” e o design auto referencial (STURMER, 2019a, p. 34 – 35).

A segunda perspectiva é a do *marketing*, explicada por meio dos conceitos de *buyer persona* e *inbound marketing*. O primeiro visa entender o comportamento e as necessidades dos consumidores com o objetivo de construir estratégias de comunicação e de *marketing* para estimular o consumo do serviço ou do produto ofertado. O *inbound marketing* – em português “marketing de entrada” ou “marketing de atração” – cria conteúdo projetado para os clientes considerados estratégicos. O objetivo é resolver os problemas e atender as necessidades desse público específico, a fim de estimular sua confiança e promover a credibilidade da organização, seus projetos, produtos e/ou serviços. O *inbound marketing* é baseado no *marketing* de conteúdo, ou seja, a produção de conteúdo é formada a partir dos interesses do consumidor.

Após a explanação dos conceitos nesse subcapítulo, pode-se dizer que, independente da área do conhecimento, os autores referenciados possuem concepções similares do que é persona: um arquétipo fictício que representa um grupo/aglomerado de usuários real ou em potencial para um produto ou um serviço. Esse perfil deve ter representações e descrições, mais completas e críveis possíveis, embasadas em dados coletados com pessoas reais a partir de técnicas qualitativas de pesquisa (em sua maioria).

2.3.1 Construção das Personas: mapeamento de estudos e métodos

A ideia de construir produtos e serviços com base no que as pessoas reais precisam e desejam parece óbvio, mas colocar os usuários (ou seja, informações sobre os usuários) realmente no centro do processo de planejamento é extremamente difícil. Primeiro não é natural estar centrado no outro. A tendência de comportamento humano é ser egocêntrico, ou seja, estar baseado nos próprios desejos e necessidades. Outras vezes, procuramos usuários que são como

nós para fornecer *feedback* sobre nossos projetos. Segundo, o grupo de usuários é formado por indivíduos complicados e variados. É preciso um grande esforço para entender suas necessidades, desejos, preferências e comportamentos. Terceiro, a comunicação entre os pesquisadores de usuários e os que projetam os produtos/serviços/ações deve ser efetiva e contínua para um real impacto das personas no projeto. Se as informações importantes sobre os usuários não estiverem disponíveis no momento certo, de difícil acesso ou ilegíveis, as equipes de desenvolvimento acabam baseando as ideias no que é mais fácil e menos importante, maior ou menor investimento para construir, por exemplo. Os aspectos citados resultam em produtos inadequados (PRUITT; ADLIN, 2006).

Várias abordagens foram sugeridas para construção de personas úteis nas diversas áreas de conhecimento que aplicam a metodologia. Em sua apresentação original, Cooper (1999) não detalha quais tipos de dados devem ser usados para criar personas e também não indica claramente até que ponto elas devem se basear na pesquisa do usuário. Porém, publicações posteriores apontam que as características e objetivos relevantes das personas devem estar fortemente embasados nos dados dos usuários reais (GOODWIN, 2002; VASARA, 2003).

Para Goodwin²⁴ (2002) a utilidade das personas no design de produtos e serviços interativos se tornou amplamente aceita nos últimos anos. Porém, a falta de publicações sobre quais as informações realmente compreendem uma persona, deixou espaço para muitos equívocos metodológicos. A autora propõe a coleta de dados comportamentais, observados através de entrevistas, como também de dados demográficos. Ao encontrar variáveis comportamentais similares ou idênticas, existe um padrão que forma a base das personas. A comunicação dos perfis é composta por texto que deve descrever o ambiente de uso em potencial do serviço ou produto; bem como um dia de trabalho típico da persona (ou outro período relevante); soluções e frustrações atuais; relacionamentos relevantes e objetivos. A narrativa é o formato adequado para transmitir atitudes, necessidades e problemas da persona em relação ao contexto pesquisado. Todos os aspectos da descrição devem ser vinculados aos dados reais. Caso isso não ocorra, a autora conclui que “não é uma persona – é um projeto de escrita criativa que não deve ser usado para tomar decisões críticas” (GOODWIN, 2002, [on-line]; MARTINS; VANZ, 2021).

²⁴ Executiva de design de produto. Palestrante e professora. Consultora de design, norte-americana. Escritora do livro *Designing for the Digital Age: How to Create Human – Centered Products and Services*.

Grudin e Pruitt (2002) propõem a união de métodos quantitativos e qualitativos para obter uma visão mais completa dos perfis, melhorar a seleção, enriquecimento e evolução de conjuntos de personas. Os dados são coletados a partir de técnicas etnográficas, pesquisas de mercado, entrevistas, observações e assim por diante, resultando numa série de documentos que descrevem/narram as personas. A figura 4 demonstra o exemplo de documento para formar e comunicar cada persona.

Figura 4: Índice para o documento básico da persona

<p>Visão geral - Patrick Blakeman (pequeno empresário) Conheça Patrick, sua empresa e família.</p> <p>Um dia na vida Siga Patrick por um dia típico.</p> <p>Atividades de Trabalho Veja a descrição do cargo e a função de Patrick no trabalho.</p> <p>Atividades domésticas e de lazer Obtenha informações sobre o que Patrick faz quando não está no trabalho.</p> <p>Metas, medos e aspirações Entenda as preocupações de Patrick sobre sua vida, carreira e o negócio.</p> <p>Habilidades, conhecimentos e habilidades de computador Saiba mais sobre a experiência de Patrick no computador.</p> <p>Tamanho e influência do mercado Entenda o impacto que pessoas como Patrick têm em nossos negócios.</p> <p>Atributos Demográficos Leia as principais informações demográficas sobre Patrick e sua família.</p> <p>Atributos de tecnologia Tenha uma ideia do que Patrick faz com tecnologia.</p> <p>Atitudes de Tecnologia Reveja a perspectiva de Patrick sobre tecnologia, passado e futuro.</p> <p>Comunicando Saiba como Patrick se mantém em contato com as pessoas.</p> <p>Considerações Internacionais Descubra como é Patrick fora dos EUA.</p> <p>Citações Ouça o que Patrick tem a dizer.</p>

Fonte: GRUDIN; PRUITT, 2002, p. 5 – tradução nossa.

Os estudiosos consideram que as características de cada persona devem ter como suporte os dados coletados na pesquisa com usuários, essa conexão deve estar explícita e saliente para tornar os perfis credíveis e visíveis – indo ao encontro do citado por Goodwin (2002) e Vasara (2003). A representação e a comunicação desses perfis devem ser multifacetadas, multimodais, contínuas e progressivas.

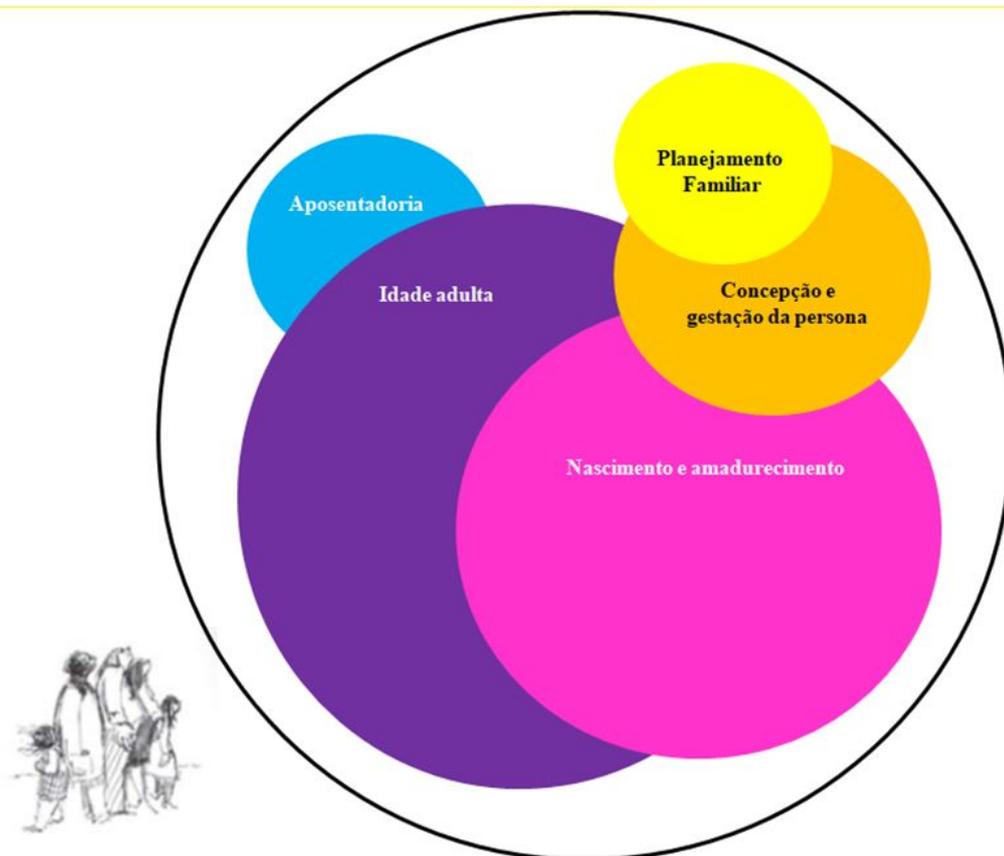
Vasara (2003) divide o processo de construção de personas em: 1) Pesquisa de base (*Background research* – inglês) ocorre antes de iniciar o estudo do usuário real, coleta dados em documentos, manuais, materiais de marketing e arquivos; 2) Pesquisa de usuário (*User Research* – em inglês) que utiliza a entrevista semiestruturada como principal método de coleta; 3) Realização de entrevistas mais específicas e breve observação de como os usuários utilizam o serviço ou produto investigado. Como instrumento de coleta o autor considera que é importante alguns dos pesquisados manterem um diário semanal, para descobrir mais sobre como eles usam o serviço ou produto em diferentes situações. Como variáveis a serem investigadas estão atitude em relação à tecnologia, conhecimento sobre a tecnologia e frequência de uso. As entrevistas e as sessões de observação são gravadas com equipamento de imagem e som; e para análise são descritas e transcritas. Para comunicação apresentam uma ficha com informações detalhadas sobre as personas construídas. Entre duas e três páginas com descrições dos perfis de usuário, narrando um dia de trabalho típico em suas vidas. As imagens para representar as personas são escolhidas em revistas com o objetivo de representar o perfil com imagens realistas do usuário-alvo (MARTINS; VANZ, 2021).

Chapman e Milham (2006) refutam a metodologia de construção personas, pois não há literatura que a avalie sistematicamente. Alegam que é imprecisa, já que não existe como verificar se o grupo de arquétipos representa apropriadamente a população de interesse. Para eles as personas estão fora do método científico, pois não houve um estudo adequado sobre a confiabilidade, validade ou utilidade da proposta metodológica. Junto a isso, sugerem que o uso de persona pode envolver as equipes em conflitos políticos sobre quem entende e representa melhor os/as clientes. Concluem que o método não deve ser considerado uma fonte de dados (MARTINS; VANZ, 2021).

Pruitt e Adlin (2006) apresentam o ciclo de vida da persona que promove a ideia de que os perfis devem estar na mente da equipe durante todo o processo de desenvolvimento do produto/serviço. Para tanto, a construção da persona se divide em cinco fases: 1) planejamento familiar (levantamento inicial de dados da organização/iniciativa para o desenvolvimento das personas); 2) concepção e gestação da persona (que se trata especificamente da construção dos perfis); 3) nascimento e amadurecimento (aborda o lançamento e a comunicação do perfil criado para a equipe de criação do produto); 4) idade adulta (que se refere ao uso/aplicação efetiva dos

perfis no processo de desenvolvimento do produto); 5) terminando com a aposentadoria (nessa fase são verificados quais dos perfis devem ser descartados ou reusados). Veja a figura 5:

Figura 5: As cinco fases do Ciclo de Vida das Personas



Fonte: PRUITT; ADLIN, 2006, p. 48 – tradução nossa.

O desenvolvimento da personalidade começa com o planejamento familiar. Essa é a fase de pesquisa e análise que precede a criação real de personas. Nesse momento o investigador se concentra em levantar informações sobre o contexto na perspectiva da organização/iniciativa.

Logo, se parte para a coleta de dados sobre o usuário real ou em potencial. Trata-se da fase de concepção e gestação da pessoa, a construção dos perfis. Durante essa fase é construído o esqueleto baseado nas informações obtidas na etapa anterior. O esboço permite identificar previamente quantas personas podem ser criadas e como priorizá-las no desenvolvimento da ideia.

Para tanto, Pruitt e Adlin (2006) consideram importantes as seguintes informações, veja a figura 6:

Figura 6: Estrutura de dados para a comunicação dos esqueletos

Menino, de 10 a 13 anos

Uso do computador na escola

- Tem acesso a um computador compartilhado em sua sala de aula ou um 'laboratório' de informática compartilhado por toda a escola
- Tem pelo menos uma tarefa no computador por semana
- Acha o uso do computador na escola "chato"

Uso de internet em casa

- Compartilha um computador doméstico com a família
- Usa internet para jogar e (às vezes) e para fazer trabalho escolar

Interesses / Atividades

- Gosta de conversar sobre jogos com os amigos
- Gosta mais de videogames do que de jogos de computador
- Participa de vários esportes organizados

Fonte: PRUITT; ADLIN, 2006, p. 208 – tradução nossa.

A figura 7 demonstra outra proposta que inclui uma fotografia para representar o esqueleto.

Figura 7: Esqueleto concluído de Oden, o "usuário ocasional" da *Microsoft Corporation*



Oden, o "usuário ocasional".

Como Oden usa X Docs?

- Oden é um planejador sênior que usa o X Doc, pois é o software da empresa;
- Ele usa formulários criados por Nick. Cara ou Dawn;
- Oden usa regularmente alguns formulários, como status de relatório, relatório de despesas ou avaliação de desempenho)

O que é importante para Oden?

- Para Oden, quanto mais fácil X Docs, melhor para ele.
- Ter documentos preenchidos anteriormente com despesas ou códigos de departamento é uma grande vitória porque ele precisa procurar informações continuamente;
- Oden quer instruções embutidas, dicas de ferramentas e maneiras fáceis para obter as informações de que precisa no formulários com precisão;
- Ele espera mensagens de erro para ajudá-lo a encontrar e resolver problemas;
- Oden às vezes trabalha em casa e viaja ocasionalmente, por isso ele deseja poder trabalhar offline;

Outras ferramentas comumente usadas: word e excel;

Oden representa um usuário final típico e posições semelhantes incluem gerentes de projeto, analistas e gerentes de departamento.

Fonte: PRUITT; ADLIN, 2006, p. 229 – tradução nossa.

A priorização se concentra nos objetivos imediatos ou nos resultados mais acessíveis. Priorizar não significa abandonar os interesses dos esqueletos de baixa relevância. Se as partes

interessadas compreendem que todos os esqueletos são importantes, dividida em três prioridades (alta, média e baixa) em relação à iniciativa. A priorização pode indicar de três a cinco principais esqueletos por valor de prioridade, a serem transformados em personas completas.

A quantidade e os tipos de dados coletados para a primeira e a segunda fase dependerão de quanto tempo e dinheiro estão disponíveis para a pesquisa. Os autores confiam que qualquer foco nos usuários é melhor do que nenhum (PRUITT; ADLIN, 2006). As fontes de dados são primárias ou secundárias. Como fontes de dados primárias podem ser consideradas qualquer levantamento feito a partir de observação direta dos usuários (comportamentos, ações, pensamentos ou sentimentos; também podem incluir pesquisas anteriores feitas com os grupos de usuários (reais ou em potencial). Fontes de dados secundárias são aqueles fornecidos por terceiros, como por exemplo, um colega de equipe distante que tem algum contato com os usuários, mas utiliza memória e experiência para fazer declarações ou inferências “fatuais” sobre esses usuários (PRUITT; ADLIN, 2006).

Os dados podem ser de tipo quantitativo, geralmente, provenientes de um grande número de usuários e coletados por meio de um método que promove eficiência e rigor (por exemplo, um questionário, entrevista telefônica estruturada ou análise de *logs* do servidor). Esses podem ser encontrados numa variedade fontes. Como exemplo, relatórios e documentos de pesquisas anteriores (se houver); artigos relevantes em periódicos profissionais, acadêmicos, da indústria e de negócios; artigos de jornais ou revistas especializadas; pesquisa governamental e institucional sobre a população; blogs, grupos de notícias, etc. (PRUITT; ADLIN, 2006).

Os dados do tipo qualitativos, geralmente, são provenientes de um número menor de usuários e coletados por técnicas de pesquisa que promovem a compreensão mais profunda do pesquisado. Entre essas estão visitas observacionais ao local ou pesquisa etnográfica, entrevistas individuais estruturadas ou não estruturadas; e grupos focais, que podem produzir dados qualitativos e/ou quantitativos, isso depende de como a pesquisa é conduzida (PRUITT; ADLIN, 2006).

Pruitt e Adlin (2006) listam um roteiro para a coleta de dados, composto por questões-chave que buscam a compreensão dos usuários: como são, quais outros interesses possuem, em que atividades se envolvem, quais comportamentos eles têm, quais ações eles tomam, quais são seus objetivos, aspirações e medos, quais conhecimentos e habilidades eles parecem ter? A segunda fase é a concepção e gestação, onde as personas são criadas a partir dos dados coletados

anteriormente. As informações são filtradas, organizadas e priorizadas para a construção dos perfis. Os perfis podem ser criados a partir de três categorias: com base nas funções do usuário; com base nos objetivos do usuário e com base em segmentos de usuário.

O tempo necessário para a fase de concepção e gestação vai depender do cronograma do projeto, da quantidade de dados coletados e dos objetivos. Exemplo, personas baseadas em suposições podem ser geradas em menos de um dia; ou em meses, pois depende da análise de uma quantidade considerável de dados. Na maioria dos casos, a criação das personas baseadas em dados dura duas semanas. Portanto, é útil planejar todas as etapas de concepção e gestação a fim de encurtar o tempo alocado para cada uma. Pruitt e Adlin (2006) sugerem considerar o tempo que as personas poderão ser úteis para o projeto para determinar (ou justificar) a quantidade de tempo que você gastará criando personas. Eles relatam que algumas personas podem ser úteis por vários anos.

Como cada projeto, produto e equipe são diferentes, não há um número exato de personas para criar. Para tanto, o tipo de produto que está construindo, a natureza do público de interesse, as informações levantadas e os detalhes das metas de negócio devem responder as seguintes questões: quantas personas eu preciso, quais personas eu preciso, quais personas devem ser primárias ou secundárias e como faço para usar personas secundárias sem projetar "para todos"

Pruitt e Adlin (2006) citam Cooper e Reimann (2003) ao defenderem a ideia de que se deve encontrar mais de um perfil primário. Alegam que criar apenas um perfil é ótimo na teoria, mas complicado na prática, pois é difícil convencer uma equipe executiva de negócios que todos os esforços de desenvolvimento devem visar uma única pessoa. Diferente dos autores, Cooper e Reimann (2003) compreendem que ter mais de um perfil primário comprova que o escopo do projeto é provavelmente muito amplo, o que pode prejudicar seu desenvolvimento.

A terceira fase é o nascimento e o amadurecimento, ou seja, a comunicação das personas. A descrição completa é composta por dados detalhados a fim de dar-lhes personalidade e contexto, incluindo fotos para dar vida às personas. No mínimo, o documento deve incluir as seguintes informações essenciais: nome; objetivos; papel; cargo; função; atividades; comportamentos; descrição do ambiente; habilidades; conhecimento e detalhes pessoais.

Embora apresentar os perfis possa parecer simples, geralmente não é. Para isso é preciso divulgar informações de forma progressiva e estratégica sobre cada persona. Para a apresentação

e comunicação dos perfis construídos Pruitt e Adlin (2006) citam casos reais da aplicação da metodologia e suas respectivas formas. Veja o exemplo na figura 8:

Figura 8: Exemplo de comunicação da persona construída para a Biblioteca Carnegie apresentada através de um Quadro de História e Narrativa



Naomi Masterson
 Não usuário
 Situação atual
 Frequência de visita: Nunca
 Busca de ajuda com a equipe: Não
 Gênero feminino

Naomi tem 39 anos, é casada e tem dois filhos. Ela trabalha em tempo integral como Diretora de Vendas da PhilTek Sistemas em Oakland. Ela está viajando, mas costuma ficar no escritório até tarde da noite. Ela mora com sua família em um loft renovado no lado sul da cidade. Ela gosta de comprar livros da Barnes & Noble e, ocasionalmente, alugar filmes da Blockbuster. Frequentemente, ela pede livros e DVDs online usando Amazon ou Netflix.

- Experiência atual em biblioteca: Naomi nunca foi à biblioteca e só pensa nela como uma instituição "bolorenta" desde a infância (ela acha que é improvável que eles tenham qualquer coisa "fresca". Ela também menciona que a biblioteca não está aberta durante as horas que ela gosta de pesquisar.

Objetivo Principal: quer o mais vendido
 Objetivo Secundário: conferir Cds ou DVDs e obter respostas sobre saúde e história da família

Fonte: PRUITT; ADLIN, 2006, p. 385 – tradução nossa.

Pruitt e Adlin (2006) consideram que o uso de banco de fotos/imagens para a representação de personas merece atenção, pois podem parecer muito elegantes, já que as pessoas neles tendem a ser modelos profissionais. Os autores sugerem à utilização de fotos que representem o mais fiel possível as pessoas reais. Exemplo, figura 9:

Figura 9: Persona criada para a organização G4K

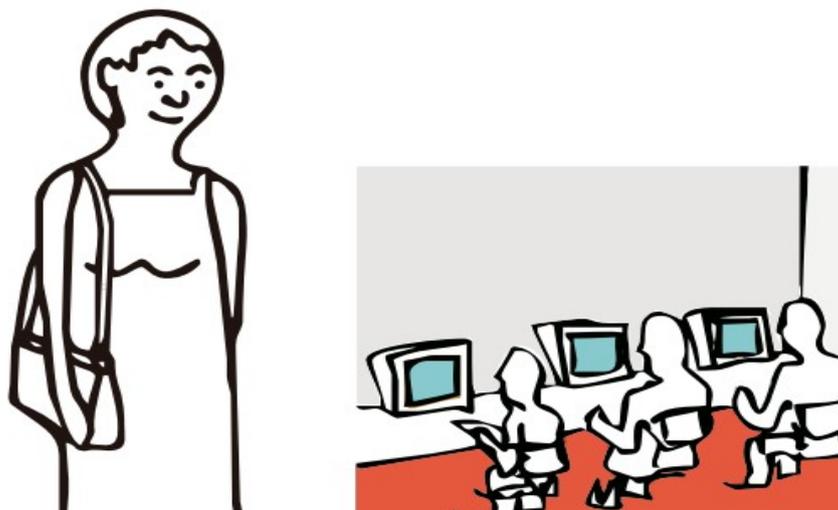


Colbi Chandler (a criança criativa) é uma charmosa garota de sete anos que adora fazer qualquer coisa imaginativa, engenhosa ou divertida. Ela gosta de ler e escrever e adora a escola. Colbi será uma visitante ocasional do site G4K, atraída por causa do conteúdo criativo e por causa de suas amigas que também visitam o site. Colbi é a irmã mais nova de Austin.

Fonte: PRUITT; ADLIN, 2006, p. 254 – tradução nossa.

A ilustração é considerada uma alternativa. A figura 10 representa Claire, criada para um projeto de design de portal para um cliente africano:

Figura 10: Persona criada para um portal africano



Fonte: PRUITT; ADLIN, 2006, p. 251.

Cooper, Reimann e Cronin (2007) mencionam que se o profissional não tiver conhecimento detalhado dos usuários para os quais está projetando, atingir os objetivos organizacionais que estão impulsionando as atividades tem pouca chance de sucesso, não importa o quão habilidoso seja esse. Não é possível obter informações reais sobre os usuários analisando números provenientes de um estudo quantitativo, como por exemplo, uma pesquisa de mercado (embora esse tipo de dado possa ser importante para responder a outros tipos de perguntas). O conhecimento mais profundo dos usuários só pode ser alcançado por técnicas de pesquisa qualitativa. Dedicar tempo à compreensão da população de usuários pode fornecer informações valiosas que não são reveladas por meio de pesquisas de mercado tradicionais.

A quarta fase é a idade adulta, onde as personas são inseridas no processo de desenvolvimento da iniciativa. As personas estão totalmente formadas e agora começarão a se desenvolver nas mentes da equipe. Porém, até agora existem poucos métodos documentados sobre essa fase e como ela de fato acontece. Sabe-se que é possível usá-las para planejar o produto ou serviço; explorar e avaliar soluções; lançar de produtos e serviços. A quinta fase é a de aposentadoria ou reutilização das personas criadas no projeto em andamento ou em similares.

É uma fase considerada crítica e muitas vezes ignorada pelas partes interessadas, pois na maioria dos casos o orçamento das iniciativas não prevê o término do ciclo de vida dos perfis (PRUITT; ADLIN, 2006).

Cooper, Reimann e Cronin (2007), consideram que existem muitos tipos de técnicas para desenvolver uma pesquisa qualitativa, cada uma das quais pode desempenhar um papel importante na compreensão do público estratégico. As consideradas mais úteis são: revisão de literatura; entrevistas com as partes interessadas; entrevistas com clientes e usuários; observação do usuário/estudos de campo etnográficos e após ter o protótipo de *software* desenvolvido, auditorias de produto.

A revisão da literatura é o levantamento de informações em fontes como: resultados e métricas de pesquisas com usuários, artigos de periódicos comerciais e técnicos, entre outros. Essa literatura guiará o desenvolvimento das perguntas para fazer entrevistas e posteriormente correlacionar com os dados compilados do usuário. A entrevista com as partes interessadas (*stakeholders em inglês*) pretende levantar aspectos importantes. Entre eles: expectativas, objetivos, capacidades, limitações do contexto técnico e orçamentário em torno do produto e; da visão que esse público tem do produto e dos potenciais usuários. Essas informações devem ser harmonizadas com as que serão coletadas dos usuários, a fim de guiar a construção das personas. Na perspectiva dos autores, partes interessadas são indivíduos, organizações ou instituições que promovem, coordenam e possuem algum tipo de autoridade e/ou responsabilidade perante o projeto. Entrevistar cada parte interessada isoladamente é mais eficaz, em vez de um grupo. A configuração individual promove franqueza por parte da parte interessada e garante que as visões individuais não sejam perdidas na multidão. As entrevistas duram no máximo uma hora, a não ser que seja identificado que uma das partes é uma fonte de informação excepcionalmente valiosa (COOPER; REIMANN; CRONIN, 2007).

Para Cooper, Reimann e Cronin (2007), usuários são os atores que efetivamente utilizam o serviço. Ao entrevistá-los, tem-se como foco levantar informações sobre seus objetivos; papel; processo de decisão para o uso; suas frustrações com as soluções atuais; problemas e vocabulário relacionados ao serviço. Os usuários devem ser o foco principal do esforço de *design*. São as pessoas que estão utilizando pessoalmente o produto ou serviço para atingir um objetivo e uma meta. Os usuários são pessoas que se encaixam como bons candidatos para uso no futuro, pois possuem necessidades que podem ser atendidas com o produto e ou serviço que está sendo

desenvolvido. A entrevista com o usuário busca compreender o contexto do produto ou do serviço na vida do mesmo (quando, por que e como é ou será usado), perspectiva do usuário, quais tarefas e atividades o produto atual precisa executar, metas e motivações para usar o produto, como os usuários pensam sobre seus trabalhos e atividades, chamado pelo autor de modelo mental; problemas e frustrações com produtos atuais (ou um sistema análogo, se não existir um produto atual). Independente da identificação, ambos os grupos devem ser entrevistados, pois cada um tem sua própria perspectiva sobre o produto. A intenção de entrevistar esse público é compreender seus objetivos, suas frustrações, motivações, inspirações, comportamento, problemas e vocabulário relacionados ao produto (COOPER; REIMANN; CRONIN, 2007).

A observação do usuário é outra técnica qualitativa considerada importante para Cooper, Reimann e Cronin (2007), pois entendem que a maioria das pessoas é incapaz de avaliar com precisão seus próprios comportamentos. A observação direta proporciona obter resultados mais ricos em relação aos aspectos levantados na etapa de entrevista com o usuário. Muitos profissionais de usabilidade utilizam recursos tecnológicos, como áudio ou vídeo para capturar o que os usuários dizem e fazem durante a entrevista. Porém, é preciso tomar cuidado para que essas tecnologias não sejam intrusivas.

Cooper, Reimann e Cronin (2007) entendem que talvez o mais eficaz para coletar dados qualitativos do usuário seja a combinação das entrevistas e da observação, o que chamam de entrevistas etnográficas. A técnica tem como objetivo entender os comportamentos e os rituais dos grupos de pessoas que interagem individualmente com produtos e serviços, um nível micro. A etnografia, em sua essência, tem como objetivo tentar entender comportamentos e rituais sociais de uma cultura inteira, um nível macro. As entrevistas etnográficas são aplicadas num processo de diversas etapas, desde a identificação dos candidatos e formação da hipótese da persona até a análise dos dados para encontrar tendências e padrões de comportamentos dos usuários entrevistados. São baseadas num modelo de aprendizado, onde o entrevistado é o mestre e o pesquisador o aprendiz – observar e fazer perguntas ao usuário. A parceria entre observado e observador é essencial. A entrevista e a observação devem ter o tom de uma exploração colaborativa, alternando entre a observação do trabalho e a discussão de sua estrutura e detalhes. A entrevista deve ser guiada por um roteiro a fim de direcioná-la sutilmente para capturar dados relevantes para os objetivos de pesquisa. Como também, é o momento para incentivar a narrativa

por parte do observado. Ou seja, incentivá-lo a contar histórias específicas sobre suas experiências com um produto ou serviço: como eles o usam, o que eles pensam disso, com quem mais eles interagem ao usá-lo, para onde vão e assim por diante. Histórias detalhadas desse tipo geralmente são as melhores maneiras de entender como os usuários se relacionam e interagem com os produtos. A interpretação por parte do pesquisador também é aspecto importante. O comportamento dos usuários, seu ambiente e o que eles dizem são pontos que devem ser observados e interpretados num conjunto. Os entrevistadores devem ter cuidado para evitar pressupostos baseados em suas próprias interpretações dos fatos sem verificar essas suposições com os usuários (COOPER; REIMANN; CRONIN, 2007).

Grupo focal também é uma técnica qualitativa citada pelos autores. Muito usada na área de *marketing* para investigar a intenção de compra ou não de um produto ou serviço, também é útil para medir reações iniciais à forma de um produto, sua aparência visual ou desenho industrial. Os grupos focais também podem reunir reações a um produto que os entrevistados estão usando há algum tempo. São formados por usuários representativos, geralmente escolhidos para corresponder a segmentos demográficos previamente identificados do mercado-alvo. São reunidos em uma sala onde é aplicado um conjunto estruturado de perguntas com opções de resposta pré-estabelecidas. Frequentemente, a reunião é gravada em mídia de áudio ou vídeo para referência posterior. É importante ressaltar que a atividade em grupos tende a conduzir ao consenso, a opinião da maioria se torna a do grupo. No entanto, um dos objetivos da pesquisa dos usuários com as personas é perceber e entender todos os diferentes padrões de comportamento que um produto deve atender. Grupos focais tendem a sufocar exatamente a diversidade de comportamento e opinião dos indivíduos.

Cooper, Reimann e Cronin (2007) não descartam a união dos métodos qualitativos e quantitativos. A pesquisa quantitativa responde perguntas sobre "quanto" ou "quantos". A qualitativa descobre sobre o que, como e por que em detalhes, é um reflexo das complexidades reais de situações humanas reais. Os comportamentos humanos são muito complexos e sujeitos a muitas variáveis, qualquer tentativa de reduzir o comportamento humano às estatísticas provavelmente negligenciará nuances importantes dos usuários. Desse modo, consideram que os dados, que podem apoiar e complementar a criação de personas, são levantados a partir de (em ordem aproximada de eficácia): entrevistas com usuários; informações sobre usuários fornecidas

por partes interessadas; dados de pesquisa de mercado, como grupos focais e pesquisas; modelos de segmentação de mercado e de revisões de literatura e estudos anteriores (em caso de).

As personas são derivadas de padrões observados durante entrevistas e observações de usuários reais ou em potencial (e às vezes clientes) de um produto. Criar personas críveis e úteis requer análise em detalhes e uma síntese criativa. Existem vários métodos eficazes para identificar padrões de comportamento em pesquisa e transformá-los em arquétipos de usuários úteis, a transparência e o rigor desse processo são aspectos fundamentais. Os principais passos são: identificar as variáveis comportamentais; mapear os sujeitos da entrevista a partir dessas variáveis; identificar padrões de comportamento significativos; sintetizar as características e os objetivos relevantes; verificar se há redundância e integridade; expandir a descrição dos atributos e comportamentos para então constituir os tipos de persona. Colagens combinadas com narrativas cuidadosamente escritas, são uma maneira eficaz de transmitir os aspectos emocionais e experienciais de uma persona. Veja o exemplo de apresentação na figura 11:

Figura 11: Formato de apresentação das personas



Fonte: COOPER; REIMANN; CRONIN, 2007, p. 103.

Aquino e Filgueiras (2008) compreendem o processo de construção de personas similar ao ciclo de vida apresentado por Pruitt e Adlin (2006). Nesse contexto, “as personas possuem um ciclo de vida que envolve passos desde a concepção, passando pela aplicação e utilização e chegando a fase de reutilização ou aposentadoria” (AQUINO; FILGUEIRAS, 2008, p. 3). Como

técnicas de coleta de dados citam questionários dirigidos e pesquisa etnográfica. Informações imaginadas são consideradas para o processo de construção de personas. Agregada às características comportamentais, demográficas e biográficas (idade, sexo, profissão, gênero, estado civil, nível de escolaridade, naturalidade, etnia, habilidades, etc.). Cada persona é a representação de um conjunto de usuários. O conjunto é formado pelos perfis dos sujeitos que se assemelham ou que se relacionam. As personas são comunicadas de modo individual, com imagem da persona representada por desenho e uma narrativa (breve história de vida, status socioeconômico atual, relação humano-computador, comportamento em geral, dificuldades, medos, etc.) (MARTINS; VANZ, 2021).

Miaskiewicz, Summer e Kozar (2008) afirmam que a pesquisa de usuários é comumente usada para a criação de personas; e que os métodos e técnicas aplicadas, geralmente, são qualitativas. Os autores apresentam a técnica chamada Análise Semântica Latente (em inglês *Latent Semantic Analysis – LSA*) para identificar e criar personas diretamente das transcrições das entrevistas realizadas com usuários, por meio de dados textuais. O objetivo da análise é encontrar características comuns compartilhadas por vários usuários – esses grupos de usuários semelhantes constituem o resultado personas. O processo é composto por 5 etapas. Inicia com entrevistas guiadas por roteiro, para que todos os pesquisados recebam o mesmo conjunto de questões. As respostas para cada pergunta devem ser organizadas em uma planilha, sendo que as perguntas são alocadas nas linhas e o entrevistado na coluna. Ao final, o documento permite a fácil análise e comparação das respostas. As semelhanças e o espaço semântico das respostas dos entrevistados para cada pergunta são identificados, resultando em matrizes de palavras e significados. Logo, os entrevistados similares são reunidos em grupos, o mais homogêneo possível para formar uma persona. A narrativa deve conter detalhes sobre familiares, amigos, posses, status socioeconômico, gênero e assim por diante. Como também descrever os objetivos, necessidades e frustrações da persona ou aspectos considerados relevantes para o produto que está sendo projetado (MARTINS; VANZ, 2021).

LeRouge *et al.* (2013) utilizou técnicas qualitativas na pesquisa-ação, entre elas observação direta, visitação de locais que os investigados frequentavam, revisão de documentos, grupos focais, entrevistas semiestruturadas e discussões informais com o grupo de investigados. O processo e a aplicação resultaram em personas que representavam os diferentes perfis de uma população envelhecida de pacientes (público-alvo da investigação). As personas resultantes

guiaram a melhoria dos planos de design, desenvolvimento e implementação de um aplicativo de telefone inteligente para ajudar a população diabética chinesa em envelhecimento crônico, a fim de proporcionar aos usuários a autogestão da doença (LEROUGE *et al.*, 2013).

Nielsen e Hansen (2014) defendem o estudo do usuário como metodologia para a construção das personas. O processo inicia com o desenvolvimento de hipóteses de personas, criadas a partir do levantamento inicial do máximo de conhecimento possível sobre os usuários. A partir da construção de cenários (histórias de como a persona utiliza o produto/serviço) é possível prever situações causadas e vivenciadas pelos mesmos, para que assim sejam realizados os ajustes necessários. Compreendem que a descrição das personas deve incluir características e informações físicas, sociais e psicológicas. A semelhança entre as características são usadas para criar grupos de pessoas, que formam a descrição completa de uma persona. A comunicação inclui nome e descrição, a qual costuma ter entre uma e duas páginas, com várias subseções que descrevem as características, vida, comportamentos, objetivos e preferências da pessoa. Muitas vezes uma foto acompanha a descrição (MARTINS; VANZ, 2021).

Marshall *et al.* (2015) utilizam o processo descrito por Pruitt e Adlin (2006) para construção das personas. O grupo focal é tido como a técnica para levantamento de dados com os usuários. A persona é apresentada com um breve relato com dados demográficos, metas e desafios em relação ao produto ou/e serviço. A comunicação traz imagens da persona, um retrato de pessoa real e uma frase de impacto que a define (MARTINS; VANZ, 2021).

Madsen *et al.* (2017) aplicaram a metodologia para identificar as necessidades e as motivações de docentes universitários para inovar processos educacionais. A técnica de coleta de dados foi uma entrevista semiestruturada, aplicada *on-line*. Os participantes foram indicados a partir da técnica bola-de-neve. Um primeiro e-mail convite foi enviado para uma lista-chave. Posteriormente, foram recebidas recomendações de novos indicados, dispostos a participar. Após concluir todas as entrevistas, o pesquisador tomou notas adicionais sobre objetivos, motivações e necessidades, escrevendo citações ilustrativas importantes. De acordo com os autores nomes e imagens auxiliam na comunicação de quem são as personas, como se fossem pessoas reais (MARTINS; VANZ, 2021).

Sturmer *et al.* (2019a) debatem, de modo exploratório, o conceito de personas junto às abordagens sobre públicos no contexto da comunicação organizacional. A discussão inicial, na perspectiva de RP, propõe fomentar o debate sobre como as personas podem contribuir para

encontrar "caminhos mais específicos para identificar as características dos públicos com os quais uma organização se relaciona, atrelando tais características à estratégia de comunicação" (Sturmer *et al.* 2019a, p. 25). Os autores apresentam a compreensão de personas em duas perspectivas. Os autores mencionam que na comunicação e no marketing a construção das chamadas *buyer personas* é baseada em pesquisa de mercado e em informações da base de clientes de uma organização. A coleta de dados pode ser realizada por meio de diferentes técnicas, como por exemplo, enquetes e entrevistas com os públicos (clientes reais, potenciais e de fora da base de dados da organização, que podem se tornar público-alvo); pesquisa de mercado; dados demográficos reais (gênero, idade, localização, função exercida, etc.). Aspectos da personalidade, preferências, interesses e dificuldades dos indivíduos devem ser considerados. Em relação ao número de personas os autores consideram:

Cooper (2004) conta que, a partir de entrevistas com representantes de possíveis grupos de usuários, são criados, normalmente, entre três e doze personas diferentes para cada projeto [...] Algumas personas são, inclusive, projetadas apenas para deixar claro que o sistema não está sendo desenvolvido para elas. (STURMER *et al.*, p. 34 2019a).

Sturmer *et al.* (2019b) afirmam que existem vários *templates* que guiam a construção das personas. Porém, eles convergem em seus propósitos, todos visam desenvolver um olhar focado nas necessidades dos clientes. Entre os modelos citados estão plataformas digitais que oferecem um jogo de perguntas-chave já elaboradas, como por exemplo, o “Fantástico Gerador de Personas”. Essa plataforma tem como objetivo a criação de personas para negócios e ao responder todas as questões pré-estabelecidas é gerado um arquivo com as informações inseridas pelo pesquisador. Sturmer *et al.* (2019a) apresentam um modelo para mapear as informações e construir as personas, formado por quatro seções. 1: Quem? Essa seção traz informações sobre o *background* da persona – sua experiência profissional, opções de carreira, núcleo familiar. Também abrange gênero, idade, renda, moradia e preferências (em relação ao objeto pesquisado). 2: O que? Busca responder às seguintes indagações. Quais os objetivos da persona e seus desafios? Como a organização, produto ou serviço poderá ajudá-la a vencê-los e superá-los? 3: Por quê? Essa seção identifica as frequentes certezas e as incertezas do indivíduo a respeito de suas dificuldades e desafio. 4: Como? Nessa última etapa são desenvolvidas mensagens-chave para as personas, descrevendo a solução que a organização oferta para esse consumidor e o argumento para as vendas.

Os autores consideram que a construção de uma estrutura de perguntas-chave facilita a coleta de dados junto aos consumidores reais ou em potencial. E que, “É possível inferir que a abordagem de personas pode ter potencial de adaptação sob o olhar das relações públicas, abrangendo os demais segmentos de públicos de uma organização e colocando sobre eles uma análise mais aprofundada e humana” (STURMER *et al.*, 2019a, p. 40).

Sturmer *et al.* (2019b) apresentam o conceito de Mapa da Persona Organizacional. A proposta une a construção de personas à perspectiva de públicos de interesses organizacionais de RP e é entendida como um guia para reconhecimento desses públicos. O Mapa é baseado em três etapas. A primeira é o Mapa do Cenário Organizacional (figura 12) que adota as classificações de Simões (1995; 2001) para os públicos, considerando-os como de *opinião*, de *comportamento*, de *decisão* e de *consulta* e se estrutura da seguinte forma.

Figura 12: Estrutura do Mapa do Cenário Organizacional



Fonte: STURMER *et al.*, p. 140, 2019b.

A segunda etapa contempla o Desenho da Persona (figura 13). O objetivo é construir os perfis dos públicos de uma organização numa perspectiva mais ampla, ou seja, além da persona do cliente consumidor/usuário. Entre esses públicos estão a comunidade, as entidades de classe, os parceiros, os colaboradores, os acionistas e diretores, fornecedores, imprensa, etc. Sua estrutura é constituída pelas seguintes perguntas-chave:

Figura 13: Estrutura do Desenho da Persona

<div style="text-align: center; background-color: #4a7ebb; color: white; padding: 10px; margin-bottom: 10px;">FOTO</div> <p>Nome: Faixa etária: Gênero: Onde mora: Onde trabalha: Cargo/Função: Faixa salarial: Educação:</p> <p>Característica da personalidade:</p>	<p>PÚBLICO Que público essa persona representa?</p>	<p>ROTINA O que faz no dia-a-dia?</p>
	<p>VALORES Quais são os meus valores? O que a persona mais preza?</p>	<p>INTERESSES Quais são os seus hobbies ou interesses?</p>
	<p>OBJETIVOS Qual o objetivo primário e secundário da persona.</p>	<p>CONTEÚDO Que tipo de conteúdo consome: Onde e em qual formato? Quais redes sociais acessa?</p>
	<p>PROBLEMA Quais são suas dores ou necessidades?</p>	<p>SOLUÇÃO Como a organização afeta a vida dessa persona, considerando suas dores/necessidades?</p>

Fonte: STURMER *et al.*, p. 142, 2019b.

A terceira etapa é o Mapa do Relacionamento, o qual propõe mapear o tipo de relação que a organização busca estabelecer com os públicos de interesse. O mapeamento é baseado na identificação das expectativas e dos objetivos da persona e da organização a respeito desse relacionamento. Como citam os autores essa fase:

[...] é o ponto principal de conexão entre o conceito de públicos de interesse organizacional e a abordagem de personas [...] retoma a abordagem de Simões (1995) a respeito de públicos de interesse, na qual o tipo de poder exercido pelo público sobre a organização é o critério principal para identificar e segmentar públicos. O tipo de poder é retomado para dar base ao mapeamento e definição das expectativas e objetivos do público sobre a organização, e vice-versa” (STURMER *et al.*, p. 144 – 146, 2019b).

O plano tático de comunicação é desenvolvido após o mapeamento das expectativas e objetivos dos relacionamentos. A figura 14 demonstra como é constituída a estrutura da etapa 3.

Figura 14: Estrutura do Mapa de Relacionamento

PÚBLICO Qual público é representado pela persona?	TIPO DE RELAÇÃO (Negócios, social, política, legal, parceria)	PODER Qual o pode exercido sobre a organização? <input type="checkbox"/> Decisão <input type="checkbox"/> Consulta <input type="checkbox"/> Comportamento <input type="checkbox"/> Opinião	EXPECTATIVA Persona Organização	OBJETIVO DA RELAÇÃO Persona Organização
PROBLEMA Qual o problema de comunicação a ser resolvido?	MENSAGEM-CHAVE Qual a mensagem que queremos comunicar para a persona?		LINGUAGEM Qual a linguagem/ vocabulário/tom de voz a ser utilizada com a persona?	
CONTEXTO Qual o contexto da persona para determinar a mensagem-chave, linguagem e canais de comunicação?			PONTOS DE CONTATO Quais os canais de comunicação para alcançar a persona?	
ACOMPANHAMENTO Como se dará o acompanhamento da relação organização-persona? Quais as rotinas que devem ser estabelecidas para manter o relacionamento?		FREQUÊNCIA Qual a frequência do contato que será estabelecido entre organização e persona?		

Fonte: STURMER *et al.*, p. 145, 2019b.

O modelo de Mapa da Persona Organizacional propõe desenvolver o processo comunicacional a partir das necessidades dos indivíduos que formam o público, para assim estabelecer o relacionamento entre as organizações e esses grupos. A referência não menciona modelos de apresentação das personas construídas. Apenas modelos dos quadros de coleta da informação.

Como podemos verificar os autores apresentados descrevem diversas metodologias e sugerem diversas variáveis para a construção das personas, uns de modo mais completo e detalhado do que outros. Em sua maioria mencionam que as técnicas qualitativas são essenciais na etapa de coleta de dados. Alguns outros consideram os dados quantitativos como agregadores de informação. Também se percebe que os autores concordam que as informações devem ser levantadas diretamente com os usuários reais ou em potencial, a fim de tornar as personas credíveis e analiticamente corretas.

A partir dessa síntese, observa-se que grande parte dos autores que publicam sobre a metodologia é de nacionalidade estrangeira e que ela é aplicada por pesquisadores de diversas áreas do conhecimento quando precisam investigar perfis. No Brasil, têm-se as publicações de Sturmer *et al.* (2019a) e (2019b) que se relacionam especificamente com a área de relações públicas e apontam algumas contribuições que a construção de personas pode oferecer à atividade de identificação e mapeamento de públicos, abrindo caminho para o debate.

Como mencionado na subseção 2.2, o grupo de usuários é um dos diversos públicos que compõem o ambiente relacional das iniciativas de acesso aberto aos dados de pesquisa. Além dele, têm-se outros grupos que podem ser formados por políticos, cidadãos, fornecedores de plataforma, instituições de ensino e pesquisa, editoras, agências de financiamento, etc.. Esses coletivos podem ser compostos por uma variedade de perfis e entendê-los de modo amplo e homogêneo pode tornar suas compreensões equivocadas e estereotipadas.

Compreende-se que a distorção da imagem de um público é capaz de afetar diretamente as tomadas de decisão de forma negativa, uma vez que pode não corresponder ao real contexto. Assim sendo, a metodologia de construção de personas ao centrar a construção dos arquétipos em dados coletados com indivíduos reais e sugerir uma série de modelos para a apresentação e compreensão desses, pode ampliar a discussão na área de RP sobre métodos e técnicas para a pesquisa de públicos.

À vista disso, o próximo texto apresenta os conceitos de públicos a partir de autores que estudam sobre o objeto e a prática. A sequência de referências visa oferecer mais subsídios para a reflexão sobre como a metodologia de construção de personas pode beneficiar a atividade de identificação e mapeamento de públicos em RP.

2.4 PÚBLICOS EM RELAÇÕES PÚBLICAS: CONCEITOS, CLASSIFICAÇÕES E TÉCNICAS DE PESQUISA

Primeiramente, explana-se que os conceitos, as classificações e as técnicas de pesquisa utilizadas para identificação e mapeamento de públicos nas bibliografias de relações públicas são variados. Este texto busca apresentar um pouco dessa diversidade trazendo visões de diferentes estudiosos, que não necessariamente são da área. A seleção considerou autores que estudam sobre públicos, escrevem sobre o conceito e a atividade de identificação, mapeamento e classificação

desses ao longo dos anos. Com isso, pretende-se demonstrar um pouco da abrangência desse objeto.

O texto começa com as referências internacionais e depois as nacionais em ordem cronológica, da mais antiga para a mais recente. Esta subseção forneceu subsídios para incitar a reflexão sobre como a metodologia de construção de personas pode contribuir no processo de reconhecimento do público investigado nessa tese. Essa perspectiva é motivada pelo segundo balizador dessa investigação e pela inferência de Strumer *et al.* (2019a), sobre o potencial de agregar valor à essa atividade de RP por meio da abordagem de personas. Portanto, o foco foi identificar quais as técnicas sugeridas por esses autores para a prática e em qual perspectiva a identificação e o mapeamento são aplicados – se na da organização ou na dos públicos. Ou seja, essa subseção não tem o objetivo de discutir se os conceitos e as classificações existentes são ou não adequadas.

Hirschman (1970), pesquisador alemão da área de economia explicou a eficiência das organizações pelas pressões do mercado a partir da capacidade das pessoas – compradores, usuários, empregados – de influenciar no ambiente dessas. O autor aponta três classificações para públicos: de *voz*, de *saída* e de *lealdade*. As ideias tiveram origem numa observação realizada por ele a respeito do transporte ferroviário na Nigéria. A partir de sua pesquisa que visava fatores econômicos, o autor identificou como os diferentes agentes influenciavam positivamente ou negativamente no sucesso empresarial. Os públicos de *voz* buscam ser ativos perante as práticas organizacionais, influenciando políticas e resultados. Os de *saída* caracterizam os que se ausentam do sistema relacional. Os de *lealdade* buscam contribuir para conquista dos objetivos da organização, sem interferir nas tomadas de decisão. Atores internos e externos da organização podem integrar esses três tipos de público. Por exemplo, clientes e/ou colaboradores podem ser considerados público de *voz* como também de *saída*. Apesar disso, o autor não aborda o conceito do objeto e nem a metodologia aplicada para essa identificação e classificação.

Gurnig e Hunt (1984), teóricos de RP norte-americanos, compreendem que público é um grupo de pessoas que enfrenta um problema semelhante, reconhece a existência desse problema e se organiza para fazer algo a respeito. O problema comum que cria e identifica o público geralmente é alguma consequência que a organização/iniciativa tem sobre esse ou vice-versa. Para os autores os públicos são sempre específicos, sempre têm algum assunto em comum, portanto, não podem ser gerais. À vista disso, os autores consideraram que os profissionais de

relações públicas devem ter uma teoria elaborada de públicos para a realização de pesquisas que visam identificá-los e mapeá-los.

Nessa perspectiva Gurnig e Hunt (1984) propuseram a Teoria Situacional dos Públicos. O modelo composto por três variáveis independentes e duas dependentes é o método mais complexo para a identificação e segmentação dos públicos (MÍGUEZ-GONZÁLEZ, 2006). As variáveis independentes são: reconhecimento de problemas que geram o público, reconhecimento de restrição (obstáculos) que vêm a afetar o comportamento dos indivíduos desse público e nível de envolvimento desses públicos com os assuntos da organização e suas iniciativas, podendo ser esse passivo ou ativo. As variáveis dependentes se referem à busca da informação pelos membros que possuem um comportamento de comunicação ativa, que pressupõe que o indivíduo, além de processar as informações que recebe, mobiliza-se para obter mais informações. Como também ao processamento da informação ou comportamento passivo de comunicação, que implica que o indivíduo se limita a processar a informação que lhe chega aleatoriamente sobre um assunto. Combinando valores altos ou baixos de cada uma dessas variáveis, obtêm-se os tipos de público (GRUNIG; HUNT, 1984; MÍGUEZ-GONZÁLEZ, 2006).

Para classificar os públicos, Grunig e Hunt (1984) afirmam que reconhecer os tipos de vínculos estabelecidos entre a organização e seus públicos é essencial. A partir disso, os autores apresentaram quatro classificações para públicos. *Ativo*: formado por indivíduos que são afetados pelo problema, o reconhecem e fazem algo a respeito. *Consciente*: que são afetados por um problema semelhante e reconhecem sua existência e causa. *Latente*: composto por pessoas que enfrentam um problema semelhante, mas não o detectam. E o *Não-público*: composto por indivíduos que não mantêm nenhum vínculo com a organização e não se encaixam em nenhuma das outras classificações. Os públicos *ativos* geralmente são altamente críticos, por isso é importante se comunicar com públicos *conscientes* antes que eles se tornem *ativos*. Quanto aos *não públicos* e aos públicos *latentes* (embora esses últimos possam evoluir à medida que se familiarizam com o tema), raramente se interessarão por problemas que não os afetam e se comunicarão sobre eles. Portanto, não é necessário investir muitos recursos na tentativa de se comunicar com eles (GRUNIG; HUNT, 1984).

Observa-se que, apesar da Teoria Situacional dos Públicos apresentar um modelo para identificação e mapeamento do objeto com variáveis a serem observadas, não indica de modo efetivo técnicas para coleta, análise e apresentação dos dados. Junto disso, embora tenha marcado

o curso da pesquisa de relações públicas ao longo dos anos, não existe consenso entre os estudiosos da área sobre ela. Alguns a consideram um modelo ultrapassado, outros acreditam que a pesquisa existente é extensa o suficiente (MÍGUEZ-GONZÁLEZ, 2006).

Os pesquisadores de gestão e negócios norte-americanos Friedman e Miles (2006) trazem para a discussão o conceito da palavra – *stakeholders* – e uma série de modelos normativos para as teorias e as ferramentas de gestão dos públicos considerados vitais para a sobrevivência e o sucesso de uma organização ou iniciativa. Os autores listam 75 noções para o termo em ordem cronológica de 1963 até 2003, todas são centradas na organização. Na tradução da palavra para o português, *stakeholder* significa parte interessada. A definição mais antiga do termo entende que *stakeholders* são grupos sem cujo apoio a organização deixaria de existir, creditada a um memorando do *Stanford Research Institute* publicado em 1963. De acordo com o documento os interesses dos *stakeholders* deveriam ser compreendidos pelos executivos, a fim ajustar os objetivos organizacionais a esses. O conceito clássico da palavra é “qualquer grupo ou indivíduo que pode afetar ou é afetado pela realização dos objetivos da organização” (FREEMAN 1984, p. 46). O autor considera que esses grupos de indivíduos que compõem o ambiente relacional das organizações podem ser vistos como oportunidades ou ameaças de acordo com o contexto. Compreende-se que a concepção de Freeman (1984) vai ao encontro do que conceitua Simões (2001) sobre o objeto – público. Desse modo, essa pesquisa entende que independente da classificação do público, todos podem ser considerados partes interessadas, conforme a ocasião e a estratégia da iniciativa. Ampliando a concepção de Cooper, Reimann e Cronin (2007) para o termo.

A classificação de públicos foi registrada já em 1930 pela *General Electric Company* (GEC) que identificou quatro grupos principais que afetavam suas atividades e objetivos – acionistas, funcionários, clientes e público em geral. Em 1940 a *Johnson & Johnson* reconheceu quatro “partes interessadas estritamente empresariais” – clientes, funcionários, gerentes e acionistas. O reconhecimento de terceirizados como público crescia e o número de grupos reconhecidos aumentava no âmbito empresarial. Friedman e Miles (2006) mencionam que os públicos mais comuns a serem considerados pelas organizações são: acionistas, clientes, fornecedores, distribuidores, funcionários, comunidades locais, sindicatos ou associações comerciais, ativistas, concorrentes, governo(s), reguladores e outros formuladores de políticas, financiadores, mídia, público em geral, parceiros, gerações futuras e passadas.

Os autores apontam que muitas empresas usam pesquisas para avaliar as necessidades, principalmente de funcionários e consumidores. As pesquisas são consideradas democráticas, pois solicitam opiniões que de outra forma não seriam ouvidas. Para Friedman e Miles (2006) as pesquisas com o público representam níveis mais sérios de engajamento, pois a organização solicita ativamente o *feedback* desses. Os estudiosos consideram que quando os públicos são informados e participam do processo de tomada de decisão, é mais provável que concordem com as ações, levando a um maior grau de confiança. Ações estratégicas colaborativas beneficiam o alcance de objetivos mutuamente benéficos, uma vez que cada uma das partes traz uma habilidade e recurso particular que juntas favorecem o todo.

Entre as técnicas citadas para as pesquisas estão: grupos focais, entrevistas e painéis consultivos. Entrevistas podem fornecer informações detalhadas sobre as percepções do indivíduo e também permitem uma comunicação bidirecional o que pode reduzir/evitar mal-entendidos. Grupos focais e painéis consultivos podem ser altamente eficazes para consultas reais. Porém, Friedman e Miles (2006) observam que a representação é algo a ser considerado nessas ações. O grupo deve ser devidamente representado e seus participantes precisam estar aptos a falar sobre o contexto.

Em 1989, no Brasil, o pesquisador de RP Andrade conceituou público como um grupo de indivíduos que se reúne espontaneamente. Esses grupos possuem abundância de informação, oportunidade de discussão e buscam encontrar uma ação em comum para expressar a opinião coletiva, baseada nas múltiplas atitudes e entendimentos de cada integrante. Andrade (1989) menciona sobre a importância de se conhecer o público e suas características. No entanto, não aborda técnicas de pesquisa para isso. Classificou então os públicos em *interno*, *externo* e *misto*. Por muito tempo, as definições foram apoiadas no critério geográfico/espacial que o objeto público tomava diante das organizações, apresentando a divisão simplificada nos três tipos citados. Considerada limitada, essa concepção incitou a revisão dos aportes teóricos por parte de alguns autores brasileiros da área, ampliando ainda mais o rol de visões. Era necessário compreender outros fatores, os quais poderiam ou não interferir na relação entre as organizações e seus públicos. Desde então essa relação é percebida a partir dos critérios de poder e de influência em um sistema social (SIMÕES, 1995; STEFFEN, 2008; FRANÇA, 2012; MARTINS, 2019; MARTINS; ROCKEMBACH, 2020).

Simões (1995), pesquisador brasileiro da área de RP, apresentou quatro classificações de públicos: de *decisão*, de *comportamento*, de *opinião* e de *consulta*. Essa concepção é baseada no conceito de Lucien Matrat, considerado o “pai” das relações públicas na França e autor do Código de Atenas, aprovado em 1965, pela Associação Internacional de Relações Públicas (IPRA) (FRANÇA, 2012, p. 28). A primeira classificação identifica os públicos que detêm poder de determinar algo sobre a existência da organização e suas iniciativas, juridicamente e legalmente, autorizando suas atividades. A segunda, os que executam as atividades da organização, cumprindo com suas finalidades. E a terceira espelha os que têm poder de influência sobre as ações da organização, mas não necessariamente, fazem parte de sua estrutura interna. Por último, estão os públicos de *consulta* que interferem de algum modo na tomada de decisão, pois suas opiniões são consideradas importantes no processo.

Em 2001, Simões reafirma as classificações trazidas em 1995 e conceitua públicos como grupos de indivíduos ou organizações que são influenciados ou influenciam as ações das organizações no alcance de seus objetivos. A relação entre organização e os públicos ocorre dentro de um sistema social sustentado por um processo de comunicação que possui a informação como matéria-prima e é baseado em sete dimensões: cultural; econômica; política; ideológica; histórica; jurídica e filosófica. Esse processo se configura em uma linha tênue de cooperação e conflito entre as partes. Ele ocorre de modo não-linear e pode apresentar as seguintes características: satisfação; insatisfação; boato; coligações; pressão; conflito; crise; arbitragem e convulsão social. As referências de Simões (1995 e 2001) também não abordam a proposta metodológica para a realização da identificação, do mapeamento e da classificação dessas definições.

Para Fortes (2003) – pesquisador de RP – o público pode ser qualquer grupo que possui interesse, impacto real ou potencial sobre as ações da organização. O autor cita uma lista de grupos que podem ser entendidos como públicos conforme níveis de proximidade espacial e de relacionamento. Entre esses estão: acionistas, administração superior, governo, empregados, entidades assistenciais, etc. E adota a classificação de *interno*, *externo* e *misto*. Diferente de Andrade (1989), Fortes (2003) leva em consideração o tipo de envolvimento do relacionamento a ser estabelecido, o grau de dependência e a capacidade de causar impactos favoráveis ou desfavoráveis no destino da organização para a classificação. Ou seja, não se trata de uma categorização unicamente geográfica, a partir da distribuição física dos grupos mais ou menos

próximos da organização. Fortes (2003) aborda o termo *stakeholder*, entendendo que esse se refere a indivíduos e grupos de interesse que exercem pressão sobre as ações estratégicas das organizações e podem ser afetados de diferentes maneiras pelas decisões. O autor apresenta uma série de tipos de pesquisa que podem ser realizadas para traçar o perfil dos públicos, apreciar seu comportamento, conhecer suas aspirações, atitudes e opiniões. Entre as que podem ser realizadas com diferentes públicos estão: pesquisa etnográfica, participante e pesquisa-ação, pesquisa de atitude, grupos focais, pesquisa bibliográfica, enquetes e conversas informais (FORTES, 2003).

Steffen (2008), pesquisadora brasileira da área, faz um estudo dos conceitos sobre públicos a partir da análise de conteúdo de 27 obras publicadas no Brasil entre 1960 e 1997. A autora apresenta dois enfoques para o conceito, o de *comunicação* e o de *poder*. No primeiro, o relacionamento entre organização e públicos é conduzido pelo discurso e como ele acontece, entendendo a atividade de RP como técnica de comunicação. Esse enfoque tem como propriedades essenciais mais representativas que: o público seja atingido por uma linguagem específica, relacione-se com a organização por meio da comunicação e esteja dentro da esfera de uma instituição. O segundo foco está vinculado à tomada de decisão. Essa visão aponta que a rede de públicos pode reagir sobre qualquer ação realizada pela organização – a favor ou contra ela. Como também cooperar com os interesses legítimos da mesma, ser influenciado por um programa de RP, ter expectativas em relação à organização e relacionar-se em situações de conflito. Nesse enfoque a atividade é compreendida como função gerenciadora das relações de poder. No entanto, a autora não trata sobre o processo metodológico dessas referências. Ou seja, foca exclusivamente no aporte teórico conceitual das definições.

Soares e Lomando (2010), estudiosas de RP, trazem o conceito de Simões (2001) para públicos – “conjunto de pessoas ou organizações cujos interesses são afetados e afetam as ações da organização na consecução de sua missão” (SIMÕES, 2001, p. 54). As autoras compreendem que para classificar os públicos, tanto os gerais quanto os estratégicos, é necessário analisar e refletir sobre dois aspectos principais: 1) a política de relacionamentos adotada pelas organizações e 2) o mapeamento amplo por meio de critérios de classificação. Para isso, Soares e Lomando (2010) apresentam algumas técnicas para a atividade. São elas: pesquisa documental em relatórios, arquivos, filmes e apresentações; entrevistas quantitativas que podem ser aplicadas via telefone, e-mail ou pessoalmente; grupos de discussão ou grupos focais como opções para coleta de dados qualitativos, como também as pesquisas de opinião e de clima para aprofundar o

conhecimento sobre os públicos. As autoras descrevem brevemente algumas das técnicas, não aprofundando a proposta metodológica no que tange à aplicação.

Numa perspectiva empresarial, o pesquisador brasileiro de RP França (2012) define públicos como grupos organizados de diferentes setores (públicos, sociais, econômicos) que podem autorizar a constituição ou oferecer o suporte necessário para o desenvolvimento dos negócios das organizações. Ao compreender que a essência de RP é ter a capacidade de se relacionar bem com os diversos públicos – visando vantagens mercadológicas e institucionais – o autor considera fundamental o mapeamento detalhado das relações estabelecidas entre organização – públicos e de suas significações no contexto operacional da organização. O estudioso salienta que “Todo trabalho de classificação exige reflexão e metodologia no procedimento a ser adotado” (FRANÇA, 2012, p. 95).

Ao defender a necessidade da pesquisa sobre o conceito do objeto, França (2012) apresenta a metodologia aplicada por ele para identificar os públicos com os quais as organizações – nacionais e multinacionais, de grande porte na região sudeste do Brasil – se relacionam e como se dá o relacionamento entre esses agentes. O autor menciona que as técnicas qualitativas são as mais seguras quando se têm esses objetivos. Para tanto, considerou entrevistas com executivos da alta administração das organizações. O instrumento de coleta de dados utilizado foi um roteiro formado por 25 perguntas. França (2012) observou elementos fundamentais, sendo eles: identificação dos tipos de públicos e os objetivos da relação, grau de importância, duração, modo, prioridade e expectativas dos relacionamentos. As variáveis trabalhadas pelo autor foram: relação, relacionamento, inter-relacionamento, interdependência, participação e motivação.

Após a coleta de todos os dados é possível construir tabelas indicativas de relacionamentos conforme os objetivos da pesquisa. O autor cita quatro modelos.

Modelo 1: tipo de público, tipo de relacionamento, objetivos da empresa, classificação, resultados esperados (expectativas da empresa), expectativas do público.

Modelo 2: tipo de público, tipo de relacionamento, objetivos da empresa, classificação, resultados esperados pela empresa.

Modelo 3: tipo de público, tipo de relacionamento, objetivos da empresa, classificação, natureza temporal da relação, resultados esperados, expectativas do público.

Modelo 4: Relação dos públicos com a organização. Aplicam-se os mesmos parâmetros, porém, em ordem inversa: tipo de público (empregados), tipo de relacionamento com a empresa, objetivos dos empregados em relação à empresa, classificação; expectativas dos empregados em relação à empresa (FRANÇA, 2012, p. 96).

Para França (2012), o conhecimento dos públicos ocorre logicamente a partir de três critérios de relacionamento. 1) Grau de dependência jurídica ou situacional da organização perante seus públicos, a fim de viabilizar sua constituição, existência e permanência no ambiente, mantendo sua produtividade. 2) Grau de participação dos públicos nos negócios da organização, na defesa de seus interesses e na sua promoção institucional. 3) Grau de interferência que determinado público pode exercer sobre a organização e suas atividades. A partir da conceituação lógica é possível formar redes de relacionamento específicas divididas em três categorias, conforme se detalha a seguir.

Na categoria 1 estão os públicos *essenciais*, ligados juridicamente à organização, ou não. Ela depende deles para sua existência e atuação. Segmentam-se em *constitutivos*, que fornecem recursos básicos para a organização e a empreendem executando as atividades-fim; e em *não-constitutivos*, que viabilizam a existência da organização, mantendo a produtividade e colaborando para as atividades-fim. Na categoria 2, estão os públicos *não-essenciais*: redes de interesse específico que possuem maior ou menor grau de participação na organização e prestam apenas serviços externos. A categoria 3 abrange as *redes de interferência*: concorrentes que atuam com os mesmos objetivos e características da organização; redes de comunicação de massa (impressa, eletrônica e digital) que podem interferir na opinião pública e redes de ativistas e ideológicas, estruturadas nacional e internacionalmente, atingem a opinião pública em defesa de diferentes causas. França (2012) aponta que a classificação de públicos pode variar conforme cada tipo de organização e sua área de atuação.

Para Lima (2015) o conceito de públicos está no centro das reflexões das relações públicas. Ao reconhecer outros modos organizativos, como organizações do terceiro setor, o autor compreende o conceito de públicos nas noções de sociabilidade, acontecimento e mobilização social, pensando-o como forma de ação coletiva. Numa visão dialética abrange a dimensão mobilizadora e sua constituição como ator social, com potência de ação que possibilita a permanente construção da realidade. Para tanto, o público necessita da construção de uma identidade coletiva, que permita coerência e unidade (LIMA, 2015).

Para compreender o caráter de acontecimento dos públicos, Lima (2015) sugere dois caminhos. Primeiramente, a partir da atividade de mapeamento de públicos, fundamental na prática de RP. O segundo, como forma de ação coletiva, a partir das discussões sobre os movimentos sociais ou os públicos políticos. Porém, não aprofunda o debate sobre métodos e

técnicas para o processo. De todo modo, o estudioso menciona que ao categorizar os públicos, a atividade de relações públicas os reconhece a partir do ponto de vista da organização.

Considerando os processos de comunicação pública, Henriques (2017) se refere a públicos "como formas abstratas e dinâmicas de experiência e de sociabilidades que se formam em função da problematização de acontecimentos e ações que afetam os sujeitos" (HENRIQUE, 2017, p. 57). O autor compreende públicos como entes relacionais e elementos centrais do processo, apresentando três dimensões distintas e interligadas para sua constituição, apresentadas no próximo parágrafo.

A cognitiva possibilita que público se perceba como tal, num processo de autoformação e autoafirmação que abarca a compreensão de problemas numa dimensão coletiva e pública. Vários problemas desafiam os públicos nessa dimensão, entre eles: a falta de informação prestada aos públicos, de disponibilidade de dados e de transparência; a distribuição desigual dos saberes e a superabundância da informação na atualidade que pode diluir excessivamente os conhecimentos sobre uma questão como também desviar o foco sobre aspectos centrais relevantes. A dimensão performática se refere às diferentes possibilidades de expressão do público – desde os modos mais difusos na conversação cotidiana até as ações coletivas mais organizadas – como por exemplo, protestos, manifestações, passeatas, etc. A dimensão organizativa trata das diferentes formas que os públicos tendem a assumir. Um público toma forma, antes de tudo, pela imaginação (de si mesmo, dos outros sobre ele e dele sobre os outros). Nessa dimensão está o dilema da institucionalização dos públicos em estruturas estáveis, o que tende a engessar a sua movimentação. Porém, a publicação de Henriques (2017) não visou o debate das práticas metodológicas para identificação e mapeamento dos públicos.

A partir do exposto, verifica-se que os conceitos para o objeto público consideram diferentes contextos e que em sua maioria estão centrados na organização. Ao observar as técnicas mencionadas no decorrer do texto, percebe-se que alguns autores contribuem de modo prático para uma metodologia de pesquisa de públicos. França (2012) contempla desde a coleta de dados até a apresentação dos resultados. Porém, não descreve com exatidão sobre as técnicas utilizadas, focando sua discussão nos critérios de relacionamento e suas variáveis. Apesar de o autor considerar a perspectiva dos públicos ao citar o Modelo 4 para construção de tabelas indicativas de relacionamento, ele foca sua coleta de dados na alta administração das organizações observadas.

Fortes (2003), Friedman e Miles (2006) e Soares e Lomando (2010) colaboram sugerindo tipos de pesquisa, técnicas qualitativas e quantitativas para o processo de identificação e mapeamento de públicos em RP. Através das referências, percebe-se que as técnicas qualitativas são as mais mencionadas, incluindo os roteiros como instrumento de pesquisa e as entrevistas como técnica de coleta de informações. No entanto, entende-se que as ideias dos autores ainda não convergem entre si e são centradas apenas na organização. As publicações deixam espaço para debate e reflexão mais aprofundados sobre como é possível desenvolver uma metodologia para a atividade de identificação e mapeamento de públicos de modo a considerar efetivamente os representantes dos coletivos. Além disso, o contexto dessa tese que considera um sistema - conjunto de organizações que possuem o objetivo em comum de promover iniciativas em prol do acesso aberto aos dados de pesquisa – aponta a amplitude de contextos que a atividade pode ser desenvolvida.

Como visto na seção 2.3.1, a metodologia de construção de personas é aplicada em diferentes campos de estudo, como por exemplo, computação, negócios, tecnologia da informação e de modo incipiente nas relações públicas. As propostas demonstram a diversidade de caminhos que podem ser seguidos. Do mesmo modo que os autores de RP, os de personas também priorizam as técnicas qualitativas para levantamento de dados. Porém, descrevem com mais exatidão as etapas do processo. Isso beneficia a aplicação prática da metodologia possibilitando mais confiabilidade aos dados. Observa-se que, quaisquer que sejam os esforços, todos estão centrados no público.

Desse modo, é possível considerar que as propostas metodológicas de construção de personas podem beneficiar e renovar as práticas de identificação e mapeamento de públicos nesse sentido. A sugestão de Sturmer *et al.* (2019b) demonstra essa possibilidade ao aproximar os assuntos “construção de personas” e “públicos organizacionais”. Os autores descrevem os tipos de dados a serem coletados e as variáveis a serem observadas para a construção da persona organizacional. O processo sugerido amplia a visão sobre os públicos e suas diversas características já que busca contemplar o contexto dos sujeitos, incluindo experiência profissional, desafios e dificuldades, entre outros aspectos. Outros quesitos considerados importantes da metodologia de construção de personas, pois se entende que atuam como agentes de aperfeiçoamento na atividade de RP são: 1) O conceito de arquétipo de Cooper, Reimann e Cronin (2007) – personas são compostas de dados observados de indivíduos reais observados em

contextos semelhantes na vida real; 2) O esqueleto de Pruitt e Adlin (2006) que concede uma imagem prévia do público, pois é desenhado a partir de dados factuais preliminares obtidos na fase de planejamento familiar. O esboço permite identificar aspectos conhecidos sobre determinado público e apontar os ainda desconhecidos ou que precisam de aprofundamento, servindo como guia para o conhecimento do público; 3) O modelo de apresentação desses arquétipos, formado por narrativas compostas por imagem e texto, que intencionam facilitar a comunicação e o entendimento desses perfis por parte dos interessados. Nesse contexto, essa investigação propõe essa mudança de perspectiva para a atividade de RP, agora centrada no público – que nesse caso será o usuário. Ele é um dos diversos públicos que podem ser identificados no contexto da ciência aberta – subcapítulo 2.1.

O debate e a aplicação incipiente das personas na área de RP permitem novos olhares e práticas visando aperfeiçoamento do processo. “É complexa a relação das organizações com seus públicos. Para compreendê-la, é necessária uma pesquisa que possa analisar as diferentes formas dessa interação e que forneça dados suficientes para o conhecimento mais profundo dos diferentes públicos” (FRANÇA, 2012, p. 48).

Na próxima seção são demonstrados dados que buscam fornecer a perspectiva do perfil dos usuários brasileiros de serviços de acesso aberto aos dados de pesquisa, a partir de algumas referências nacionais. Esses dados foram fundamentais para essa investigação, uma vez que serviram como guias para a coleta e a análise dos dados.

2.5 PERFIL DOS USUÁRIOS DE SERVIÇOS DE ACESSO ABERTO AOS DADOS DE PESQUISA NO QUE TANGE A ECOLOGIA: A PERSPECTIVA TEÓRICA BRASILEIRA

Esta seção traz o conjunto preliminar de informações sobre usuários de dados de pesquisa ecológicos abertos no Brasil, tendo como bases principais as respostas do questionário aplicado pelo GT – RDP Brasil em 2018 e as publicações de Anjos (2016) e Vanz *et al.* (2018). O levantamento realizado pelo GT teve como objetivo mapear os potenciais usuários nacionais de serviços de acesso aberto aos dados de pesquisa. O questionário aplicado *on-line* foi composto por 27 questões objetivas e se dividiu da seguinte forma: oito questões demográficas, 13 questões relativas às práticas dos pesquisadores sobre dados de pesquisa e seis que buscaram identificar as percepções em relação ao compartilhamento, uso e reuso de dados. O instrumento foi enviado para um corpus formado por pesquisadores de todo o Brasil, “contatados por meio de correio

eletrônico de instituições de pesquisa, coordenadores dos programas de pós-graduação, líderes de grupo de pesquisa cadastrados no Diretório do CNPq” (VANZ *et al.*, 2018, p. 3).

A partir dos 4.703 respondentes, 105 se identificaram como pertencentes à área de Ciências Biológicas – ecologia e é nesse corpus que as observações são embasadas. Algumas perguntas do questionário permitiam respostas de múltipla escolha, por esse motivo alguns resultados superaram o número de 105 registros. Maior parte, 68 pessoas se identificaram como sendo do gênero masculino e 37 do feminino. Em relação ao nível de escolaridade, 46 eram doutores, 35 pós-doutores, 17 mestres e sete graduados. Esses indivíduos eram pertencentes às mais diversas instituições, em sua maioria (85) do setor acadêmico público. Boa parte (64 pessoas) se identificou como professor pesquisador, 24 como aluno de graduação e 17 como pesquisadores. Dos 105 investigados, 32 possuíam entre 61 e 70 anos, 26 entre 51 e 60 anos, outras 26 pessoas, de 41 a 50 anos, 13 entre 31 e 40 anos e apenas oito com mais de 71 anos. Nesse corpus, os dados de pesquisa são produzidos preponderantemente em grupos de até cinco pessoas – com 62 respostas, seguidos por 16 respondentes que produzem seus dados em grupos de seis até 12 pessoas, nove indivíduos informaram que a produção ocorre de forma individual e seis em grupos com mais de 12 pessoas.

Em 2018, 70 dos 105 compartilhavam seus dados de pesquisa. Sendo que 64 partilhavam apenas parte desses e somente seis, todos os dados de pesquisa. O compartilhamento dos dados acontecia por meio de repositórios caracterizados como temáticos (33 menções), de revista/editor (47), da própria instituição (26), como também nacionais (26). Os motivos apresentados por esses 70 indivíduos para o compartilhamento foram publicação em periódico ou em anais de evento (45 citações), por iniciativa própria (39), por exigência do principal financiador (21) e por solicitação da instituição em que atuam (oito). Observa-se que dos 105 respondentes, 23 alegaram não compartilhar nenhum de seus dados de pesquisa em 2018. Dentre as principais razões apresentadas por esses, para não compartilhar, estão o entendimento de que é preciso publicar os resultados antes de compartilhar os dados (16 menções), a falta de exigência de compartilhamento (sete), a falta de infraestrutura (seis) e a não padronização dos dados (seis). A maioria das 105 pessoas informou conhecer pouco sobre a gestão de dados de pesquisa (64 citações) e somente 16 afirmam conhecer bastante sobre o processo. Nesse grande grupo também é possível verificar que o acesso livre aos repositórios de pesquisa ocorre em certas ocasiões, ou seja, não é uma prática frequente entre investigadores da área. Muitos dos dados compartilhados

ainda precisam da permissão do pesquisador responsável para acesso (55 citações), somente 12 afirmam que outros podem acessar seus dados livremente. Das 105, 80 pessoas afirmaram que a instituição onde atuam nunca disponibilizou pessoal especializado para preparar os dados de pesquisa para serem compartilhados, 75 apontam que suas instituições nunca nem ofereceram orientações específicas para a ação e 61 indivíduos alegam que suas instituições nunca disponibilizaram um repositório para o compartilhamento. Quando questionado sobre o acesso aos dados de outros pesquisadores, 41 respondentes afirmaram utilizar algum repositório para a ação. Entre os indicados apareceram o *Ocean Biogeographic Information System (OBIS)*, a *Global Biodiversity Information Facility (GBIF)* e a Rede *speciesLink*. Sobre o uso efetivo de dados compartilhados por outros grupos em suas pesquisas, 51 mencionaram já ter utilizado e 33 nunca aproveitaram.

Retomando as observações sobre o grande grupo de 105 respondentes, 59 concordam em utilizar dados de outros pesquisadores se estiverem acessíveis e 25 concordam parcialmente. Esses indivíduos consideram ser apropriado criar novos conjuntos de dados a partir de dados compartilhados (58 citações), 61 estão dispostos a ajudar outros pesquisadores ao compartilhar seus dados e 53 confiam que o compartilhamento de dados pode ampliar o reconhecimento acadêmico, uma vez que divulgam a pesquisa e possibilitam o aumento das citações. Os investigados também se demonstram dispostos a dedicar parte de seu tempo e esforço para preparar os dados para serem compartilhados (51 registros). Porém, pode-se dizer que ainda há desconfiança no processo de compartilhamento de dados, 47 indivíduos concordam total ou parcialmente que suas ideias de pesquisa poderiam ser copiadas ao compartilhar. Os respondentes também afirmam esperar retornos de outros pesquisadores quando esses utilizarem os dados compartilhados. Dentre esses retornos estão: acordo mútuo de reciprocidade (48 citações), agradecimento (58) e citação nas publicações resultantes (63), ter acesso a uma cópia do artigo publicado (67) e a todos os demais produtos resultantes da pesquisa, incluindo material educacional, apresentações, etc. (54).

Em relação aos tipos de repositório que deveriam preservar os dados de pesquisa, na opinião dos 105 investigados da área da ecologia, 68 entendem que seus dados deveriam ser preservados em repositórios temáticos, 64 em repositórios de revistas, 63 em repositório da instituição e 44 em outro tipo de repositório de dados. Na opinião dos respondentes as agências

de financiamento e as instituições de filiação devem oferecer serviços de apoio à gestão dos dados de pesquisa.

A seguir, são concentradas as observações em relação aos dados de pesquisa produzidos pelo grupo de 105 indivíduos da ecologia que responderam a *survey* do GT – RDP Brasil (2018). Conforme os respondentes, os dados são em sua maioria observacionais (69 menções) e experimentais (53). Dados coletados de sensores e equipamentos também são utilizados pelos pesquisadores da área (37 registros), seguidos por entrevistas (17), questionários (17), dados documentais (15), dados automaticamente gerados por programa de computador (14) e dados de simulação (10 citações). Esses dados estão em diversos formatos, entre eles: planilhas (79), banco de dados (67), textos (57), imagens (45), animações, modelos ou simulações (15), slides (11). Dados cartográficos, mapas, *software*, algoritmos, áudios, vídeos e websites também aparecem, mas com menos de 10 registros cada um. O volume estimado dos dados de pesquisa produzidos por cada um desses pesquisadores, em um ano, varia entre 1 a 50 gigabytes (30 registros) e dados maiores que 1 gigabyte (24). Observa-se que, 21 indivíduos não souberam responder a questão, o que indica a necessidade de investigar sobre o assunto.

Também numa perspectiva brasileira, é apresentada a referência de Anjos (2016). A autora investigou quais as percepções e práticas adotadas pelos pesquisadores de 15 programas de pós-graduação em ecologia, em seis estados do nordeste brasileiro no que tange a gestão dos dados científicos. Conforme a publicação a área da ecologia está em constante desenvolvimento, necessitando do uso e da preservação de dados, já que esses são fontes primárias, podendo ser utilizados e reutilizados em pesquisas futuras, a fim de compreender e/ou certificar fenômenos. Como objetivos, a autora teve traçar o perfil dos pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro; identificar as ferramentas adotadas na realização do plano de gestão de dados científicos de pesquisa por esses pesquisadores; averiguar a natureza e origem dos dados científicos coletados; verificar se os pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro fazem uso de metadados; identificar as práticas de compartilhamento de dados empregadas e 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.

Por meio de um questionário semiestruturado, elaborado e executado pelo *software Lime Survey*, Anjos (2016) coletou informações de 58 investigados-alvos da pesquisa (de um total de

303 pessoas que receberam o instrumento via e-mail). Dezesesseis respondentes pertencem à faixa etária entre 35 e 54 anos, 25 pessoas à de 35 a 44 e 17 indivíduos à de 45 a 54 anos. Boa parte, 44 respondentes, afirmaram fazer a gestão dos dados. Os pesquisadores, que relataram não fazer a gestão dos dados científicos coletados nas suas pesquisas, apresentaram as seguintes justificativas: falta de conhecimento de ferramentas adequadas; não tiveram a ideia de compartilhar e não pesquisaram sobre como fazer ou não possuem informações sobre como proceder. Os investigados alegam que a natureza da análise de suas pesquisas são mistas, ou seja, utilizam-se de dados qualitativos e quantitativos. Em relação à origem dos dados de pesquisa, percebe-se que eles podem variar entre observacionais, computacionais e experimentais.

Sobre planos de gestão de dados: 25 pesquisadores afirmaram desenvolver um plano ao longo da condução de suas pesquisas e 19 negaram realizar algum processo para isso. Algumas das justificativas dos que negaram ter plano de gestão foram: deixo a cargo dos orientandos; desconhecimento; não existe um planejamento específico, a tabulação e o armazenamento dos dados são feitas em planilhas; falta de tempo; alega não conseguir aplicar as ferramentas necessárias, mas possui interesse; etc. Os pesquisadores que fazem plano de gestão apontam utilizar as seguintes ferramentas: Excel, Google Drive, Mendeley, ECOCEM, *Access*, banco de dados, *Open Office*, Trello, Capes, *Science Direct* e Bireme.

Dos 44 pesquisadores, somente 13 fazem o uso de metadados. Os padrões citados foram: data, localidade geográfica, comportamentos, dados climáticos, dados de outros estudos, unidades de medida, equipamento utilizado no registro e métodos utilizados. Três deles não souberam informar os padrões de metadados utilizados.

Quando questionado sobre compartilhamento de dados, os respondentes demonstram concordar com o compartilhamento e reuso das informações de suas e de outras pesquisas, mas consideram importante que os dados sejam citados quando reutilizados. As razões apresentadas para o não compartilhamento de dados coletados pelos pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro foram: a falta de infraestrutura técnica, a ausência de clareza dos padrões de metadados, a falta de tempo, a ausência de pessoal de apoio, ausência de clareza no que diz respeito aos deveres do usuário dos dados e a inexistência de incentivo para disponibilização dos dados.

Em relação a quando os dados poderiam ser disponibilizados para o público e qual seria o tempo necessário para isso, boa parte (37 investigados) concorda que esse tempo deve ser

definido pelo próprio pesquisador e que em média é necessário de dois a quatro anos para isso. Alguns informam que esse período pode variar dependendo do quão sensível são os dados. Anjos (2016, p. 47) observa que “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”.

De modo geral, a autora conclui que os pesquisadores dos programas de pós-graduação em ecologia do nordeste brasileiro possuem conhecimento acerca do uso, da gestão e do compartilhamento de dados. Para eles esse processo é um incentivo para o crescimento da Comunicação Científica (ANJOS, 2016).

A partir do exposto nesse subcapítulo, percebe-se a diversidade de aspectos que podem compor o perfil do grupo de usuários da ecologia e seus diferentes níveis de conhecimento, interesse e influência para adoção e difusão do acesso aberto aos dados de pesquisa no âmbito brasileiro. Os dados fornecem subsídios importantes que guiam a investigação dessa pesquisa. Porém, mesmo com essas informações existe espaço para investigar mais profundamente aspectos sobre motivações, estímulos, medos, dúvidas, características, informações demográficas e sociais desses usuários. Dados esses fundamentais para a construção das personas e a compreensão de seus contextos.

Nesse contexto, percebe-se inicialmente que o grupo de usuários pode ser um público de voz no processo de construção das iniciativas. Ou seja, parte ativa durante o desenvolvimento do projeto, influenciando principalmente na efetividade de resultados. Como também de *lealdade*, já que busca contribuir para conquista dos objetivos das iniciativas, sem interferir nas tomadas de decisão (HIRSCHMAN, 1970).

No próximo capítulo são descritos os procedimentos metodológicos dessa pesquisa.

3 PROCEDIMENTOS METODOLÓGICOS

Uma vez que não existe consenso sobre os estágios que compõem os processos de construção de personas e de identificação e mapeamento de públicos, essa tese uniu técnicas científicas para responder o problema e atingir os objetivos apresentados. A proposta foi baseada nas referências Grudin e Pruitt (2002); Fortes (2003); Pruitt e Adlin (2006); Friedman e Miles (2006); Cooper; Reimann e Cronin (2007); Soares e Lomando (2010), França (2012) e Sturmer *et al.* (2019a) e (2019b).

Assim sendo, a metodologia foi composta por uma série de procedimentos que buscou ampliar o conhecimento sobre o perfil do público usuário das Ciências Biológicas - ecologia de serviços de acesso aberto aos dados de pesquisa, permitindo a compreensão de múltiplos aspectos que caracterizam esse ator e seu contexto. Considerando as premissas da construção de personas e do estudo de usuário a investigação foi centrada nos indivíduos que formam esse coletivo e não nos que promovem as iniciativas (TARAPANOFF, 1995; HERNÁNDEZ SALAZAR, 1997). Ou seja, apesar de considerar as informações provenientes das partes promotoras no levantamento e na análise de dados, o foco principal foi descobrir sobre os perfis que compõem o grupo de usuários que acessam as plataformas, a fim de compartilhar e reusar dados de pesquisa. Para tanto, esse capítulo explana os processos de pesquisa bibliográfica, de reuso de dados do GT – RDP Brasil (2018), de análise de conteúdo e movimento de categorização.

3.1 PESQUISA BIBLIOGRÁFICA

Esta fase levantou o referencial para fundamentação teórica dos temas que guiam a pesquisa – dados abertos de pesquisa, dados ecológicos, estudo de usuários, públicos em relações públicas e personas. Considerada o primeiro passo da pesquisa científica, essa etapa permitiu o contato direto com o que foi escrito e publicado sobre as temáticas (LAKATOS; MARCONI, 1992).

O levantamento bibliográfico, com caráter exploratório-descritivo, foi dividido em dois vieses. Primeiramente, considerou para a busca as palavras-chave: ciência aberta, dados abertos ecológicos, acesso aberto aos dados de pesquisa ecológicos, usuários de dados de pesquisa, compartilhamento e reuso de dados ecológicos, dados abertos de pesquisa, *data sharing* e dados

abertos. Essa etapa incluiu a busca de artigos, teses, dissertações, dicionários e relatórios publicados entre 2002 e 2022, na Base de Dados em Ciência da Informação (Brapci), na *Library Information Science Abstracts (LISA)*, na *Web of Science (WoS)* e no Google Acadêmico. As referências permitiram mapear alguns dos grupos envolvidos no contexto da ciência aberta e priorizar o coletivo de usuários, entendendo-o como um público, conforme explicado nessa tese. Esse bloco originou os capítulos e subcapítulos que abordaram os temas: ciência aberta, dados de pesquisa, dados ecológicos de pesquisa, mecanismos de compartilhamento e perfil dos usuários de serviços de acesso aberto aos dados de pesquisa na ecologia, numa perspectiva teórica brasileira.

O segundo viés adotado focou na busca de artigos, teses, dissertações e relatórios publicados entre os anos de 2002 e 2020, no portal de periódicos e na base de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); no Google Acadêmico e no repositório digital Lume, da Universidade Federal do Rio Grande do Sul (UFRGS). As palavras-chave utilizadas abrangeram alguns dos temas norteadores dessa pesquisa, sendo elas: origem das personas, metodologia de personas, *personas methodology*, estudo de usuário, públicos em relações públicas, personas e relações públicas. Os resultados desse levantamento deram origem aos capítulos e subcapítulos sobre o usuário como público e persona, a metodologia de construção de personas e os conceitos, classificações e técnicas de pesquisa para o objeto público em RP.

As referências informativas remissivas, citadas nas publicações encontradas em ambos os vieses, também foram consultadas. As de interesse foram devidamente incluídas no referencial. O compilado de estudos permitiu fazer o levantamento inicial de informações sobre o contexto pesquisado para o desenvolvimento das personas (PRUITT; ADLIN, 2006). O que possibilitou a percepção inicial sobre os usuários real ou em potencial de serviços de acesso aberto aos dados de pesquisa, a fim de iniciar o processo de compreensão desses. O conhecimento gerado pela pesquisa bibliográfica possibilitou pensar sobre outras perspectivas, que poderiam complementar os dados até então obtidos sobre o coletivo da ecologia e suas práticas de compartilhamento e reuso de dados de pesquisa. Na sequência são apresentados quais aspectos foram considerados para a seleção dos entrevistados e como foi composto o processo de coleta de informações.

3.2 CORPUS ENTREVISTADO, TÉCNICA E INSTRUMENTO DE COLETA DE DADOS

Conforme visto nos subcapítulos 2.3.1 e 2.4 dessa tese, os autores de personas e públicos indicam uma série de caminhos que auxiliam na definição do corpus, da técnica e do instrumento de coleta a ser utilizado para pesquisa. À vista disso, o corpus selecionado (descrito a seguir) foi estabelecido por meio de amostragem intencional não probabilística. Os elementos foram escolhidos pelo critério da investigadora, levando em consideração o aporte teórico, a pergunta norteadora e os objetivos dessa investigação.

Nas amostras não-probabilísticas (intencionais), tal definição é feita a partir da experiência do pesquisador no campo de pesquisa, numa empiria pautada em raciocínios instruídos por conhecimentos teóricos da relação entre o objeto de estudo e o corpus a ser estudado (FONTANELLA *et al.*, 2011, p. 389).

Este momento tratou da concepção das personas e aspirou levantar dados para a construção dos arquétipos (PRUITT; ADLIN, 2006). Os participantes dessa etapa foram enquadrados em dois perfis: a) perfil usuário e b) perfil parte interessada. O primeiro, foi originado pelo reuso dos dados gerados pela *survey* do GT – RDP Brasil em 2018, disponíveis para acesso em <https://cedap.ufrgs.br/xmlui/handle/20.500.11959/161>. Salienta-se que a análise desses dados gerou parte da seção 2.5 da presente tese, além de prover os nomes dos entrevistados. A observação dos dados compartilhados permitiu mapear 105 respondentes que se identificaram como integrantes das Ciências Biológicas – ecologia e visou atender ao objetivo específico (b) dessa tese, mapear as principais características dos perfis que formam o público usuário das Ciências Biológicas - ecologia, identificados na *survey* do GT – RDP Brasil (2018), a fim de aprofundar o conhecimento sobre esses. O usuário foi um público entendido previamente como de *voz*, uma vez que demonstrou ser ativo nas práticas de compartilhamento e reuso, podendo influenciar políticas e resultados de iniciativas que visam o acesso aberto. E de *lealdade*, já que buscava contribuir para conquista dos objetivos da mesma, sem interferir nas tomadas de decisão (HIRSCHMAN, 1970). Dos 105 respondentes, foram selecionados os 40 indivíduos que responderam “Sim” na questão 3.5 “Você aceitaria participar de uma entrevista para explorar as questões abordadas neste questionário em maior profundidade?”. Logo, foram filtrados os pesquisadores que afirmaram compartilhar todos ou parte de seus dados de pesquisa na época – pergunta “2.7 Quanto dos seus dados de pesquisa você atualmente compartilha em acesso

aberto?”, com 38 menções. O passo seguinte identificou os que já utilizaram, de fato, dados compartilhados por outros grupos em suas pesquisas, questão 2.13 do questionário, com 26 citações. Desses foram selecionados os indivíduos que concordavam em compartilhar todos os seus dados de pesquisa em um repositório de acesso aberto sem restrições, questão 3.1. Esse grupo contou com quatro professores pesquisadores, três pesquisadores e um aluno de graduação. Das oito pessoas mapeadas, somente cinco aceitaram participar da entrevista. Alguns dos participantes sugeriram novos nomes e contatos de pessoas interessadas no movimento do acesso aberto aos dados de pesquisa. Os convites para participação foram enviados por e-mail, porém, não se obteve retorno positivo de algumas indicações.

O perfil parte interessada contemplou pessoas ligadas a instituições que promovem e/ou coordenam ações e iniciativas em prol do acesso aberto aos dados de pesquisa, ou possuem algum tipo de autoridade e/ou responsabilidade sobre as ofertas, independente da área do conhecimento – tendo como base o conceito de Cooper, Reimann e Cronin (2007). Esses indivíduos foram indicados pela rede de referências do GT – RDP Brasil e pelos usuários entrevistados, a fim de atender o objetivo específico (a) desse estudo. Portanto, esses indivíduos deveriam promover e/ou coordenar ações e iniciativas em defesa do movimento brasileiro de acesso aberto aos dados de pesquisa. Esse perfil foi previamente identificado como um público de *decisão*, de *comportamento* e de *consulta* perante as iniciativas, uma vez que possuem poder de determinar algo sobre a existência das mesmas; que executam ações, a fim de cumprir com suas finalidades e que interferem de algum modo na tomada de decisão, pois suas opiniões são consideradas importantes no processo (SIMÕES, 1995). Como também, um público de *voz*, já que buscam ser ativos diante das práticas diárias, influenciando políticas e resultados das iniciativas (HIRSCHMAN, 1970).

Ao total, foram enviados 16 convites: nove para usuário e sete para parte interessada. Obteve-se retorno positivo de sete indivíduos. O quadro 2 apresenta a síntese do corpus:

Quadro 2: Síntese do corpus investigado

Nº	Tipo de Perfil	Área de atuação	Data da Entrevista
1 ²⁵	Parte Interessada	TI	10 e 17/09/2021
2	Usuário	Ecologia	04/10/2021
3	Usuária	Ecologia	05/10/2021
4	Usuário	Ecologia	18/10/2021

²⁵ A entrevista de número 1 ocorreu em dois dias, devido a motivos particulares da pessoa entrevistada.

5	Usuário	Ecologia	04/11/2021
6	Parte Interessada	Ecologia	14/02/2022
7	Usuário	Ecologia	02/06/2022

Fonte: elaborado pela autora.

Os participantes foram identificados por número, a fim de preservar suas identidades e impossibilitar a associação, direta ou indireta, a um indivíduo. O corpus total foi considerado suficiente para análise, uma vez que os relatos apresentaram padrões similares e complementares durante a pré-análise, conforme será explicado na subseção 3.3.

A técnica empregada foi entrevista semiestruturada, a segunda mais utilizada para a coleta de dados em estudos de usuários após o questionário e; também indicada para a identificação e o mapeamento de públicos. O contato síncrono permitiu a interação entre entrevistado e entrevistadora, possibilitando captar reações e sentimentos em relação às pautas, dando um maior grau de confiabilidade aos dados coletados. Por ser uma técnica face a face foi possível esclarecer terminologias usadas pelos entrevistados e pedir detalhes das respostas fornecidas, quando foram detectados fatos interessantes (DA CUNHA, 1982). As entrevistas foram executadas isoladamente para estimular a franqueza nos discursos e garantir que as visões individuais não fossem perdidas (COOPER; REIMANN; CRONIN, 2007). Os encontros foram mediados pela plataforma *Microsoft Teams* e duraram em média uma hora. Os depoimentos foram gravados em áudio e vídeo, transcritos para editor de texto do Google Drive e analisados manualmente em planilha do Excel, conforme será explicado na subseção 3.3.

Para elaborar os instrumentos de coleta, o esqueleto do usuário foi constituído. Para tanto, os dados dos 105 indivíduos da ecologia mapeados na *survey* do GT – RDP Brasil (2018) foram filtrados e observados de modo mais minucioso. As respostas do questionário apontaram alguns indícios sobre aspectos demográficos, práticas dos pesquisadores sobre dados de pesquisa e suas percepções em relação ao compartilhamento e reuso de dados. Esses dados foram articulados com as referências nacionais de Barros (2005), Barros *et al.* (2005), Drucker (2011), Tabarelli *et al.* (2013), Luizão *et al.* (2013), Anjos (2016) e Vanz *et al.* (2018). A partir disso, foi possível identificar a Persona Ecologia (PE) como principal usuária de serviços de acesso aberto aos dados de pesquisa ecológicos e as possíveis carências de informação em relação a esse perfil e seu contexto. Observa-se que os dados da *survey* e do quadro teórico não apresentaram subsídios que indicavam outros perfis prévios para representar o grupo de usuários. O esqueleto apresentou a seguinte ótica, veja o quadro 3:

Quadro 3: Esqueleto Persona Ecologia

Perfil: características que formam o <i>background</i> do indivíduo.	Doutor, entre 40 e 50 anos, pertencente ao gênero masculino, professor pesquisador do setor acadêmico público.
Prática: da pesquisa, do compartilhamento e reuso de dados de pesquisa.	Seus dados de pesquisa são produzidos em grupos de até cinco pessoas. Ele produz uma variedade de dados, a depender do tipo de projeto que está desenvolvendo ou orientando. A natureza da análise de suas pesquisas é mista, ou seja, utiliza-se de dados qualitativos e quantitativos. Esses dados podem ser observacionais, computacionais, experimentais, coletados por sensores e equipamentos, registros documentais e de texto, entrevistas, imagens/fotografia, simulações, mapas, áudios, vídeos, planilhas, etc. Num período de um ano o volume estimado desses dados pode variar de 1 a 50 gigabytes. Ele faz a gestão dos dados de sua pesquisa, mas precisa conhecer mais sobre esse processo. Geralmente, desenvolve um plano de gestão de dados ao longo da condução de suas pesquisas, utilizando as seguintes ferramentas: Excel, Google Drive, Mendeley, ECOCEM, Access, banco de dados, <i>Open Office</i> , Trello, Capes, <i>Science Direct e Bireme</i> . Porém, em alguns casos, ou os planos ficam a cargo dos orientandos, ou não é elaborado um plano específico. Os dados acabam sendo armazenados em planilhas, por causa da falta de tempo. Ele é disposto a dedicar parte de seu tempo e esforço para preparar os dados para serem compartilhados. Ele compartilha parte dos dados produzidos em repositórios temáticos, de revistas, de instituições não-governamentais, entre outros, mas os interessados precisam de sua permissão para acessar. Os bancos de dados ecológicos crescem em volume, amplitude e complexidade. Geralmente são iniciativas promovidas a partir de redes colaborativas formadas por investigadores que depositam seus dados e reusam dados de outros pesquisadores. Ele entende que o compartilhamento dos dados, total ou parcial, depende de quão sensível são esses dados, que a decisão de quando compartilhar deve ser definido por ele próprio, mas que em média é necessário de dois a quatro anos para isso. É preciso publicar os resultados antes de compartilhar os dados. Ele não costuma utilizar metadados para identificar os dados compartilhados. Porém, considera alguns padrões essenciais ao compartilhar seus dados: data, localidade geográfica, comportamentos, dados climáticos, dados de outros estudos, unidades de medida, equipamento e métodos utilizados no registro. Ele utiliza os dados compartilhados por outros pesquisadores para criar novos conjuntos de dados. Para acessar esses dados, pode utilizar repositórios como <i>Ocean Biogeographic Information System (OBIS)</i> , a <i>Global Biodiversity Information Facility (GBIF)</i> , a Biblioteca Digital Georreferenciada do Programa Pesquisas Ecológicas de Longa Duração (BDig-PELD) e a Rede <i>speciesLink</i> .
Motivação/Expectativas ao compartilhar dados.	Ele é disposto a ajudar outros pesquisadores ao compartilhar seus dados e confia que o compartilhamento de dados pode ampliar o reconhecimento acadêmico, uma vez que divulgam a pesquisa e possibilitam o aumento das citações. Junto a isso é motivado pela publicação em periódico ou em anais de evento, desde que seja por iniciativa própria. Ele espera retorno de outros pesquisadores quando esses utilizarem os dados compartilhados. Como por exemplo: acordo mútuo de reciprocidade, acesso à cópia do artigo e a todos os demais produtos resultantes da pesquisa, incluindo material educacional, apresentações, etc.; agradecimento e citação nas publicações resultantes. Ele entende que seus dados deveriam ser preservados em repositórios temáticos, em repositórios de revistas ou em repositório da instituição onde atua. Ele acredita que as agências de financiamento e as instituições de atuação devem oferecer serviços de apoio à gestão dos dados de pesquisa.
Medo/insatisfações.	Falta de infraestrutura técnica; ausência de clareza dos padrões de metadados; falta de tempo; ausência de pessoal de apoio; falta de transparência no que diz respeito aos deveres do usuário dos dados; inexistência de incentivo para disponibilização dos dados, ausência de políticas eficazes para compartilhamento; financiamento limitado; métodos de coleta não padronizados; desatenção para documentar os dados nos mecanismos de compartilhamento; ausência de articulação entre as diversas redes que promovem iniciativas para facilitar abordagens analíticas e de visualização integradas. Ele ainda

	desconfia do processo de compartilhamento, pois entende que suas ideias de pesquisa poderiam ser copiadas ao compartilhar. A instituição onde ele atua nunca disponibilizou pessoal especializado para preparar os dados de pesquisa para serem compartilhados, nem ofereceram orientações específicas para a ação ou disponibilizaram um repositório para o compartilhamento.
--	--

Fonte: elaborado pela autora.

O esboço facilitou a organização dos roteiros, a fim de promover o aprofundamento em tópicos importantes para responder o problema e os objetivos dessa pesquisa. À vista disso, a coleta de dados foi guiada por perguntas-chave que abrangeram pautas complementares entre si, a fim de estimular os (as) entrevistados (as) a fornecerem subsídios padronizados para a análise de conteúdo. Os roteiros e Termos de Consentimento Livre e Esclarecido (TCLE) são encontrados nos Apêndices A, B, C e D desse estudo. A similaridade entre os instrumentos de coleta aplicados nas entrevistas com os usuários e as partes interessadas foi proposital, com a intenção de possibilitar a análise de dados de modo complementar e comparativo.

O levantamento buscou por dados qualitativos e quantitativos. Os primeiros buscaram informações sobre vocabulário e outros aspectos sociais do grupo entrevistado, fornecendo respostas úteis para questões importantes que levam a soluções superiores e mais assertivas (COOPER; REIMANN; CRONIN, 2007). Possibilitaram compreender sobre expectativas, objetivos, motivações, limitações e frustrações para uso de serviços de dados abertos de pesquisa. Inspirada em Sturmer *et al.* (2019b) indicadores como rotina do dia a dia, tipo de conteúdo de interesse, onde e em qual formato também foram investigados. Soares e Lomando (2010) apontam que os dados qualitativos são úteis para aprofundar o conhecimento sobre o público. Já os dados quantitativos se referem a informações demográficas básicas sobre o usuário. Incluem variáveis como idade; gênero; estado civil; renda; formação; etc. É toda característica concreta que ajuda a restringir o segmento de mercado no qual as pessoas do público estratégico se encaixam melhor. Os dados quantitativos facilitaram a identificação do grupo num ambiente heterogêneo.

Todas as informações coletadas nessa investigação são confidenciais, interessaram os dados coletivos e não aspectos particulares de cada entrevistado. A participação nessa pesquisa não trouxe complicações legais de nenhuma ordem e os procedimentos utilizados foram aprovados pelo Comitê de Ética, registro número CAAE 42051020.0.0000.534. Nenhum dos procedimentos adotados ofereceu riscos à dignidade dos participantes. O risco previsto ao realizar a pesquisa foi de o participante ser identificado mesmo com todos os cuidados de sigilo adotados.

Havia opção de não concordar em participar ou desistir em qualquer momento, sem nenhum prejuízo. A investigação não teve nenhum tipo de despesa para os entrevistados, bem como não ofereceu nenhum tipo de pagamento pela participação.

Observa-se que a técnica de grupo focal não foi classificada para esse estudo, pois um dos pré-requisitos para aplicação da mesma é que os entrevistados não se conheçam e não façam parte de um universo comum, o que não condiz com a realidade desse estudo (MOYA; RAIGADA, 1998; CANALES; PEINADO, 2007). Grupos de discussão e entrevistas em grupo foram descartados, a fim de preservar a opinião individual de cada entrevistado. As técnicas de observação e de etnografia são indicadas por alguns autores que abordam a metodologia de construção de personas, já que promovem a descrição mais completa e fiel das práticas sociais dos indivíduos e suas redes relacionais (coletividades) (MATTOS, 2011; PINHEIRO; SZANIECKI; MONAT, 2016; POLIVANOV, 2013). No entanto, não foram aplicadas por motivos financeiros e de traslados, uma vez que demandam observação direta sobre o contexto pesquisado. Situação essa, agravada pela pandemia do COVID entre os anos de 2020 e 2022. Uma alternativa para a observação seria realizá-la por intermédio de câmeras instaladas no ambiente do pesquisado. Porém, esse estudo considerou que a ação criaria a sensação de invasão do espaço do mesmo. Além disso, implicaria também em investimentos financeiros e tecnológicos.

3.3 ANÁLISE DE CONTEÚDO E MOVIMENTO DE CATEGORIZAÇÃO

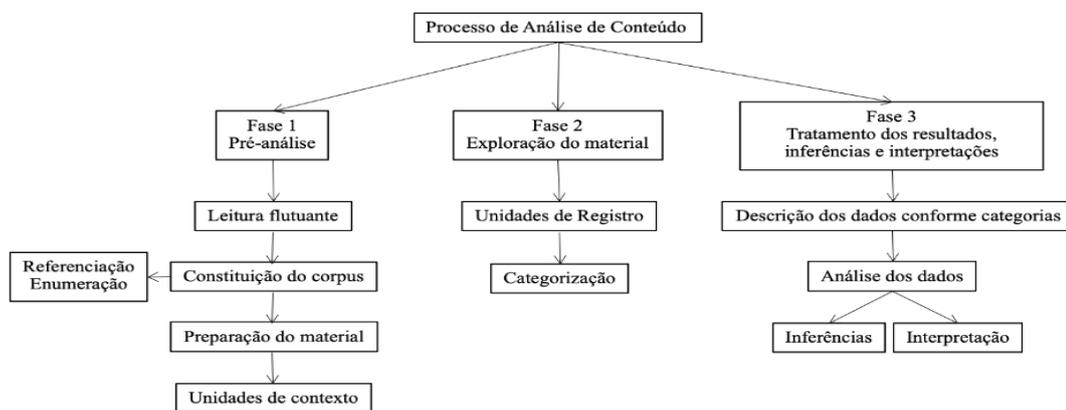
Nesta seção é explicado como aconteceu o processo de AC, relatando o movimento da categorização e a seleção dos dados considerados para a compreensão e construção da persona, tendo como principais referências Bardin (2016) e Rodrigues (2021). Observa-se que apesar das referências consideradas para compor a metodologia abordarem uma série de procedimentos para a coleta dos dados, não mencionam da mesma forma o processo de análise desses – Grudin e Pruitt (2002); Fortes (2003); Pruitt e Adlin (2006); Friedman e Miles (2006); Cooper; Reimann e Cronin (2007); Soares e Lomando (2010), França (2012) e Sturmer *et al.* (2019a) e (2019b). Nesse contexto, compreende-se que a análise é uma etapa da segunda fase do ciclo de vida das personas que trata da construção dos perfis (PRUITT, ADLIN, 2006). A opção pela AC foi

motivada pelos tipos de dados gerados por essa investigação e pelo entendimento de que esse tipo de análise pode beneficiar a concepção das personas, conforme explicado a seguir.

A AC é uma técnica de pesquisa que trabalha com a palavra, permitindo de forma prática e objetiva produzir inferências e interpretações da comunicação de um texto, por meio dos conteúdos manifestos e latentes nos discursos, reproduzíveis em determinado contexto social. Por procedimentos sistemáticos de descrição do conteúdo das mensagens visa obter indícios (quantitativos e qualitativos) sobre o tema e o objeto investigados. O texto é entendido como meio de expressão do sujeito, onde a pesquisadora busca categorizar as unidades de texto (palavras ou frases) que se repetem ou se completam, inferindo uma expressão que as representem (CAREGNATO; MUTTI, 2006).

A AC partiu dos textos gerados pelas entrevistas semiestruturadas e se dividiu em três fases: 1) pré-análise, 2) exploração do material e 3) tratamento dos resultados – inferência e interpretação (ver figura 15). Todo o movimento de categorização levou em consideração a pergunta norteadora, os objetivos, o quadro teórico e os instrumentos de coleta utilizados nessa investigação – regras da pertinência e da produtividade (BARDIN, 2016; RODRIGUES, 2021).

Figura 15: Resumo do processo de AC e suas fases



Fonte: elaborado pela autora a partir de Bardin (2016) e Rodrigues (2021).

A primeira fase contemplou as seguintes etapas. Primeiramente, a leitura flutuante sobre as transcrições das entrevistas que ocorreu concomitante a coleta e permitiu a constituição do corpus analisado de modo processual. No decurso, ocorreram as etapas de referenciação e enumeração, que identificaram os entrevistados e suas respostas mantendo-os anonimizados. A etapa de preparação do material contou com a organização dos documentos gerados na coleta em

seus diferentes formatos (áudio, vídeo e texto) em pastas salvas no Google Drive, a fim de permitir a consulta sempre que necessário durante o processo de análise. O movimento não-estruturado de ida e vinda nos materiais é considerado pertinente na AC, pois possibilita a observação da entrevistadora sobre os conteúdos produzidos, desde o início da coleta até a disposição final dos depoimentos.

Em sequência, uma planilha de Excel foi elaborada para cada entrevistado (a), elencando as perguntas-chave dos instrumentos de coleta e os respectivos depoimentos na íntegra. À vista disso, partiu-se para a leitura dos relatos estruturados e no decorrer foi-se sublinhando as unidades de contexto (UC). Essas são os índices, os recortes dos textos (palavra, frase ou sentença) que corresponderam à pergunta norteadora e aos objetivos dessa tese. Os discursos nem sempre seguiram a ordem exata das pautas dos roteiros, por vezes, os indivíduos misturavam os assuntos durante o relato.

Quadro 4: Exemplo de planilha utilizada na fase de pré-análise para a estruturação dos dados coletados nas entrevistas semiestruturadas

Entrevistado (a) N° 1		
Pergunta-chave	Depoimento na íntegra	Unidade de Contexto (UC)
Como você imagina esse usuário, quem é esse usuário.	A primeira imagem é de um <u>pesquisador</u> que tem <u>intenção de fazer uma pesquisa</u> e não... Já poderia <u>pular a parte de coleta de dados</u> , partir de um ponto anterior, que ele não precisasse afinal coletar <u>dados</u> , às vezes <u>observacionais</u> . <u>Dados que são únicos</u> , que ele não vai ter a oportunidade de coletar, então ele <u>vai utilizar dados de outros</u> . Mas, agora que tu perguntou dessa forma, também penso que as <u>agências de fomento</u> , que os <u>revisores dos periódicos</u> também são usuários se eles vão acessar esses dados, mas não penso que o objetivo deles é o reuso. Deles, entende? Então, não são os primeiros usuários que vem na minha cabeça. O primeiro que vem na minha cabeça é <u>o usuário que vai sentar com aqueles dados e daqueles dados partir para uma nova pesquisa</u> . Apesar de que, da forma de que tu falou, pensar que existem <u>outros papéis de usuário</u> , mas que não tem esse objetivo do reuso deles. Eles até <u>podem estar querendo garantir o reuso</u> de outros, mas eles não vão sentar e usar aqueles dados eles mesmos.	Pesquisador
		Intenção de fazer uma pesquisa
		O usuário que vai sentar com aqueles dados e daqueles dados partir para uma nova pesquisa
		Pular a parte de coleta de dados
		Dados observacionais
		Dados que são únicos
		Vai utilizar dados de outros
		Outros papéis de usuário
		Agências de fomento
		Revisores de periódicos
		Podem estar querendo garantir o reuso

Fonte: elaborado pela autora.

A pré-análise viabilizou reconhecer o conjunto de temas iniciais que guiaram as fases seguintes da AC. O tratamento manual dos dados facilitou a identificação dos conteúdos latentes e a extração de elementos que auxiliaram na compreensão desses indivíduos, permitindo as intuições iniciais sobre esses, suas opiniões e seus contextos. Apesar da pré-análise considerar a formulação de hipóteses e saber que o esqueleto e o quadro teórico possibilitam ideias prévias

sobre o perfil do usuário da ecologia, esse estudo optou pelo distanciamento da imagem preconcebida nesse momento. Isso foi motivado pela perspectiva de serendipidade da pesquisadora, que proporcionou uma descoberta fortuita e não planejada, conforme será visto na sequência.

Logo, partiu-se para a exploração minuciosa do material sistematizado com o objetivo de ampliar o olhar sobre os entrevistados e compreendê-los de modo mais profundo. Esse momento implicou na identificação das unidades de contexto pelas unidades de registro (UR). Essas são as menores partes do conteúdo que direcionam o movimento de categorização e foram classificadas em duas frentes. A primeira – Personagem – foi composta pelas variáveis demográficas. Entre elas: nível socioeconômico, sexo, gênero, etnia, escolaridade, nacionalidade, idade, estrutura familiar, profissão, ocupação, etc. Essa unidade teve como propósito facilitar a identificação do grupo num contexto heterogêneo, buscando caracterizar de forma factual as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa. A segunda unidade de registro são os Temas, que se referem aos assuntos identificados nos discursos e consideraram opiniões, expectativas, motivações, medos e práticas dos entrevistados em relação aos aspectos estudados. Conforme Bardin (2016) e Rodrigues (2021) essa unidade de registro é considerada a mais antiga e indicada na AC. Ela permite o envolvimento de componentes racionais, ideológicos, afetivos e emocionais no processo. A presença, a ausência e a frequência de aparição dos diferentes temas apontaram informações importantes para o objetivo analítico desse estudo (CAREGNATO; MUTTI, 2006; BARDIN, 2016; RODRIGUES, 2021).

Quadro 5: Exemplo da planilha para exploração do material

Pergunta-chave	Unidade de Contexto (UC)	Unidade de Registro (UR)
Como você imagina esse usuário, quem é esse usuário.	Pesquisador	Personagem Ecologia (PE)
	Intenção de fazer uma pesquisa	Motivações
	Pular a parte de coleta de dados	Motivações
	Dados observacionais	Tipo de dado
	Dados que são únicos	Tipo de dado
	Agências de fomento	Personagem Suporte (PS)
	Revisores de periódicos	Personagem Suporte (PS)
	Vão acessar esses dados, mas não penso que o objetivo deles é o reuso	Personagem Suporte (PS)
	Não são os primeiros usuários que vem na minha cabeça	Personagem Suporte (PS)
	O primeiro que vem na minha cabeça é o usuário que vai sentar com aqueles dados e daqueles dados partir	Personagem Ecologia (PE)

	para uma nova pesquisa	
	Outros papéis de usuário	Personagem Suporte (PS)
	Podem estar querendo garantir o reuso de outros, mas eles não vão sentar e usar aqueles dados eles mesmos	Personagem Suporte (PS)

Fonte: elaborado pela autora, baseado no Quadro 4.

O alinhamento de sentido das UC e UR foi fundamental para organizar os trechos convergentes dos depoimentos, uma vez que a ideia sobre determinada pauta pôde ser expressa por meio de diferentes palavras dos indivíduos no decorrer das entrevistas, de modo não seqüencial (RODRIGUES, 2021). As unidades de registro foram primeiramente mapeadas e justificadas tendo como base o referencial teórico, o problema e os objetivos dessa tese, fio condutor do movimento dialógico necessário para a construção da persona. As perguntas-chave contempladas nos roteiros facilitaram a identificação dos eixos temáticos (UR) e das categorias. As categorias de análise foram estabelecidas conforme os princípios da pertinência e da produtividade (BARDIN, 2016; RODRIGUES, 2021).

A partir disso, a função “filtro de texto” foi utilizada na coluna de UR das planilhas para possibilitar a observação mais detalhada dos excertos. Esse movimento possibilitou identificar aspectos que comprovaram e complementaram o conhecimento contido em Barros (2005), Barros *et al.* (2005), Drucker (2011), Tabarelli *et al.* (2013), Luizão *et al.* (2013), Anjos (2016), GT – RDP Brasil (2018) e Vanz *et al.* (2018), como também identificar novos olhares sobre o contexto a partir dos conteúdos latentes. No transcorrer da exploração, percebeu-se que os discursos apontavam um segundo perfil de usuário que exerce influência no contexto e na promoção do acesso aberto aos dados de pesquisa. Esse é o arquétipo da Persona Suporte (PS), que não apareceu nas referências. Esse movimento definiu as categorias e subcategorias a serem analisadas, conforme disposto a seguir:

Quadro 6: Categorias finais de AC

	CATEGORIA: Persona Ecologia (PE)	CATEGORIA: Persona Suporte (PS)
Subcategorias	Perfil; Prática: da pesquisa; Prática: de compartilhamento e reuso; Prática: meios de compartilhamento e reuso; Tipo de dado produzido, compartilhado e reusado; Competência para compartilhamento e reuso; Motivações para compartilhamento e reuso; Medos/Insatisfações; Expectativas/Necessidades.	Perfil; Prática: meios de compartilhamento; Medos/Insatisfações; Expectativas/Necessidades; PE na visão da PS.

Fonte: elaborado pela autora.

Conforme Bardin (2016) as categorias de AC são um movimento de classificação das unidades de contexto com critérios previamente definidos. O processo de categorização representa o resultado da síntese das comunicações, destacando os aspectos mais importantes de cada depoimento. Por conseguinte, novas planilhas foram elaboradas para reagrupar os depoimentos dos entrevistados (as) nas categorias estabelecidas.

Quadro 7: Exemplo de planilha organizada para a AC

Subcategoria: Medos/insatisfações PE em relação aos serviços de acesso aberto e à prática do compartilhamento e reuso de dados de pesquisa		
ENTREVISTADO (A)	Pergunta-chave (PC)	Unidade de Contexto (UC)
E 4	Conhecimento e uso de linguagem de metadados para ecologia, <i>EML</i> ou outra	Se todas essas informações não estiverem bem descritas nos metadados, os pesquisadores ou as pessoas da área de consultoria que podem utilizar esses dados não vão ter clareza de como esses dados foram reunidos e muito menos de como esses dados foram calculados
E 4	Como/por que meios praticam o compartilhamento e o acesso aberto	A base de dados deles, assim, os servidores, eu não sei se estão espalhados ou se estão concentrados no <i>Ártico</i> Eu não tenho certeza se tudo está lá ou se grande parte dos dados estão lá. Eu não sei se fica no norte do Canadá ou no norte da Europa, mas assim é um dos dois lugares do <i>Ártico</i> , mas é isso. Então eu fico um pouco receoso de disponibilizar tudo por ali, porque se a base cai a gente tem alguns empecilhos assim. Eu, pelo menos, tenho um pé atrás com qualquer coisa que é <i>on-line</i>
E 5	Os pesquisadores brasileiros das Ciências Biológicas – ecologia têm expectativa que suas próprias instituições ofereçam o serviço de acesso aberto ou preferem utilizar um repositório internacional?	O único receio que eu tenho é que os repositórios não pertençam a instituições privadas que depois vão querer comercializar o dado. É essa a minha única preocupação

Fonte: elaborado pela autora.

A organização permitiu identificar padrões similares e complementares nas respostas de cada entrevistado (a). Conforme Goodwin (2002) ao encontrar aspectos similares, existe um padrão de comportamento e de perfil que forma a base da persona. Esses padrões são os dados factuais que embasam os arquétipos. As inferências e interpretações apresentadas posteriormente foram elaboradas nessa perspectiva.

A partir da análise foi possível identificar dois tipos de perfil de usuário. O primeiro, considerado a persona principal é a Persona Ecologia (PE), trata-se do arquétipo que produz, compartilha e reusa dados de pesquisa ecológicos. O segundo se trata da Persona Suporte (PS),

perfil secundário que busca garantir o processo de compartilhamento e reuso oferecendo subsídios de gestão, tecnologia, orientação, entre outros. Ambos os arquétipos serão apresentados na sequência, primeiramente em texto e logo, por meio de infográficos²⁶ que intencionam dar personalidade às personas. Compostos pelos dados coletados e analisados durante essa investigação, os instrumentos narram quem e como são os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas brasileiras que visam o acesso aberto aos dados de pesquisa. Trata-se do processo conclusivo da metodologia que esboça essas personas no contexto. Os instrumentos foram inspirados nos autores Grudin e Pruitt (2002), Pruitt e Adlin (2006), Cooper, Reimann e Cronin (2007), França (2012) e Sturmer *et al.* (2019b). Relembra-se que essa última referência não sugere formas de comunicação dos perfis. Porém, o modelo sugerido para a coleta de dados – Desenho da Persona – permitiu agregar valor às apresentações.

As categorias e subcategorias identificadas na AC guiaram a elaboração dos infográficos. Os excertos das entrevistas realizadas foram rearticulados e inseridos nos instrumentos, dando voz às personas. A ferramenta utilizada para a criação dos infográficos foi o Canva.com, disponibilizada gratuitamente na internet para os interessados. Foi desenvolvido um avatar para cada persona. As representações em formato de animação foram construídas pelos aplicativos *MojiPop* e *ToonArt*, disponíveis gratuitamente na *Play Store*. As imagens/retratos que representam os arquétipos levaram em consideração as características demográficas dos perfis para a busca.

²⁶ Recurso de comunicação composto por elementos visuais aliados ao texto enxuto e reduzido para transmitir uma mensagem (MÓDOLO, 2007).

4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS: A PERSONA ECOLOGIA E A PERSONA SUPORTE

Entende-se a terceira fase da AC como a etapa final da concepção dos perfis (PRUITT; ADLIN, 2006). Ela contempla as interpretações e as inferências dos dados da pesquisa. A primeira é embasada no quadro teórico dessa tese e nos dados do GT – RDP Brasil (2018) observados. A segunda é concebida a partir dos conteúdos manifestos e latentes dos discursos, que não aparecem nas referências. A análise considerou a presença (ou ausência), a frequência, a ordem e a co-ocorrência dos aspectos estudados e gerados pelos depoimentos (BARDIN, 2016; RODRIGUES, 2021).

Esta seção busca atender ao objetivo geral dessa pesquisa, demonstrando a compreensão da Persona de Ciências Biológicas – ecologia no âmbito brasileiro. Para tanto, contempla o objetivo específico (c) dessa investigação que trata de construir as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento de acesso aberto aos dados de pesquisa. Aqui são contempladas a Persona Ecologia e a Persona Suporte. Observa-se que as expressões utilizadas para as personas (ecologia e suporte) podem ser alteradas conforme estratégias, metas e objetivos das iniciativas e pela busca de um entendimento contínuo dos arquétipos.

4.1 PERSONA ECOLOGIA (PE)

Neste momento, o esqueleto da Persona Ecologia, apresentado na subseção 3.2 foi transformado em arquétipo a partir da articulação do quadro teórico e dos dados disponibilizados pelo GT – RDP Brasil (2018) com os coletados nas entrevistas semiestruturadas. A descrição completa foi composta por informações detalhadas a fim de dar-lhes personalidade e contexto, agregando aspectos e características ao perfil prévio. A PE se confirma enquanto público usuário real de iniciativas que visam o acesso aberto aos dados de pesquisa, uma vez que realiza pesquisas para obter determinadas informações frequentemente (LIMA, 1989; LARA, 2003; CONTI, 2003; MONFASANI; CURZEL, 2006; ROZADOS; PIFFER, 2009).

A PE se configura como *stakeholder* uma vez que afeta ou é afetada pelas iniciativas de acesso abertos aos dados de pesquisa, podendo ser vista como oportunidade ou ameaça conforme

a ocasião e a estratégia dos promotores (FREEMAN; 1984; FRIEDMAN, MILES, 2006). Ela também representa um público de *opinião*, pois possui interesses nos serviços prestados, é afetada por eles e influencia na efetividade desses, mas não necessariamente faz parte da estrutura interna das instituições que os promovem (SIMÕES, 2001). Trata-se de um público *consciente*, pois seus indivíduos são atingidos por um problema semelhante, reconhecendo sua existência e causa (GRUNIG; HUNT, 1984). O perfil também foi percebido como um público de *lealdade* que busca contribuir para conquista dos objetivos das iniciativas, sem interferir nas tomadas de decisão. Assim como, um público de *voz*, uma vez que é ativo na prática de compartilhamento e reuso de dados ecológicos (HIRSCHMAN, 1970).

4.1.1 Perfil PE

O arquétipo da ecologia se identifica com o gênero masculino, está entre os 40 e 50 anos de idade, atua como pesquisador, professor de cursos de graduação e pós-graduação, servidor público e possui doutorado, conforme o esqueleto. Os dados coletados nas entrevistas permitiram agregar características importantes na composição do *background* desse perfil. Entre elas: orientação heterossexual, branco, reside e trabalha na região Sudeste do Brasil, é casado e ter filhos é uma possibilidade refletida pelo casal. Ele trabalha com dados ecológicos ao longo de sua carreira, sua base de formação e conhecimento compete aos campos da biologia e da ecologia. A ecologia é uma ciência ampla e por isso ele precisa articular conhecimentos variados, leis da termodinâmica, óptica, acústica, matemática, álgebra, física, química entre outros. Essa realidade vai ao encontro do que citam Reichman, Jones e Schildhauer (2011) sobre a diversidade de disciplinas adjacentes que interagem com a pesquisa ecológica.

Ele ocupa cargos de coordenação/gestão em sua instituição e possui renda mensal de aproximadamente R\$ 15.125,00. “Coordenador [...] Gosto da liderança, mas gosto muito de compartilhar informação” (Entrevistado N° 2). “Ex-coordenador de programa de pós-graduação” (Entrevistado N° 5). “Professor universitário não faz somente pesquisa e extensão, também faz gestão acadêmica” (Entrevistado N° 7). Ele possui como características a criatividade e a resiliência, que influenciam na sua rotina de pesquisador. “[...] ser criativo e pensar em maneiras diferentes de usar aquele conjunto de dados” (Entrevistado N° 5). “Quem trabalha na minha área

não pode achar que vai fazer a mesma coisa eternamente [...] Todo dia é uma novidade, você deve estar preparado para isso” (Entrevistado N° 7).

A leitura é uma prática constante para a PE, não apenas sobre assuntos relacionados à sua pesquisa e à ecologia, como também literatura. Ele gosta de acompanhar alguma coisa de política e futebol, praticar algo relacionado à arte (teatro, dança, música, trabalhos manuais) e esportes. Gosta de ver filmes e programas informativos, geralmente acessa essas informações por meio de plataformas *on-line*. Porém, o acesso e o uso de redes sociais não é prática comum dele. A seguir algumas falas dos entrevistados que corroboram para o perfil:

Gosto de esportes, de futebol. No momento estou em crise porque meu time está lá embaixo na tabela. Gosto de ler, então eu compro livros e leio livros variados [...] Não só especificamente do meu tema, mas gosto de ler alguns livros de filosofia, leio muitos noticiários via UOL é o principal que eu leio hoje [...] notícias, editoriais do Estadão, da Folha [...] Eu leio bastante noticiário e livro [...] Fora artigos científicos (Entrevistado N° 2).

No meu dia a dia, eu leio bastantes livros [...] livros, streaming e Youtube [...] tento evitar rede social [...] vejo alguma coisa, ou relacionada com parte de programação ou parte de inglês, alguma parte de filmes para dar uma esparecida, assim, série [...] tenho mais lido do que assistido (Entrevistado N° 4).

Assuntos que eu tenho interesse, relacionados à ciência, uma outra coisinha de futebol [...] dicas de congresso, oportunidades de emprego para compartilhar com os alunos, para divulgar também o que são os novos trabalhos [...] Informações gerais fora da academia [...] eu acesso sites de notícias, principalmente de jornais consolidados [...] Globo News (Entrevistado N° 5).

Busco informações na internet, é básico [...] não uso o Facebook, nem Instagram e nem WhatsApp [...] Estou lendo hoje, por exemplo, Coração do Mundo, um livro de 2.500 páginas [...] qualquer pessoa que queira escrever bem tem que ler literatura [...] se você está com dificuldade de entender o livro de ecologia, pega um Dom Casmurro (Entrevistado N° 7).

A Persona Ecologia promove a ideia e a prática do compartilhamento e reuso de dados de pesquisa desde o início de sua carreira, já no mestrado buscou informações de outros investigadores. Ele é a líder do processo de compartilhamento e reuso, uma vez que ocupa posição de coordenação nos projetos de pesquisa em que atua. Atualmente, compartilha seus dados de pesquisa, reusa dados de outros pesquisadores e fomenta o acesso aberto para sua equipe, seus alunos e alunas.

Sempre foi comum para mim compartilhar, dividir o conhecimento, passar o conhecimento [...] característica de formação mesmo, de querer ajudar a formar novos profissionais (Entrevistado N° 2).

[...] orientador, é ele que encabeça a parte da disponibilização dos dados pelo *data paper* na *Ecology* ou por outras fontes de dados [...] ele é bem aberto para essa questão de colocar os dados, disponibilizar os dados [...] Eu reuso dados [...] tenho publicado a base de dados, então eu também compartilho (Entrevistado N° 4).

Sou fã de uso e compartilhamento [...] sou um grande fã, a gente gera muito dado aqui no Brasil e em outros lugares [...] sou um grande fã de gerar, de compartilhar e de usar [...] sou tanto um gerador de dados quanto usuário de dados gerados por outras pessoas [...] Tento compartilhar a filosofia que é legal usar o dado que está disponível, porque se ele está disponível quer dizer que as pessoas que geraram esses dados já estão confortáveis em torná-los disponíveis (Entrevistado N° 5).

Tem sido cada vez mais corriqueiro solicitarmos como resultado final da disciplina geração de conhecimento e artigo de coisas que ninguém fez, ou porque não teve tempo ou não teve a ideia (Entrevistado N° 7).

Sua disponibilidade para a promoção do acesso aberto gera estímulos positivos na equipe na qual trabalha, demonstrando sua influência no contexto. Os demais integrantes do grupo acabam se sentindo motivados para a prática do compartilhamento e reuso.

“todo grupo de pesquisa, principalmente da parte de ecologia, são mais abertos à disponibilização dos dados” (Entrevistado N° 4).

[...] tenho um aluno e uma aluna [...] de iniciação científica e com bolsa, ambos estão trabalhando com base de dados [...] Ambos estão super confortáveis com a ideia de trabalhar com um conjunto de dados que foi gerado por outro grupo [...] (Entrevistado N° 5).

A PE está atenta à transparência, à qualidade, à responsabilidade e à acessibilidade dos dados produzidos, compartilhados e reusados na pesquisa ecológica:

Tento compartilhar essa filosofia [...] da responsabilidade de gerar um bom dado, principalmente quando você tem financiamento. Depois a responsabilidade de usar bem esse dado enquanto ele está sob sua posse exclusiva. Depois a responsabilidade de compartilhar esse dado para que futuramente outras pessoas possam usar [...] Dentro da ecologia todo mundo está muito preocupado se a ciência é um processo transparente, se é acessível (Entrevistado N° 5).

Quais são as regras de uso dos dados compartilhados, como eu posso compartilhar dados sem colocar em risco a pesquisa daquela pessoa que produziu os dados? (Entrevistado N° 7).

Os textos a seguir, discorrem sobre a prática da pesquisa, do compartilhamento e do reuso de dados de pesquisa pela Persona Ecologia. As observações complementam o contexto desse arquétipo.

4.1.2 Prática: da Pesquisa PE

Esta seção relata como se caracteriza a rotina do dia a dia de pesquisa da Persona Ecologia. Os depoimentos permitem comprovar as considerações de Michener (2006), Sagarin e Pauchard (2010) e Drucker (2011), no que tange a variedade dos contextos onde a investigação ecológica é desenvolvida.

[...] mamíferos carnívoros de todo território nacional [...] algumas das espécies que a gente trabalha se estendem a todo continente americano [...] esses animais são marcados com colares de telemetria via satélite e essa informação é enviada para nós a cada uma hora, a localização do animal [...] no caso do protocolo TEAM [...] tem sítios em todos os países, praticamente todos os países tropicais. Então, eles fazem análises globais de impacto, por exemplo, de mudanças climáticas sobre a biodiversidade. [...] áreas de amostragem [...] várias unidades de conservação na Amazônia [...] reservas extrativistas, tem parques nacionais, florestas nacionais e reservas biológicas que são monitoradas [...] a equipe vai para o campo, para uma expedição de um determinado período, às vezes 15 dias, 20 dias, depende da área [...] quão acessível é aquela região [...] Pantanal, Estação Ecológica de Taiaimã [...] a partir do momento que está tudo instalado começa a coleta de dados, aí a gente começa a monitorar [...] A partir daí é tudo laboratorial, computador (Entrevistado N° 2).

Bioma da Mata Atlântica (Entrevistado N° 4).

Os meus campos são, predominantemente, feitos em riachos pequenos, nascentes [...] desde o Estado de São Paulo ao bioma Mata Atlântica [...] região costeira do Estado de São Paulo, ou os riachos que correm para o interior [...] conduzo as pesquisas aqui no meu laboratório, junto ao grupo de pesquisa, os alunos e as alunas [...] Os nossos objetos de pesquisa são, por exemplo, um grupo taxonômico preferido [...] borboletas [...] árvores [...] insetos [...] mamíferos [...] É uma possibilidade, ou ir ao campo ou manipular um experimento em laboratório (Entrevistado N° 5).

Nordeste, no Rio Grande do Norte [...] de ambiente praiano até caatinga [...] Brasília [...] cerrado [...] Manaus, Amazônia [...] água, floresta, microorganismos [...] Bioeconomia [...] Ecologia humana, ecologia de populações de bichos, de plantas [...] Ecologia do meio ambiente físico, vai de microorganismos até biosfera [...] Pico da Neblina, que é um parque nacional, que é uma reserva indígena [...] entender o que acontece na primeira camada atmosférica, que vai de 0 até 600 m de altura, dependendo do dia, da hora e da estação do ano [...] quem trabalha com modelagem e previsão, normalmente não vai para campo [...] sou experimental, sou campo [...] Maranhão, Amapá, Rondônia e Amazonas (Entrevistado N° 7).

As investigações envolvem saídas de campo, como também o trabalho interno, realizado em laboratórios e computadores, indo ao encontro do que afirma Zimmerman (2008). Conforme os relatos, as observações ocorrem ao longo das diversas regiões do Brasil, permitindo interpretar que os estudos ecológicos brasileiros são de larga escala, corroborando com o que afirmam Reichman, Jones e Schildhauer (2011) quando mencionam o âmbito internacional. Como também

que a PE aborda questões complexas que vão desde o gene à biosfera, conforme Michener e Jones (2012). As palavras-chave citadas pelos indivíduos também demonstram a abrangência dos aspectos estudados na ecologia. Entre elas: serviços ecossistêmicos; biodiversidade; carbono; ecologia do movimento; rádio telemetria; felinos; onça pintada; uso e ocupação do espaço; fragmentação de habitats; ecologia de comunidade; ecologia de anfíbios; modelos conjuntos de comunidades de espécie; ecologia; modelagem estatística; ecologia teórica, ecossistemas aquáticos e ecologia do meio ambiente físico. A única palavra citada mais de uma vez nos relatos foi biodiversidade (MARTINS; VANZ, 2022).

Aspecto importante que floresceu nos relatos foi o estabelecimento de parcerias para a realização das investigações ecológicas. Os dados do GT – RDP Brasil (2018) e a publicação de Vanz *et al.* (2018), que subsidiaram o esqueleto, mencionam que os dados são produzidos predominantemente em grupos de até cinco pessoas, mas não indicam outras formas de cooperação. Os excertos abaixo aprofundam o conhecimento sobre essa realidade no contexto da Persona Ecologia.

Parcerias internacionais (Entrevistado N° 2).

O que eu faço é buscar uma colaboração com um co-orientador ou uma co-orientadora [...] De uns tempos para cá, devido ao crescimento da qualidade da ciência brasileira, as parcerias se fortaleceram [...] Recentemente, a gente tem trabalhado mais como parceiros do que como fornecedores de dados [...] Desde quando eu sou aluno de doutorado eu tenho parcerias com colegas de outras instituições de fora do Brasil (Entrevistado N° 5).

[...] grande esforço do governo americano através da *National Aeronautics and Space Administration* (NASA) com o governo brasileiro, através de uma série de instituições, para fazer um estudo de longo prazo, pelo menos de cinco anos, para entender tudo que fosse relacionado com a interação da biosfera Amazônica com a atmosfera [...] dado era coletado no campo e transmitido via satélite para o *National Oceanic Atmospheric Administration* (NOAA) nos Estados Unidos [...] Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (IBGE), ela só tem 1.000 hectares, na época ela era muito estudada, tinham mais de 400 cientistas trabalhando (Entrevistado N° 7).

A partir dos depoimentos é possível sugerir que a cadeia de relações sociais estabelecida pela PE envolve uma série de atores, os quais se articulam e influenciam na prática da pesquisa. Sua rede é formada por parceiros nacionais e/ou internacionais que podem ser pesquisadores individuais, grupos ou instituições de pesquisa.

4.1.3 Prática: de Compartilhamento e Reuso PE

Neste texto são apresentados aspectos que complementam o conhecimento prévio obtido a partir do quadro teórico e dos dados que formaram o esqueleto da PE, sobre a prática do compartilhamento e do reuso de dados de pesquisa. Podemos interpretar que de modo geral, o arquétipo compartilha seus dados e reusa os de outros pesquisadores ao longo dos anos de atuação, sendo ator ativo no processo de acesso aberto, confirmando o conhecimento de Vanz *et al.* (2018).

Promove constantemente e já há bastante tempo [...] questões ecológicas que requerem o uso de outras fontes de informações a gente também busca essas fontes que estão compartilhadas [...] diariamente [...] na medida que eu vou recebendo esses dados, eu posso fazer esse *update* no *Movebank* diariamente [...] é uma coisa que vem sendo bastante discutida no meio acadêmico, no meio da pesquisa o compartilhamento de dados (Entrevistado (a) N° 2).

A gente reuniu durante o meu mestrado e agora estou trabalhando no doutorado, dados secundários, de artigos. Eu fiz uma grande busca de artigos que realizavam a amostragem de comunidades [...] a gente foi na coleção científica da *Dado anonimizado* e a gente compartilha para o digital (Entrevistado (a) N° 4).

Comecei a pensar nisso quando eu estava no doutorado. Meu doutorado foi de 2007 à 2010 e eu comecei a reutilizar conjuntos de dados que tinham sido gerados no laboratório que eu fazia parte [...] comecei a pensar nisso quando eu estava estruturando meu projeto [...] comecei a entrar em contato com pessoas que eu sabia que tinham dados parecidos com aqueles que eu precisaria e poderia usar para estabelecer parcerias. [...] comecei a pensar nisso, como é lidar com um conjunto de dados criados por outra pessoa? Esse conjunto de dados fica no computador dessas pessoas, porque não estão disponíveis? Porque não é fácil conseguir isso? [...] comecei a pensar que eu poderia usar os conjuntos de dados de uma outra maneira. Naturalmente, diferente da que tinha sido usada originalmente. [...] comecei a compartilhar quando eu virei docente, porque aí eu era o responsável por gerar o dado [...] depois que eu virei responsável por um laboratório (Entrevistado (a) N° 5).

Tive a honra de participar como aluno e como técnico, talvez da primeira iniciativa brasileira de acesso aberto aos dados que foi o Programa LBA. Ele começou a ser idealizado entre 96 e 97. De 98 a 99 ele foi implementado no Brasil. LBA é um acrônimo em inglês de Programa de Grande Escala da Biosfera-Atmosfera na Amazônia (Entrevistado (a) N° 7).

No entanto, ainda que a Persona Ecologia pratique o compartilhamento e o reuso de dados no decurso de sua carreira, os depoimentos a seguir possibilitam compreender que ela considera alguns critérios para a disponibilização. Ao observar os dados da *survey* do GT – RDP Brasil (2018) percebe-se um leve desacordo entre as respostas fornecidas na época e os depoimentos

atuais. Naquele momento os cinco indivíduos concordavam em compartilhar todos os seus dados de pesquisa em um repositório de acesso aberto sem restrições.

Os dados são abertos, você pode colocar algumas restrições, por exemplo, de que aqueles dados fiquem disponíveis a partir de uma determinada data, totalmente disponíveis. Ou, ele precisa ser requisitado ao coordenador do projeto, mas de forma geral os que a gente depositou nessas plataformas já estão todos disponíveis. [...] compartilha tudo [...] hoje, a gente consegue compartilhar os dados que a gente coleta. (Entrevistado (a) N° 2).

A gente não compartilha automaticamente. A gente só compartilha depois da tese ou da dissertação ter sido defendida e ou depois que o primeiro artigo foi publicado. A gente nunca compartilha o dado antes de usar o dado. [...] a maioria das pessoas disponibiliza o dado, somente depois que ela publica o artigo [...] Todos os conjuntos de dados que a gente gera, que estão sob minha responsabilidade, que foram gerados com financiamento público são, eventualmente, compartilhados [...] é sempre depois que os alunos e as alunas decidem que já usaram aquele conjunto de dados. Estão confortáveis. Ou quando a gente submete o primeiro artigo para a revista [...] É o dado que, geralmente, está associado a uma dissertação e a uma tese de doutorado. A pessoa que está responsável quer ser a primeira pessoa a usar esse conjunto de dados. Ela foi financiada para fazer isso. Enfim, ela quer gerar esse dado, quer usar, quer tornar público as suas descobertas. Ela quer fazer isso em primeira mão (Entrevistado (a) N° 5)

Existem níveis, o que pode ser liberado e o que não pode ser liberado. Isso depende dos acordos. [...] Se o dado está bloqueado pelo cara que produziu os dados, vai se perguntar: você pode liberar esses dados para fulano? [...] fazer um acordo se vai ter ou não co-autoria ou algo parecido. [...] Normalmente, todos os dados não são compartilhados [...] Herbários não tem problemas com isso, coleções biológicas não tem problema com isso. Colocou na coleção, não quero saber se você ainda não publicou. Você só vai colocar na minha coleção biológica aquilo que você está disposto a liberar de informação [...] A gente colocou como regra de negócio do sistema, que aquela informação molecular e de uso biotecnológico daquele organismo só vai estar disponível [...] depois de quem depositar disser: está liberado [...] para que a pessoa não fique segurando informação, pré-estabelecemos em sistema um padrão igual a proteção de propriedade intelectual, quando você deposita uma tese na biblioteca. Você pediu para bloquear parte da informação, ok. Essa parte da informação vai ser bloqueada por dois anos. Se você não pedir mais dois anos, ela vai ser automaticamente liberada. Depois do quarto ano você não pode pedir novamente para bloquear (Entrevistado (a) N° 7).

Desse modo, os relatos acima vão ao encontro do mencionado por Vanz *et al.* (2018) como motivo para não compartilhar os dados, confirmando que a PE entende que é preciso publicar os dados antes de compartilhar e sua permissão é necessária para acesso.

Pode-se inferir que o PE é responsável e sente confortável em fazer o envio dos dados de pesquisa para as plataformas de acesso aberto, pois entende que é o mais apto para a tarefa e para garantir a confiabilidade dos dados.

[...] esse líder de projeto vai fazer esse *upload* na sua conta pessoal, mas algumas coisas são a conta da instituição [...] eu preciso entrar, baixar essas informações do sistema, porque eu recebo isso via e-mail. Eu baixo essas informações, entro na plataforma do *Movebank* e lá na plataforma, no meu perfil, eu subo a informação, eu faço o *upload* da informação. [...] a gente precisa fazer o *upload* desses dados [...] Se são dados que eu conheço, eu prefiro eu fazer porque eu vou fazer as etapas da limpeza da informação, tudo direitinho [...] Às vezes, sobrecarrega um pouco eles terem que fazer esse *upload*, mas eu me sinto confortável fazendo [...] eu preciso sempre dar uma olhadinha [...] Consolidar a informação, simplesmente dar uma checada (Entrevistado N° 2).

Até o momento fui eu [...] Como eu trabalho mais um pouco essa parte de dados, de *GitHub*, essa parte da disponibilização eu meio que ajudo as outras pessoas [...] se eu, eventualmente, passasse isso para outra pessoa, talvez isso se perca ao longo do tempo (Entrevistado N° 4).

O usuário – o produtor do dado – é obrigado a colocar na plataforma, na base de dados (Entrevistado N° 7).

Porém, ele considera a importância de outros profissionais auxiliarem no processo de compartilhamento, conforme será visto na subcategoria Expectativa/Necessidades PE. Mesmo compartilhando seus dados, a PE mantém práticas internas nos grupos, equipes, instituições e respectivas pesquisas – para armazenamento dos dados.

A gente tem, às vezes, o mesmo dado depositado em alguns locais [...] A gente fez um repositório interno compartilhando no Google Drive essas planilhas (Entrevistado (a) N° 2).

[...] ao disponibilizar uma planilha de dados no Google eu sempre tenho ela física no meu computador (Entrevistado (a) N° 4).

A PE compreende a importância do uso de metadados na hora da prática do compartilhamento e reuso.

Vou descrever o dado "esse projeto é o projeto tal, da espécie tal, os responsáveis são fulanos, ele começou em tal ano e a previsão de terminar é tal ano" (Entrevistado (a) N° 2).

Os metadados são fundamentais. Um dado sem metadado, praticamente não tem muita utilidade (Entrevistado (a) N° 4).

Quando a gente compartilha o conjunto de dados ele não vai só a planilha. Inclusive, os repositórios pedem que você forneça o metadado, que é o dado do dado. Uma descrição de como aquele dado foi coletado, o que significa as unidades de medida. A temperatura foi em Celsius foi em Fahrenheit (Entrevistado (a) N° 5).

O padrão é sempre o mesmo. Dados mínimos, metadados dos dados e o seu miolo, que é específico para cada área (Entrevistado (a) N° 7).

A partir do exposto, é possível confirmar que a PE se caracteriza como público uma vez que ela é afetada e afeta a efetividade do processo. Configura-se, então como um público de *voz*, de *lealdade* e de *opinião*. Sua influência pode intervir nos resultados das iniciativas que visam o compartilhamento e reuso de dados ecológicos, seja de forma positiva ou negativa. A partir dessa realidade, compreende-se que as opiniões desse perfil deveriam ser consideradas importantes na tomada de decisão das iniciativas (público de *consulta*) (HIRSCHMAN, 1970; SIMÕES, 2001). A não distinção ou entendimento das especificidades desse ator influencia nas soluções a serem propostas, pois ao valorizar apenas as tecnologias utilizadas como pontos essenciais das estratégias as iniciativas possuem “visão turva” sobre a realidade desse arquétipo (MONFASANI; CRUZEL, 2006; ROSSI, 2012).

A instauração de parcerias também aparece nessa subcategoria, demonstrando mais uma vez a diversidade de agentes que se envolvem e influenciam na prática.

O principal é a formação de rede colaborativa [...] quando a gente fez o nosso trabalho de compilar dados de telemetria de onças pintadas, a gente reuniu pesquisadores do mundo inteiro para compartilhar esses dados (Entrevistado N° 2).

Parcerias [...] eles entraram como co-autores nesse artigo final que reuniu todos os dados [...] as relações sociais que são feitas através da disponibilização dos dados, elas acabam sendo bastante proveitosas, tanto para gente de cá que está juntando as informações, quanto para quem está de lá disponibilizando as informações (Entrevistado N° 4).

Parceria entre agências de fomento [...] uma chamada internacional [...] tem muita parceria que é independente da coleta de dados. Às vezes, um dado que você já coletou no passado, você pode reutilizar usando uma abordagem nova, mais criativa. Talvez você precise de um equipamento que você não tem. Aí estabelece uma parceria com uma instituição estrangeira. Ou você coleta um dado aqui no Brasil, isso faz parte de uma tese de doutorado, e o aluno quer fazer um estágio no exterior, isso também estabelece uma nova parceria. [...] Se eu conseguisse compilar vários conjuntos de dados dos meus colegas e das minhas colegas, naturalmente, esse pessoal topava e virava parceiro e parceira. [...] sempre tive parceiros e parceiras que me ajudaram (Entrevistado N° 5).

Projeto Observatório de Torre Alta (ATTO – *Amazon Tall Tower Observatory*) uma torre na Amazônia com 326 m de altura que tem cooperação com o *Max Planck Institute* da Alemanha [...] tem uma lista enorme de projetos que envolvem muitas pessoas, onde o compartilhamento de dados acaba acontecendo (Entrevistado N° 7).

Os fragmentos de texto vão ao encontro do que cita Albagli (2019) sobre o complexo universo de atores que são englobados na prática da ciência aberta. Essa investigação constata que as parcerias fazem parte da realidade da PE e demonstra sua importância no decorrer da prática da pesquisa, do compartilhamento e reuso de dados.

4.1.4 Prática: Meios de Compartilhamento e Reuso PE

Os mecanismos de preservação e compartilhamento utilizados pelos pesquisadores são fundamentais no processo de disponibilização e reuso de dados. O quadro teórico dessa pesquisa revela que os meios utilizados para o processo podem estar associados a instituições governamentais, organizações não-governamentais, universidades, bibliotecas, redes de observatórios ambientais, comerciais e empresas sem fins lucrativos (MICHENER; JONES, 2012). Que os repositórios digitais (RDs) são bases de dados *on-line* que armazenam arquivos de diversos formatos (IBICT, 2018a). No contexto da ciência aberta, os RDs intitulam os tipos de provedores de dados que gerenciam as informações da ciência e formam um tipo alternativo de comunicação científica. Esses repositórios de acesso aberto constituem um sistema de informação (SI) científica, em ambiente digital, que reúne, armazena, organiza, preserva, recupera e dissemina a informação (LEITE, 2009). SIs organizam os dados de pesquisa, englobam os processos de coleta, processamento, armazenamento e recuperação de informação pelos indivíduos interessados (TARAPANOFF, 1995, p. 126).

A partir dos dados coletados nessa pesquisa é possível interpretar que os mecanismos de compartilhamento utilizados correspondem aos conceitos citados. Os trechos abaixo apontam alguns desses:

O governo federal vem trabalhando muito com base de dados de biodiversidade, que também seriam uma base de dados aberta [...] *GitHub* que é a plataforma que a gente compartilha tanto dados, quanto os códigos que a gente escreve para fazer as análises [...] repositórios internacionais, alguns deles como *Dryad*, *Zenodo* [...] armadilhas fotográficas a gente usa a plataforma *Wildlife Insights* que nos auxilia a fazer a identificação das espécies [...] É uma plataforma específica para o protocolo TEAM (*Tropical Ecology Assessment & Monitoring Network*). [...] Você pode aportar qualquer dado de armadilhamento fotográfico nessa plataforma. E a gente compartilha então nessa plataforma já os dados imediatamente. [...] *Wildlife Insights* que é para as imagens de fotografia. E *Movebank.org* que é para os dados de telemetria [...] *Dryad* é uma que a gente pode acessar bastante com os dados também disponibilizados. O *Dryad* tem *Digital Object Identifier (DOI)*, então a gente vai pelo *DOI* e já acha a informação [...] A gente deposita no Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) (Entrevistado N° 2).

Grade dos *scripts* que eu gero e trabalho disponibilizados na base do *GitHub* [...] dados espaciais, de ecologia de paisagem, a gente basicamente entrou na base da Fundação Brasileira de Desenvolvimento Sustentável (FBDS) [...] (CRIA) Centro de Referência em Informação Ambiental e o *speciesLink*, eu acho que é uma grande base de dados e mais recentemente tem o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR) [...] são servidores *on-line* que têm esses dados [...] a gente tem usado, justamente, o *Mapbioma* [...] uma base fundamental que está na Google é tudo processado pela Google [...] outras informações que eu retiro dos artigos, eles vem de buscas sistematizadas e não sistematizadas [...] *Web of Science*, *Scopus* e o próprio Google Scholar que traz informações. Então não necessariamente a gente tem uma fonte

só [...] a gente tem um que é mais artigo e outro que é mais base de dados on-line que ou está em servidor *on-line* ou em algum servidor de alguma ONG, que a gente acessa e pega esses dados. [...] *GitWeb* [...] *Zenodo* (Entrevistado N° 4).

Zenodo que é um lugar bastante simples, comporta diferentes formatos de dados. Os que eu compartilhei estão lá. [...] Alguns deles são pagos e algumas dessas revistas inclusive pagam o serviço. [...] repositório que se chama *Figshare*. Esse repositório é pago e algumas revistas pagam [...] repositório que eu já usei chamado *Zenodo* é livre, é um repositório geral [...] iniciativa que chama *Open Science* que eu acho que também serve como repositório. Eu não tenho certeza absoluta [...] outro que não é um repositório, é mais um lugar de controle de versões e de desenvolvimento de códigos chamado *GitHub*. É mais para código, mas ele suporta conjuntos de dados pequenos [...] Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR), que é uma plataforma brasileira baseada nesses formatos internacionais, baseada no *Darwin Core* [...] A FAPESP, no final dos anos 90, criou um programa que depois virou um programa institucional chamado BIOTA-FAPESP. A ideia era fazer um levantamento da biodiversidade de São Paulo. [...] vários grupos, peixes, aranhas, borboletas. O que você imaginar [...] Associado ao BIOTA-FAPESP também foi criado um sisteminha de armazenamento de dados. Você entra nesse sistema e digita, por exemplo, o nome de uma espécie. Ele gera um mapa com todos os locais de ocorrência dessa espécie, quem coletou, associado à qual projeto [...] Na Biologia e na Ecologia, especificamente, tem várias iniciativas nacionais, porque a gente gera muito dado. Nosso campo é coletar bichos, coletar plantas (Entrevistado N° 5).

Normalmente, essas bases de dados são próprias [...] *speciesLink* – o mais famoso [...] o relatório do Painel Intergovernamental sobre Mudanças Climáticas (IPCC, na sigla em inglês) ele é baseado em plataformas como a do Programa de Grande Escala da Biosfera-Atmosfera na Amazônia (LBA) [...] têm suas próprias bases que se conectam com outras bases, que pegam ou o metadado, ou o dado certificado, ou o dado bruto [...] *speciesLink* é maravilhosa. Não é bem um repositório, é um sistema que conecta repositórios diferentes e a pessoa consegue trabalhar ou com os próprios metadados dessas coleções, ou com os próprios dados brutos [...] link de um dropbox da vida que você vai e pega os dados que eles coletaram [...] dados coletados por equipamentos e vou colocar numa base geral, em regra, esse tipo de base permite que o depositante diga qual é o nível de liberação dos dados (Entrevistado N° 7).

Entre os mecanismos internacionais mencionados apareceram: *Figshare*, *GitHub*, *GitWeb*, *Dryad*, *Movebank*, *Open Science*, *speciesLink*, *Zenodo* e *Wildlife Insights*. Entre os brasileiros: o do Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio), as bases da Fundação Brasileira de Desenvolvimento Sustentável (FBDS) e do Centro de Referência em Informação Ambiental (CRIA), a iniciativa MapBiomas do SEEG/OC (Sistema de Estimativas de Emissões de Gases de Efeito Estufa do Observatório do Clima), o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR), sistema do BIOTA-FAPESP e a base do Programa de Grande Escala da Biosfera-Atmosfera na Amazônia (LBA). Exemplos de interfaces na figura 16:

Figura 16: Exemplo de interfaces de dois mecanismos utilizados pela PE

Rede *speciesLink*

MRM • Sair • Português ▾

speciesLink

ALCB - Herbário Alexandre Leal Costa

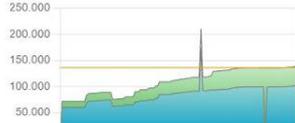
UFBA - Universidade Federal da Bahia
Departamento de Botânica
Salvador - BA - Brasil

O Herbário ALCB é um dos mais importantes do estado pois têm em seu acervo espécimes representantes de todos os ecossistemas brasileiros e coleções botânicas históricas, como as do Padre Caminhoá e Pierre Verger. Atualmente, o acervo consta de 109.700 exsicatas, entre as quais 87.603 são fanerógamas, 11.731 briófitas, sendo oito espécimes tipo oriundos da Floresta Atlântica da Bahia, 7.108 algas, 1.900 Pteridófitas (Lycophyta e Monilophyta), 555 fungos e 803 amostras de madeiras, grande parte referente aos diversos biomas da Bahia, e em menor número a outros estados do nordeste ou doações de outras regiões do Brasil. As classificações utilizadas para a organização do acervo levam em conta aquelas propostas por: Wynne (2011) para algas; Goffinet et al. (2008) para Bryophyta; Crandall-Stotler et al. (2009) para Marchantiophyta; Duff et al. (2007) para as Anthocerotophyta; Smith et al. (2004) para Pteridófitas e APG III (2009) para as Angiospermas.

Online desde: 19/07/2011
Última atualização: 27/10/2022

Número de registros

Total: 136.087 Online: 138.661 Georreferenciados: 101.807 (por município: 27.048) Com Imagens: 51.785






SiBBr

gov.br

SiBBr SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA

Sobre ▾ Atlas ▾ Participe ▾ Ajuda ▾ GBIF ▾ PT ▾

🔍 Pesquisar ...

Home > Coleções > Conjuntos de dados

SiBBr conjuntos de dados

Grande parte do conteúdo no SiBBr, tais como os registros de ocorrência, dados ambientais, imagens e estatuto de conservação das espécies, vem de conjuntos de dados fornecidos por instituições de coleta, coletores individuais e grupos comunitários. Estes conjuntos de dados estão listados nesta página.

Exibindo 531 dadosconjuntos.

Buscar Repor lista Baixar

Resultados por página: 20 Ordenar por: name Ordem de classificação: ascending ▾

- ▶ **Acari Collection - Instituto Nacional de Pesquisas da Amazônia**
Tipos de recursos: records Licença: Ver registros
- ▶ **Amblypygi Collection - Instituto Nacional de Pesquisas da Amazônia (INPA)**
Tipos de recursos: records Licença: Ver registros
- ▶ **Amostragem de florestas secundárias no Cerrado Jalapão**
Tipos de recursos: records Licença: CC-BY-NC Ver registros

Fontes das imagens: SPECIES LINK (2022) e SiBBr (2022).

O Programa Pesquisas Ecológicas de Longa Duração (PELD), citado por Barros *et al.* (2005) e Pavão *et al.* (2018a) foi mencionado em um dos relatos. Porém, nada sobre a Biblioteca Digital Georreferenciada (BDiG-PELD) (BARROS, 2005).

O projeto Ecológico que se chama PELD, Programa Ecológico de Longa Duração. É um tipo de projeto que existe em vários outros países, que é a ideia de financiar projetos de longo prazo para tentar monitorar como as coisas estão mudando ao longo do tempo [...] o PELD 1 que é o PELD da Amazônia deve ter uns 20 anos, eu acho. Ou mais até, se bobear uns 30 anos (Entrevistado (a) N° 5).

Um meio de compartilhamento não mencionado no quadro teórico que apareceu na pesquisa qualitativa foi o *data paper*. Não se trata de um repositório ou sistema de informação, mas de um instrumento que possibilita a comunicação de dados de pesquisa, já que seu objetivo principal é descrevê-los, em vez de relatar a investigação e seus resultados. *Data paper* é um documento pesquisável de metadados que explica um conjunto de dados acessível *on-line*, publicado de acordo com as práticas acadêmicas comumente aceitas. Esse meio deve incluir elementos que documentam como ocorreu a coleta, taxonomias, intervalos espaciais e temporais, cobertura regional dos registros, quem coletou, quem possui os dados, qual *software* foi usado para criar os dados, ou qual é indicado para visualizá-los, etc. Outros objetivos do *data paper* são: fornecer uma publicação de periódico citável que traga crédito acadêmico, descrever os dados de forma estruturada e legível por humanos e trazer a existência dos dados para a atenção da comunidade científica. É possível inferir que *data paper* é estratégia eficaz e útil para acesso a uma série de dados e informações relacionadas à biodiversidade e aos ecossistemas (CHAVAN; PENEV, 2011).

A gente já fez publicações de *data paper* que é uma nova modalidade de publicação que a gente vem contribuindo para diferentes *data papers*. [...] publicado em *data paper* em 2018 e está disponível para a comunidade [...] são os artigos publicados em forma de dados [...] a gente tem usado muito são os *data papers* [...] *Ecology*, que é uma revista muito importante na área da ecologia, que tem disponibilizado uma parte da revista para os *data papers* (Entrevistado (a) N° 2).

A gente publicou uma série desses *data papers*, que reúne essas informações principalmente das Comunidades, mas também alguns processos [...] no modelo de publicação desses dados, exige que você publique junto metadados e aí tem a descrição de cada coluna. Então tem o nome da coluna e a descrição das informações (Entrevistado (a) N° 4).

Tem revistas agora que tem uma sessão específica para a publicação de dados [...] Na minha área tem sido comum a publicação de um tipo de artigo chamado *data paper*. É um artigo sobre dados. Não é um artigo que tem uma hipótese, que tem um experimento. É uma descrição de um conjunto de dados. É um dado legal, geralmente, um dado que envolve uma compilação de muitos outros dados (Entrevistado (a) N° 5).

O que estou tentando fazer para que os dados fiquem públicos de alguma forma é tentar encontrar revistas que aceitem depositar os dados na revista. O que é uma novidade que tem só há uns três ou quatro anos. Você publica o artigo e coloca no apêndice os dados brutos. Como o meu laboratório não tem uma plataforma, um repositório onde eu coloque e disponibilize de uma forma mais controlada para quem quiser acessar, se a

revista aceitar colocar como anexo a planilha. Aqui estão meus dados, se você quer verificar se o que eu fiz está certo ou errado? Toma. (Entrevistado (a) N° 7).

O *data paper* deve apresentar todos os dados coletados pela pesquisa, não apenas os dados usados nos artigos. Ele é um instrumento que promove o compartilhamento de dados e traz benefícios aos pesquisadores. Entre eles: aumentar a visibilidade do trabalho de um autor e melhorar as taxas de citação, demonstrar suas contribuições para a ciência e o subsequente reconhecimento dos pares que influencia na sua reputação, oportunidades de emprego, promoção no trabalho e capacidade de garantir financiamento (OSAWA, 2019). Desse modo, infere-se que o *data paper* promove a disseminação científica, uma vez que possibilita a comunicação da informação científica e tecnológica transcrita em códigos especializados para um público especializado (ALBAGLI, 1996; BUENO, 2010).

Entre os mecanismos e iniciativas citadas no quadro teórico dessa pesquisa, não apareceram nos meios utilizados e conhecidos pela PE para compartilhamento e uso de dados os seguintes: o do *International Biological Program (IBP)*, os das redes *U.S. Long-Term Ecological Research (LTER)* e *International Long Term Ecological Research Network (ILTER)*, o do *National Center for Ecological Analysis and Synthesis (NCEAS)*, o sistema do *Ocean Biogeographic Information System (OBIS)*, a *Ocean Observatories Initiative (OOI)* e a iniciativa da *Biodiversity Observation Network (BON)*. Os relatos não apontam aspectos que justifiquem ou possam gerar inferências sobre a ausência de citações sobre essas iniciativas e meios. O que de fato os depoimentos indicam é que a PE utiliza diversos mecanismos para a prática do compartilhamento e reuso de dados de pesquisa.

Os aspectos que influenciam no processo decisório da PE em relação à prática e a qual mecanismo utilizar no processo de compartilhamento e de reuso é fluído, acontecendo espontaneamente na rotina do dia a dia das investigações.

Foi tão natural [...] não é uma forma que a gente teve um evento para discutir isso e formalizar [...] É um acordo coletivo a gente sentiu a necessidade de fazer [...] uma demanda de todos os membros da equipe, que a gente precisava interagir mais com os diferentes grupos, inclusive compartilhando dados [...] a gente faz reuniões de planejamento periódicas, tanto de planejamento como de monitoramento do planejamento estratégico, isso veio numa discussão bem natural e não há nenhuma divergência em relação a isso, felizmente não tem divergência (Entrevistado (a) N° 2).

Se algum dia a gente sentou e discutiu qual é a melhor base, qual a base de dados é mais confiável para disponibilizar os dados, a gente não fez isso. [...] São vários fatores que influenciam a decisão, o principal, talvez, seria o conhecimento da própria plataforma

que eu vou disponibilizar os dados. O segundo, talvez que eu pensaria é se essas bases de dados são confiáveis [...] eu não confio muito nessas bases de dados que são apenas *on-line* [...] Porque eu não sei se o sistema vai cair ou se alguma coisa vai acontecer (Entrevistado (a) N° 4).

Sempre dou uma olhada no que o pessoal está falando da segurança, da qualidade, do armazenamento, se é uma iniciativa sem fins lucrativos. Tem uma comunidade bastante ativa de ecologia, principalmente dessa área de biodiversidade no *Twitter*. Sigo algumas pessoas que estão envolvidas com iniciativas parecidas e que são defensoras de compartilhamento de base de dados [...] vejo o que as pessoas estão falando, recomendando e os motivos que elas citam (Entrevistado (a) N° 5).

Assim sendo, é possível interpretar que esse arquétipo é parte ativa e influencia no movimento da ciência aberta. Ele empreende a abertura do processo científico e a translação do conhecimento, incentiva a transparência da pesquisa científica e estimula a reflexão sobre atividades, processos, métodos, avaliações e produções, visando colaboração e abertura de dados. Nessa conjuntura infere-se que esse movimento da PE possibilita ampliar os impactos sociais e econômicos da ciência, reforçando o conceito de responsabilidade social científica (SANTOS *et al.*, 2017, SILVA; SILVEIRA, 2019; FIOCRUZ, 2022).

4.1.5 Tipo de dado produzido, compartilhado e reusado PE

O quadro teórico ressalta a diversidade e o crescimento exponencial do volume de dados de pesquisa produzidos na ecologia. Zimmerman (2008), Sagarin e Pauchard (2010), Reichman, Jones e Schildhauers (2011), Hunt, Baldocchi e Van Ingen (2011), Michener e Jones (2012), Anjos (2016), Vanz *et al.* (2018) e os dados do GT – RDP Brasil (2018) demonstraram que a Persona Ecologia produz e reutiliza uma grande variedade e quantidade de dados. Que esses dados são originalmente observacionais, coletados e organizados de muitas maneiras, incluindo registro manual de observações no laboratório e no campo por meio de blocos de anotações, gravadores e computadores portáteis; e a coleta automatizada de dados por meio de instrumentação de laboratório e de campo. Entre os tipos de dados citados pelo conjunto de autores estão: observações de espécies, de laboratório, de simulação, observacionais, computacionais, experimentais, documentais, dados de campo, plantas, solo, medidas químicas da água, instâncias de controle biológico, história, características da floresta, anéis de árvores, bruto ou derivado, coleções físicas, modelos, imagens, textos, questionários, etc. As entrevistas

comprovam a heterogeneidade e a magnitude dos dados ecológicos e agregam informações ao conhecimento gerado.

Dados de movimento de animais [...] algumas amostragens eram feitas em pequenos fragmentos florestais [...] varia bastante [...] questões ecológicas que requerem o uso de outras fontes de informações a gente [...] inclusive simula o compartilhamento de colegas para que a gente possa gerar esse produto que nos interessa [...] quando a gente compartilha dados de monitoramento por telemetria, são dados de localização dos animais em intervalos de tempo regular [...] são coordenadas geográficas. Então você tem a data, a hora e a coordenada [...] esses dados são específicos, [...] uso e ocupação de espaço, interações ecológicas em termos de interação intra e inter específicas. Então são várias interações estudadas com esses dados de telemetria [...] pode ser um fragmento florestal, pode ser numa unidade de conservação na Amazônia, na Mata Atlântica, no serrado, em qualquer unidade [...] duas principais formas que a gente tem hoje de dados que são compartilhados, de armadilhamento fotográfico e de monitoramento por sistema de telemetria, GPS, satélite. [...] dados volumosos [...] da última vez que eu vi os dados de armazenamento fotográfico nossos, depositados no *Wildlife Insights* são mais de dois milhões de imagens depositadas [...] só nossos são mais de dois milhões de imagens [...] os dados são brutos (Entrevistado N° 2).

Dados de métricas de paisagens são tamanho dos fragmentos, a distância entre os fragmentos, a quantidade de conectividade [...] o dado final, ele tem que ser um pouco mais “mastigado”, um pouco mais trabalhado para que ele preencha os campos, as colunas, as linhas, o formato do dado ele esteja um pouco mais sinalizado para que quem for utilizar, é, também não fique com várias colunas com nomes diferentes, ou com informações diferentes [...] massa de dados em formato de *data papers* na Ecology [...] É uma tabela de Excel, uma tabela .csv, eles ainda não estão estruturados para que alguém conecte num banco de dados e comece utilizar os dados. [...] os dados são sobre anfíbios ao longo da Mata Atlântica toda [...] tempo que o pesquisador ficou amostrando, que método que ele utilizou, que habitat que ele amostrou, informações e características de localização, como informações geográficas de longitude e de latitude, além do nome do local, nome da cidade, município, estado [...] dados de comunidades de anfíbios, amostragem dessas espécies, dados morfométricos, dessas espécies que foram encontradas e dados de ecologia de paisagem [...] *shapefiles*, que são arquivos que a gente entra lá no geoprocessamento que são vetores, onde está a floresta [...] geoprocessamento com QGIS e com Grass GIS, que é um outro *software* (Entrevistado N° 4).

Coletamos amostras de mais de 100 riachos diferentes aqui numa região do interior do Estado de São Paulo [...] insetos que vivem em riachos e rios [...] baleias [...] uso de base de dados ou simulações computacionais [...] peixes em riachos [...] uma análise baseada em sistema de informação geográfica para escolha dos locais, o uso do solo se é no meio da cana-de-açúcar, no meio da floresta, se está muito perto de uma cidade. Esse tipo de dado é gerado antes no laboratório. [...] No campo a coleta envolve mesmo a captura de organismos, que são amostrados com equipamentos específicos, armazenados, preservados e depois a gente traz isso de volta para o laboratório para poder identificar quais são as espécies, quantas tem. [...] outras medidas já terminam ali, porque são equipamentos que geram informações instantâneas. Por exemplo, sensores de temperatura, de ph, salinidade. Isso é automático, isso é um aparelho. Ele tem uma memória. É um pequeno computador que já gera o conjunto de dados. Isso é armazenado num formato típico daquele aparelho e depois a gente trata isso em programas de planilhas de dados [...] armazenado em formato de planilha [...] em vários formatos [...] Excel, txt, csv [...] planilhas de fácil manuseio e de fácil análise (Entrevistado N° 5).

Tem o dado bruto e o dado certificado [...] tem redes de coletas de dados múltiplas [...] geomorfologia, solo, relevo [...] dados de herbário [...] microorganismo exige uma checagem de informação, a comprovação molecular, você tem que fazer o DNA [...] nível de precisão [...] Fotos sempre ajudam a quem vai ver os dados, entender os dados (Entrevistado N° 7).

A partir do exposto é possível elucidar que a ecologia aborda questões em escalas espaciais, temporais e de organização biológicas, mais amplas. Essa realidade comprova as mudanças nos tipos, magnitude e complexidade dos dados de pesquisa adquiridos e analisados pela Persona Ecologia. Percebe-se que o uso de equipamentos, satélites, sensores remotos e embutidos, entre outros influencia diretamente nesse contexto. Os depoimentos comprovam o que foi citado por Tabarelli *et al.* (2013) e Luizão *et al.* (2013) sobre a diversidade de tipos de dados utilizados pelos pesquisadores brasileiros.

Confirma-se que os dados produzidos pela ecologia são heterogêneos e que podem ser classificados como *observacionais*, *experimentais* e *computacionais*, indo ao encontro das definições de Sayão e Sales (2014). Isso devido às suas características, os *dados observacionais* são os coletados por observação direta, em lugares e tempos específicos e possuem natureza instantânea. Os *dados experimentais* são obtidos em laboratório. E os *dados computacionais* são resultados da execução de modelos de computador ou de simulações.

Verifica-se também que os dados ecológicos podem ser classificados conforme a fase de pesquisa, uma vez que se enquadram nos conceitos de *dados brutos* vindos diretamente dos instrumentos científicos utilizados para coleta e *dados derivados*, uma vez que são resultados do processamento ou da combinação de dados brutos ou de outros dados (SAYÃO; SALES, 2015). Os dados também podem ser entendidos como *dados de pesquisa*, já que se tratam de registros factuais em diversos formatos (scores numéricos, documentos textuais, sons, planilhas de estatísticas, dados retirados de bases de dados públicas ou privadas, registro sonoro de uma entrevista, registro descritivo de um espécime de rocha, uma coleção de imagens digitais) utilizados como fontes primárias para a pesquisa científica. E, são unidades de informação obtidas no decorrer da pesquisa científica formatadas para torná-las apropriadas para comunicação, interpretação e processamento por computador (OECD, 2007; CIDACS, 2022). Pode-se inferir que os dados da pesquisa da ecologia estão entre os níveis 3 Estrelas – conjuntos de dados publicados em algum formato aberto (*Open Format – OF*), como por exemplo, *Comma*

Separated Value (CSV); e 4 Estrelas – uma vez que permitem criar ligações e reusar os dados disponibilizados (RAUTENBERG, 2018).

No entanto, os dados ecológicos ainda não se enquadram totalmente no conceito de *dados abertos*. Apesar de possuírem formatos variados que permitem comunicação, interpretação ou processamento visando gerar novas pesquisas; poderem ser usados, reutilizados e redistribuídos; e serem um tipo de objeto digital armazenado que pode ser encontrado através de mecanismos de busca, eles não necessariamente são acessados e compreendidos por qualquer pessoa livremente (PARHAN; BODNAR; FUCHS, 2012; DECRETO N ° 8.777, 2016; MEDEIROS, 2018; GLOSSÁRIO FIOCRUZ, 2022). Essa percepção é embasada nos depoimentos sobre como é promovida a prática, os meios e instrumentos utilizados para compartilhamento e reuso de dados da PE, que são prioritariamente direcionados a especialistas.

Essa realidade influencia diretamente no contexto para a prática da ciência aberta no campo da ecologia. Assegura-se que a variedade de tópicos estudados, dos tipos de dados coletados e das formas de armazenamento desses, afeta diretamente no processo de compartilhamento e reuso de dados de maneira integrada, fazendo do acesso aberto aos dados de pesquisa um desafio, quando tratado das Ciências Biológicas – ecologia (REICHMAN; JONES; SCHILDHAUER, 2011). Os relatos vão ao encontro do que afirmou Barros (2005), sistemas de informação em biodiversidade devem ser capazes de manipular dados heterogêneos e de gerenciar grandes bancos de dados referentes às espécies. Desse modo, compreende-se que essa conjuntura demanda políticas que identifiquem e sustentem essa diversidade, uma vez que os dados ecológicos possuem uma heterogeneidade intrínseca (SAYÃO; SALES, 2014).

4.1.6 Competência para compartilhamento e reuso PE

Como comprovado até o momento, a pesquisa ecológica é integrada por uma série variada de práticas, instrumentos, ambientes, objetos de estudo e tipos de dados. Junto a isso, confirmou-se que a PE costuma exercer um papel importante no processo, demonstrando-se ativa ao longo dos anos. Os depoimentos comprovam também que a PE possui certo domínio sobre o processo, demonstrando competências, conhecimentos e habilidades que influenciam na prática do acesso aberto aos dados de pesquisa. “A gente tem capacidade técnica de análise de dados” (Entrevistado (a) N° 2). “Tenho bastante domínio na parte de conjunto de dados” (Entrevistado

(a) N° 4). “Dou aula sobre isso. Sobre como organizar um conjunto de dados de modo que uma pessoa que não gerou esse conjunto de dados consiga entender o que é esse conjunto de dados.” (Entrevistado (a) N° 5). “Nós temos profissionais competentes” (Entrevistado (a) N° 7).

No entanto, quando questionado sobre o conhecimento e uso de uma linguagem própria para descrever os dados de pesquisa compartilhados – *Ecological Metadata Language (EML)* ou outra – percebeu-se que isso ainda não é costume desse arquétipo e quando conhecida não é utilizada com frequência no momento de catalogação.

Particularmente não conheço (Entrevistado (a) N° 2).

Não (Entrevistado (a) N° 4).

É a principal que o pessoal usa na minha área. Como eu conheço um pouquinho da *EML*, eu tento usar os mesmos campos que estão ali, mas eu não uso *EML*. Deveria (Entrevistado (a) N° 5).

Uns colegas já falaram sobre, alguns colegas da minha universidade, mas eu nunca vi, nunca usei esse padrão (Entrevistado (a) N° 7).

A pouca assiduidade da aplicabilidade do padrão parece ser motivada mais pela falta de conhecimento e familiaridade por parte da PE, do que qualquer outra justificativa. A *EML* é considerada a abordagem mais indicada e útil para caracterizar os dados heterogêneos da área. As especificações formais de metadados fornecem orientação para descrever consistentemente os diversos objetos e tipos de dados (REICHMAN; JONES; SCHILDHAUER, 2011). Desse modo, é possível inferir que esse aspecto colabora ainda mais para a disparidade das informações no campo, o que implica diretamente no processo de localização, acessibilidade, interoperabilidade e reutilização dos dados ecológicos – princípios *FAIR* (*Findable, Accessible, Interoperable, Reusable*).

A falta de padronização dos dados foi mencionada como um dos motivos para o não compartilhamento por 28, dos 105 respondentes da ecologia na *survey* do GT – RDP Brasil (2018) e por Vanz *et al.* (2018). Essa investigação complementa o conhecimento das referências e indica a importância desse fator no processo. Nas subcategorias Medos/Insatisfações e Expectativas/Necessidades esse assunto também aparece, como será visto na sequência do texto.

4.1.7 Motivações para compartilhamento e reuso PE

Esta seção apresenta as motivações do Personagem Ecologia em relação ao compartilhamento e reuso de dados de pesquisa. Os dados do GT – RDP Brasil (2018) e a referência de Vanz *et al.* (2018) indicam algumas das motivações desse usuário. Entre elas estão: a publicação em periódicos ou em anais de eventos, a exigência do principal financiador, a solicitação da instituição na qual atuam, a criação de novos conjuntos de dados, a ajuda a outros pesquisadores, a ampliação do reconhecimento acadêmico, a divulgação da pesquisa e o aumento das citações. Anjos (2016) também aponta que o usuário considera importante a citação quando os dados são reutilizados. Algumas delas são identificadas nos relatos dos entrevistados.

[...] a gente tem utilizado isso e produzido informações com base em dados depositados [...] Quando você compartilha esses dados, você também é procurado por outros grupos para integrar diferentes projetos [...] você mostra, ao mesmo tempo em que você está publicando artigos com aqueles dados e aqueles dados vão compor uma rede maior [...] muitos pesquisadores vêem a qualidade do seu trabalho e te chamam para compor a equipe (Entrevistado N° 2).

[...] utilizar esses dados para responder outras perguntas ecológicas [...] tem dados de comunidades ao longo do bioma, esses novos pesquisadores podem fazer outras perguntas tanto biológicas, quanto ecológicas para responder reutilizando esses dados que já estão disponibilizados [...] a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), ela exige que você antes de começar um projeto já disponibilize. Tem uma forma de disponibilização dos dados [...] a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) também tem, depois de certo tempo você tem que disponibilizar [...] (Entrevistado N° 4).

Se você ceder os seus dados você vai ser um co-autor e co-autora dos manuscritos, dos artigos [...] Isso é legal, pois gera uma citação para as pessoas que organizaram o conjunto de dados e estão disponibilizando. As pessoas que fizeram isso é um serviço difícil, elas ganham algum crédito, pelo menos [...] o conjunto de dados está acompanhado da citação do artigo que usou originalmente aquele dado (Entrevistado N° 5).

Assim sendo, é possível interpretar que a abertura de dados potencializa a citação das pesquisas, conforme Baynes (2017).

Os discursos apresentam novos aspectos de estímulo para a Persona Ecologia. Entre eles a ampliação da relevância do dado ecológico quando percebido em um contexto macro, uma vez que o compartilhamento possibilita a união de diferentes dados e a visualização mais completa da

informação. Esse movimento promove o avanço da ciência, uma vez que permite novas perspectivas, análises e conclusões. Junto a isso a otimização de tempo, custos e resultados.

São dados que, são poucas informações, ou poucos dados que não gerariam um artigo científico para uma pesquisa ou que também não seriam interessantes num repositório [...] você monta uma rede de pesquisa para uma determinada finalidade. Esses dados, eles vão ganhando relevância dentro do contexto [...] O que era uma informação muito importante de um fragmento, mas que era limitado daquele fragmento virou uma informação de todo o contexto [...] estando disponível, ele dá um ganho de escala para produção brasileira, para produção científica propriamente dita [...] A gente é extremamente favorável a isso e tem contribuído com essas redes e bases de dados, para que haja um avanço maior ainda no conhecimento [...] Quanto mais disponível a gente colocar essas informações para a comunidade como um todo, para a sociedade como um todo, mais vigilantes a gente tem, para o bem e para o mal [...] a pesquisa de qualidade, ela tende a sobrepor os riscos que são inerentes ao mau uso de dados. Eu acho que a gente só tem a ganhar com isso. (Entrevistado N° 2).

Está crescendo [...] retratações de artigos [...] O que é uma retratação? O artigo foi publicado, depois de um tempo alguém encontra ou um erro involuntário num código de análise, ou numa parte do dado. Depois de corrigir esse erro o resultado do artigo que foi publicado muda. Então, a revista vai lá e retrata. Retira aquele artigo e diz “olha esse artigo não tem mais essas conclusões. Elas não são mais válidas (Entrevistado N° 5).

Eles tinham que fazer uma pesquisa local de um problema específico. Eles tinham que correr atrás das coisas, não dava para ir a campo passar dois anos pesquisando, eles tinham que pegar metadados, eventualmente, ir a campo fazer uma verificação ou outra para gerar o artigo. [...] achar uma base de dados, fazer a verificação desses dados, o *checklist*, a análise para ver o que é consistente ou não [...] Se a gente não estiver aberto para visões diferentes, vamos ficar na ecologia do Século 19, do Século 20. Ela não cabe [...] Foi engraçado porque foi um trabalho que a gente fez em 15 dias, com uma ida a campo só para verificar se os nossos mapinhas estavam batendo, se nossos padrões de fato eram aqueles. Foi o trabalho mais fácil e mais rápido que eu já fiz em ecologia e só custou meio tanque de gasolina. Muito barato! [...] Até a década de 60 as pessoas iam para o cerrado do Brasil estudá-lo na época da seca, porque dava para chegar em algum lugar. As conclusões que eles chegaram sobre aquele bioma, que não era considerado bioma na época, é que ele era igual ou muito similar a caatinga brasileira. Ele era xeromórfico porque era seco. A partir da década de 80, se descobriu que na verdade tinha muita água no solo do cerrado. As folhas caíam, não por conta da falta de água e sim por outro motivo que ninguém sabia qual era. Essa resposta só apareceu na década seguinte e foi conseqüentemente ratificada, era uma questão mais evolutiva do que necessariamente climática (Entrevistado N° 7).

Os relatos também mostram que o fato de se ter financiamento público para as pesquisas, motiva a prática do compartilhamento. Junto a isso, disponibilizar se trata de uma responsabilidade fundamental da PE.

[...] eu não consigo imaginar como alguém é dono de alguma coisa assim. Você agora é dono de uma espécie? Tipo, a palmeira é minha. Você vai ter um outro tipo de palmito [...] é um bom uso do dinheiro, que geralmente é público. A gente gera o conjunto de dados com financiamento público [...] deixar ele disponível é o mínimo. Porque você

recebeu esse dinheiro para fazer isso, aqui está o dado, pelo menos. Quem quiser usar, está disponível [...] a ciência é um processo criativo, se a gente tiver um pouco de criatividade, a gente pode olhar o mesmo conjunto de dados com diferentes lentes, com diferentes olhares. Isso também é um bom jeito de usar o dinheiro público porque você está usando novamente um conjunto de dados que foi gerado com financiamento (ENTREVISTADO N° 5).

Se você recebe financiamento público, os dados não são seus. São do povo. Isso é difícil em qualquer área do conhecimento. Dizer, o seu salário é pago pela União, pelo Estado ou pelo Município e tudo o que você faz não é seu (ENTREVISTADO N° 7).

É possível interpretar que o movimento oportuniza a reflexão sobre as atividades, processos, avaliações e produções científicas indo ao encontro da percepção sobre ciência aberta da FIOCRUZ (2022). Percebe-se também, que a disponibilização dos dados possibilita a ampliação da rede, das oportunidades de pesquisa e da divulgação das informações, promovendo a translação do conhecimento, a ampliação dos impactos sociais e econômicos da ciência e a responsabilidade social científica (SANTOS *et al.*, 2017).

A disponibilização dos dados de pesquisa também é incentivada pelas revistas que o PE tem interesse em publicar. Elas não definem em qual meio, mas por vezes sugerem plataformas.

Revistas de alto impacto, eles exigem. Não só que você deposite os dados como você coloque os códigos de análise que você fez para referência, quem quiser repetir a análise [...] de livre escolha (Entrevista N° 2).

Geralmente, as revistas [...] que eu submeto os artigos, ou para *Ecology*, ou... Enfim, essas revistas elas, geralmente pedem os códigos que a gente utiliza, um código de programação esteja em alguma base aberta [...] as revistas, geralmente pedem para que a gente disponibilize, tanto os *scripts* que a gente usou para fazer as análises, as estatísticas, quanto um repositório mesmo de dados *on-line* para a gente disponibilizar que dados a gente utilizou. [...] Quando a gente submeteu o artigo para *Ecology*, eles estavam exigindo que a gente colocasse no *GitHub* e gerasse um *DOI* pelo Zenodo [...] têm revistas inclusive que têm o próprio repositório de dados para inclusão dos dados (Entrevista N° 4).

Hoje, as revistas, principalmente em ecologia, é praticamente impossível você publicar um artigo sem fornecer os dados [...] a revista fala “olha a gente só pública com os dados abertos”. [...] As revistas estão trabalhando para que não só os dados sejam compartilhados como os códigos de análise sejam compartilhados. [...] Boa parte das revistas, hoje, está pedindo os dados e esse roteiro de análise, de modo que fique disponível para quem quiser refazer as suas análises. Para saber se tem um erro e se aquilo que você está mostrando no seu artigo pode realmente ser reproduzido com aquele dado e com aquele roteiro [...] não conheço nenhuma revista que exige um repositório específico [...] Eu conheço revistas que sugerem alguns repositórios (Entrevista N° 5).

Assim sendo, confirma-se o citado por Vanz *et al.* (2018) quando aponta que o compartilhamento ocorria por meio de repositórios de revista/editor. Diante do exposto, é possível inferir a influência desse agente na promoção do compartilhamento e do reuso de dados de pesquisa ecológicos.

4.1.8 Medos/Insatisfações PE

Um dos aspectos identificado também nessa subcategoria foi a não padronização dos dados ecológicos e como essa influencia negativamente no compartilhamento e reuso da PE.

Essa é uma parte bem crítica, de você disponibilizar os dados depois que eles são coletados. Se eles não estiverem bem estruturados para passar para outra pessoa, isso pode ser perdido [...] falta de estrutura do próprio dado, quanto das estruturas dos metadados. Se tudo isso não estiver organizado, isso pode se perder [...] Se eu, eventualmente, passasse isso para outra pessoa, talvez isso se perca ao longo do tempo [...] Se todas essas informações não estiverem bem descritas nos metadados, os pesquisadores ou as pessoas da área de consultoria que podem utilizar esses dados não vão ter clareza de como esses dados foram reunidos e muito menos de como esses dados foram calculados (Entrevistado (a) N° 4).

Esses projetos estavam gerando uma quantidade grande de dados. Só que esses dados estavam sendo armazenados de jeitos diferentes, em formatos diferentes (Entrevistado (a) N° 5).

Quando você junta americanos, europeus, asiáticos, russos para trabalhar com a América do Sul, África ou Oceania existe um problema básico quando você vai compartilhar dados que é a padronização dos dados. [...] Se você for pegar a base de dados do Jardim Botânico do Rio de Janeiro e comparar com a base de dados do INPA você tem indivíduos e dados armazenados lá da mesma região, mas quando você vai juntar as informações para trabalhar você, às vezes, tem muita dificuldade porque falta informação que é importante. [...] Quando eu pego um trabalho acadêmico em ecologia, ou um trabalho técnico de dados ecológicos e comparo com outro trabalho, se eu não tiver a informação de como os dados foram coletados, o risco de você errar na comparação é grande. Porque a precisão pode ser diferente, porque a forma como você coletou é diferente. [...] Se você entrar no *speciesLink* hoje e pedir para verificar a distribuição de todos os cágados no Brasil. Você vai ver cágados há 100 km dentro do Oceano Atlântico na margem norte do nordeste brasileiro. A pessoa que fez aquele grande esforço de coleta não prestou atenção em duas variáveis importantes quando se usa o GPS, *datum* e aferimento se o GPS estava dentro ou não do local onde estava marcado (Entrevistado (a) N° 7).

Os relatos vão ao encontro do afirmado por Drucker (2011), os diferentes tipos de dados ecológicos tornam sua compreensão complexa, principalmente se reutilizados por outra pessoa, que não participou do levantamento original. Diferente do que menciona Peters *et al.* (2014), a PE possui a cultura de compartilhamento de dados e de certa forma está preparada para lidar com

questões de escala regional a continental devido à rotina de seu dia a dia de pesquisa. Porém, confirma-se que os métodos de coleta não padronizados e a desatenção para documentar a proveniência é fator relevante no processo. A partir da fala dos entrevistados, é possível reiterar a ideia de Michener (2015) de que os pesquisadores não fornecem metadados suficientes para apoiar a interpretação e a reutilização dos dados, indicando que a padronização dos dados ecológicos ainda é um desafio para o avanço de uma ciência mais aberta na área.

A PE apresenta algumas insatisfações em relação às infraestruturas e serviços atuais para compartilhamento e reuso.

Dependendo da plataforma, se você entra com uma diferença pequena no horário, muda totalmente a características dos dados. Então, isso pode influenciar em quem vai baixar essa informação para analisar (Entrevistado (a) N° 2).

As bases não são tão estáveis, assim, para a gente ter uma confiança sem ter uma base física [...] O que falta talvez nessa organização, disponibilização dos dados seja a questão dos bancos de dados relacionais mais estruturados, como o SQL. Acho que é isso ainda que falta para o metadado ficar bem completo. [...] Muitos dados que eu não tive acesso. O artigo foi publicado e a pessoa tinha uma tabela específica e eu não consegui acessar os dados. Aquilo virou um NA, não estava disponível (Entrevistado (a) N° 4).

Já vi locais que fornecem os dados e se você quiser usar você precisa vincular uma instituição. Não é que tem que ser a instituição que faz o cadastro, mas você tem que estar vinculado a uma instituição. [...] A gente gera muito dado repetido. Temos muitas iniciativas isoladas que são parecidas. Elas têm um poder menor de gerar síntese e ao compartilhar os dados – essas iniciativas de síntese, tipo os *data papers* e os repositórios – permitem que a ciência avance de um jeito que isoladamente demora muito mais (Entrevistado (a) N° 5).

Servidores com *backups*, com redundância, em lugares diferentes [...] na padronização tem um problema grave que é cada coluna que eu colocar a mais gera uma quantidade x de *bytes*. [...] Meu laboratório não tem uma plataforma, um repositório onde eu coloque e disponibilize de uma forma mais controlada para quem quiser acessar (Entrevistado (a) N° 7).

Barros (2005) afirmou que até aquele momento não existiam sistemas adequados para coletar, compartilhar e analisar dados ecológicos. Essa realidade é confirmada nos dias atuais a partir dos depoimentos. Pode-se interpretar que apesar de os ecologistas produzirem uma quantidade considerável de dados, em escalas mais amplas, a partir de redes colaborativas que promovem mecanismos de compartilhamento e reuso, ainda não existe a cultura de agregação desses, os sistemas disponíveis se mantêm limitados e não integram os dados. A infraestrutura cibernética existente ainda apresenta escassez de sistemas de gerenciamento de metadados fáceis

de usar, gerando dificuldades significativas na avaliação da qualidade e proveniência dos dados. Isso pode ser motivado pela capacidade reduzida de financiamento para curadoria e compartilhamento de dados e pela falta de articulação não somente dos sistemas existentes, como também de pessoas e instituições envolvidas no contexto. Infere-se que as instituições promotoras e mantenedoras desses parecem não estabelecer uma relação articulada entre si. Isso pode implicar diretamente no processo do acesso aberto, uma vez que os dados ainda não se conectam de modo eficiente. Desse modo, compreende-se que as ações que promovem o compartilhamento de dados ecológicos são variadas. Porém, essas práticas ainda evoluem lentamente (HEIDORN, 2008; HAMPTON *et al.*, 2013; MICHENER, 2015). Assim sendo, é possível compreender que ainda permanece a necessidade de uma infraestrutura computacional aberta que contemple as particularidades de cada fonte e de cada dado; e sistemas de informação que possibilitem a busca e a interpretação para facilitar as atividades dos pesquisadores da ecologia.

A falta de recursos financeiros para a prática da pesquisa e do acesso aberto também se confirma como causa de descontentamento na PE.

O corte da verba da ciência, isso vai impactar diretamente a questão das pesquisas. Tanto na formação dos próximos cientistas, quem está fazendo mestrado e doutorado, quanto quem já é cientista e precisa de projetos para ir para campo quando a gente pensa em ecologia, que talvez seja o principal gasto (Entrevistado (a) N° 4).

No Brasil, a gente não vai ter financiamento e tempo para fazer coleta em várias regiões do mundo. [...] Oficializar uma parceria depende de financiamento (Entrevistado (a) N° 5).

O Brasil nunca se importou em investir em produção de tecnologia para ser nacional [...] O Brasil não oferece as condições adequadas para que o professor universitário publique rápido. [...] Os editais do Brasil não permitem que eu coloque um contrato de uma pessoa para tomar conta do meu projeto. Recebemos R\$ 400.000 e eu não tenho nenhum centavo para dizer assim: contrato fulana para me ajudar a gerenciar e pensar as coisas que nós não podemos pensar. Eu tenho que parar de fazer o que estou fazendo para poder fazer as coisas acontecerem (Entrevistado (a) N° 7).

Conforme os dados do GT – RDP Brasil (2018) e Vanz *et al.* (2018) as agências de financiamento devem oferecer serviços de apoio. A PE também apresenta ressalvas em relação à segurança e à disponibilização dos dados ao longo do tempo, principalmente em repositórios que não deixam transparente por quem são gerenciados e onde suas bases estão localizadas.

[...] ter uma base física de controle, de manutenção e de acesso nacional é fundamental, principalmente pela segurança [...] Deixar essa base de dados em mãos internacionais,

talvez não seja a melhor opção porque a gente não tem controle, principalmente Google ou Amazon. [...] Servidores que a gente não sabe onde estão e não sabe até quando vão. [...] Meu receio é ter a manutenção e a posse desses dados em território nacional para conseguir organizar, manipular e disponibilizar os dados de forma mais segura (Entrevistado (a) N°4).

O único receio que eu tenho é que os repositórios não pertençam a instituições privadas que depois vão querer comercializar o dado. É essa a minha única preocupação (Entrevistado (a) N°5).

Um projeto assinado pelos franceses, que quando eles coletaram os dados – não era permitido via satélite, era coletado em campo – os caras foram para a França sem deixar uma cópia aqui. Era sobre a bacia amazônica. O Brasil levou cinco longos anos para repatriar os dados. Isso é um problema de segurança nacional. É um problema de segurança nacional (Entrevistado (a) N°7).

Essa investigação não apontou a desconfiança da PE em relação à possibilidade de suas ideias serem copiadas ao compartilhar os dados, conforme considerado por Vanz *et al.* (2018).

4.1.9 Expectativas/Necessidades PE

Essa subcategoria foi articulada tendo como base os diversos aspectos que influenciam no contexto do acesso aberto da Persona Ecologia. É possível verificar que ela possui expectativas e necessidades em relação aos serviços. Os depoimentos abaixo demonstram o interesse dele em gerar impacto para a sociedade; ampliar a observação das pesquisas e a relevância da ciência; acelerar a comunicação do conhecimento para tomada de decisões; ter uma disponibilização que garanta o reuso por meio da padronização dos dados e uso de metadados, e; ter uma iniciativa nacional para o acesso aberto.

A gente trabalha muito com planejamento para a conservação, tomada de decisão e políticas públicas. Para nós aceleraria também o conhecimento para a melhor tomada de decisão [...] elas tem que influenciar cada vez mais que os pesquisadores compartilhem suas informações, esses dados em plataformas abertas para que haja grande escala na produção de qualidade das informações, dos artigos científicos e que vão gerar mais impacto para sociedade [...] de forma geral eu espero isso, que essas plataformas contribuam para gerar esse tipo de estímulo [...] para cada quantidade de dados tem, sei lá, dez mil perguntas que a gente pode fazer e tentar responder [...] Um pesquisador sozinho, possivelmente, não vai exaurir tudo que pode ser feito com aquela informação. Então, basicamente acho que é esse tipo de serviço (Entrevistado (a) N° 2).

[...] saber que existe essa iniciativa e que os dados estão lá e que eu posso ter acesso, isso seria um grande avanço na questão científica (Entrevistado (a) N° 4).

[...] fazer sínteses mais amplas, tempos mais longos, grupos mais diversos (Entrevistado N° 5).

A gente sonha em construir a coleção Zoológica [...] uma coleção de dados ecológicos puramente para quem vai para campo, mede, coleta e não traz nada para o laboratório (Entrevistado (a) N° 7).

Ele considera ter suporte e recursos (humanos e de treinamento/capacitação/orientação) para melhoria do processo de compartilhamento e reuso.

Um profissional para fazer isso eu acho que desocupa um pouco os profissionais que estão fazendo as análises de dados, que estão indo pro campo [...] ele que vai fazer o *upload* a gente tem que ter o cuidado de que isso seja bem triado, seja bem feito para que não gere problema, se houver um canal de comunicação [...] um treinamento básico, de um dia, sei lá, de orientação em relação a isso e porta aberta (Entrevistado (a) N° 2).

Talvez precise de alguém que direcione um formato de dados mais sistematizado, ou mais organizado para a gente utilizar esses dados daqui para frente [...] Idealmente, quando a gente trabalha com banco de dados seria alguém da Ciência da Computação [...] seria interessante essa questão, talvez ou treinamento ou algum direcionamento desse conhecimento de gerenciamento de dados. (Entrevistado (a) N° 4).

À vista disso, a Persona Ecologia tem algumas expectativas em relação à Persona Suporte e como esse segundo perfil de usuário deve ser. Os trechos abaixo demonstram essa perspectiva:

Um técnico de nível médio, com boa fluência em computadores, não precisaria ser ninguém de nível superior. Algum técnico de nível médio resolveria. [...] Pode ser de outra instituição, desde que haja um canal de comunicação facilitado. (Entrevistado (a) N° 2).

Idealmente, quando a gente trabalha com banco de dados seria alguém da Ciência da Computação [...] processos [...] simples, [...] planilha de Excel, que não está nada estruturado, acredito que mesmo alguém da graduação com um certo conhecimento de planilha, um certo conhecimento de controle de versão, que talvez seja um pouquinho importante nessa parte do *GitHub*, um pouquinho de linguagem de programação para conseguir manejar as tabelas com um pouco mais de facilidade e conseguiria fazer o upload dos dados ou mesmo uma publicação [...] o inglês é fundamental para essas bases, principalmente porque tudo está internacional. E em algum momento você vai se deparar ou uma publicação que vai estar em inglês, ou algum processo que vai requerer um pouco de conhecimento em inglês. [...] possibilidade de um canal aberto de comunicação entre os dois seja fundamental ele manter essa comunicação com os pesquisadores, que seja um e-mail ou um chatzinho "olha, essa coluna aqui ela está estranha, você pode verificar?" [...] a comunicação é fundamental. Ela não precisa ser física, mas ter essa comunicação com quem organiza os dados com quem disponibiliza os dados é importante (Entrevistado (a) N° 4).

O desafio para quem trabalha com sistema de informação voltado para a ecologia ou qualquer outra coisa. O que era fundamental e importante há 100 anos, hoje em dia é

insuficiente. O que vai ser importante daqui a 100 anos, provavelmente vai ser insuficiente o que a gente coleta hoje em dia (Entrevistado (a) N° 7).

Ele tem a expectativa da melhoria das infraestruturas e processos existentes para aperfeiçoar a prática.

[...] pode ter um processo manual ou um processo automatizado para limpeza desses dados nos campos que a gente quer lá no final [...] bancos de dados estruturais, por exemplo, uma coisa mais organizada [...] um grande servidor que reúna essas informações e deixe disponível num site, que seja fácil de acessar, seja fácil manipular, seja fácil corrigir, seja fácil implementar. [...] essa base em que os pesquisadores já tenham, quando começam a pesquisar já tenham essas informações, os campos que eles vão ter que preencher no futuro (Entrevistado (a) N° 4).

Um bom compartilhamento de dados é o dado original mais uma boa descrição de como aquele dado foi coletado e como ele deve ser entendido. [...] o conjunto de dados tem que estar tão bem organizado, que uma pessoa que não conhece como ele foi gerado, somente ao analisar a planilha, seja capaz de entender o delineamento amostral que gerou aquele conjunto de dados (Entrevistado (a) N° 5).

O desafio para quem está trabalhando com padronização é qual o limite mínimo que atenda a maior quantidade possível de áreas? E, como eu faço esse sistema? Seja ele repositório, seja ele um sistema de informação propriamente dito que vai gerar outras informações a partir dos dados, dos metadados, ou seja, lá o que for para tomadores de decisão. [...] O fato é qual o mínimo eu tenho que usar? Dentro desse mínimo que cabe todas as áreas do conhecimento da ecologia ou da ciência de um modo geral que pode usar esse tipo de dado. [...] As exigências das bases de dados também precisam melhorar [...] Essas bases de dados têm esse papel, como eu conversei com atores diferentes para que todos eles possam – até ter interpretações diferentes sobre meus dados, mas não pode ser díspar, não pode ser antagonica. Não dá para dizer que $2+2 = 4$ e o outro achar que é 7. Isso não funciona. [...] (Entrevistado (a) N° 7)

A Persona Ecologia espera uma infraestrutura nacional para o compartilhamento e reuso de dados abertos de pesquisa.

Uma iniciativa dessas certamente a gente faria depósito nessa base de dados [...] Mesmo que a gente tenha um repositório nosso, pode ser aberto (Entrevistado (a) N° 2).

Seria fantástico [...] Uma iniciativa nacional que reúna os dados, principalmente com o método que sistematize essas informações para disponibilidade e avanço científico eu acho bastante fundamental [...] Tanto para reunir o que a gente já tem de dados, quanto para organizar ele e tanto para novas pesquisas que podem utilizar esses dados [...] é fundamental para o avanço científico [...] tendo uma iniciativa nacional que reúna todos os dados é interessante alguma coisa que seja alojado ou no Ministério das Comunicações, ou no Ministério da Educação ou no Ministério, bom, seria o ideal no ministério da Ciência e da Tecnologia [...] Centrar isso ou em Brasília, ou talvez no INPI (Instituto Nacional da Propriedade Industrial) talvez, que já tem uma grande experiência em juntar dados e de trabalhar com isso [...] A necessidade de ter uma base física “de controle”, de manutenção e de acesso nacional ela é fundamental para disponibilizar

esses dados, principalmente pela segurança de a gente ter o dado na mão e poder manejar ele, quanto da manutenção ao longo do tempo. [...] ter um servidor que reúna de fato os dados (Entrevistado (a) N° 4).

Os relatos vão ao encontro do mencionado por Anjos (2016) e Vanz *et al.* (2018) sobre os motivos do não compartilhamento. Os autores apontaram que a falta de infraestrutura, de padronização dos dados, a não disponibilização de pessoal especializado para preparar os dados para disponibilização e a falta de orientações específicas sobre o processo influenciam na prática. Essa investigação confirma que essas considerações se mantêm até o momento. Logo, é possível inferir que a PE tem a expectativa de repositórios digitais desenvolvidos a partir de base de dados relacional, que contemplem dados integrados em sistema de informação estruturados que permitem diferentes solicitações por parte do usuário (CALDEIRA, 2009).

Essa pesquisa também identificou que a Persona Ecologia não tem o costume de utilizar uma linguagem própria para descrição dos dados disponibilizados, conforme citado na subcategoria Competência. Nessa perspectiva, a PE entende que algumas informações são fundamentais para a padronização, a fim de garantir o reuso e apresenta algumas expectativas/necessidades em relação a isso.

Se tivesse em campos fechadinhos para as pessoas irem preenchendo, isso talvez pudesse ser um avanço, bastante, para tanto checagem dos dados quanto disponibilização correta dos dados (Entrevistado (a) N° 4).

Um dado super bem organizadinho. Cada linha é uma informação única, as colunas estão bem descritas, tem outro arquivo de metadados (Entrevistado (a) N° 5).

Capacidade de verificação de onde se coletou (data, hora, coordenadas, latitude, longitude, *datum*) [...] colocar um metadado que é composto por dois blocos de informações. A primeira é o que significa cada coisa que você colocou, cada acrônimo significa o quê? A segunda coisa é para cada variável dessas qual é a precisão e qual é a acurácia da medida [...] Existem três tipos de padrões que são importantes para a gente entender. O primeiro é que tipo de dado que eu coloco ali. O segundo é como eu organizo e gerencio esses dados. O terceiro, como eles conversam [...] Quando a gente for pensar num sistema nacional de padronização [...] dizer qual equipamento, qual fabricante, número de série, qual foi a primeira calibração e a última calibração. [...] Ter uma base secundária de equipamentos, mostrando a história dele é fundamental, porque não é raro você ter descalibração. Ou pior, ele quebra e você não percebe [...] O metadado tem que vir com todas essas questões: unidade de medida, precisão, acurácia, limite mínimo e máximo esperado, variação, para que as pessoas possam tomar a decisão depois que voltam do campo para mesa e começam a trabalhar (Entrevistado (a) N° 7).

Os relatos demonstram a consciência da PE sobre a importância dos metadados para facilitar o gerenciamento e a compreensão das investigações ecológicas, conforme citado por Michener (2006). Os excertos também ampliam o conhecimento gerado por Anjos (2016) em relação às quais variáveis devem ser consideradas no momento da disponibilização. A autora apontou alguns, entre eles: data, localidade geográfica, dados climáticos e equipamentos utilizados para o registro. Essa subcategoria demonstra que o contexto permanece com alguns desafios para a prática efetiva do acesso aberto pela Persona Ecologia.

Na sequência, a discussão apresenta a Persona Suporte e suas nuances, segundo perfil de usuário identificado a partir da análise dos dados coletados nas entrevistas.

4.2 PERSONA SUPORTE (PS)

Antes de apresentar as inferências sobre a PS, observa-se que o arquivo de respostas do questionário aplicado pelo GT – RDP Brasil (2018) apontou alguns indícios sobre quem deveria oferecer serviços de apoio a gestão dos dados de pesquisa (questão 3.4 da *survey*). Entre os apontados apareceram instituição de atuação do pesquisador, agência de financiamento e ente governamental. Ademais, as agências de financiamento (regional e nacional) foram citadas como agente que exige a disponibilização dos dados de pesquisa, tanto na *survey*, quanto nas entrevistas dessa tese, configurando-se como um dos motivadores do processo.

Quem começa a iniciativa de colocar os dados são as agências de fomento, são os artigos, os periódicos. [...] É importante ver se elas mudam de comportamento, por exemplo, para de exigir ou começam a disponibilizar os próprios repositórios ao invés de exigir que a pessoa encontre um. Isso pode mudar a forma de desenvolvimento do ambiente em que os dados vão ficar disponíveis. Se as agências mudarem seus comportamentos, isso pode mudar esse crescimento, pode acontecer de outra forma (Entrevistado (a) N° 1).

Nessa perspectiva se observa também que as interpretações e inferências sobre a Persona Ecologia possibilitaram perceber a importância das redes colaborativas na prática da pesquisa e da ciência aberta no contexto da ecologia, uma vez que boa parte dos mecanismos utilizados para compartilhamento e reuso de dados são promovidos por meio de parcerias entre diversos atores. Desse modo, essas redes são também entendidas como agentes que oferecem suporte no decorrer do processo.

Porém, os depoimentos e a AC evidenciaram outro perfil que influencia na prática de compartilhamento e reuso de dados, entendido agora como Persona Suporte. Esse arquétipo não se configura como uma organização, instituição, ente governamental ou rede colaborativa, outros aspectos o caracterizaram de modo diferente, ampliando o conhecimento prévio dessa investigação. A PS se configura como um (a) profissional, indivíduo que auxilia na gestão dos dados e atua ativamente, facilitando o processo.

Ciência da informação, biblioteconomia, arquivologia, TI só seria necessário mais, talvez, para execução de experimentos mais complexos. O bibliotecário vai conseguir ajudar mais no preenchimento dos metadados. Não falo bibliotecário, uma pessoa que entenda de informação. [...] Eu também acredito, por exemplo, a cada área essa pessoa da informação vai ter que ter um conhecimento avançado naquela área [...] Por exemplo, agora a gente está trabalhando com dados geoespaciais, eles têm um padrão brasileiro, que tem que ser atendidos, certos metadados tem que ser preenchidos e ai? Aí, uma pessoa da matemática nunca ouviu falar desse padrão, talvez, né. Então, hoje a gente tem as bibliotecas setoriais e acho que seria o caso de ter uma pessoa da informação, um bibliotecário especializado naquela área. [...] Vai ser uma pessoa da área da informação, talvez uma pessoa de TI ajudando 10, 15 pesquisas (Entrevistado (a) N° 1).

Existe uma nova profissão sendo discutida que é do *data steward* [...] O *data steward*, tem gente que chama de administrador de dado, tem gente que chama de custódia de dados [...] Que auxilia na gestão de dados não só durante o projeto, mas antes. Dinamarca, Holanda, Reino Unido tem gente que é contratada como *data steward* nas instituições, não só nas de pesquisa nas universidades, como algumas instituições de pesquisa privada. [...] Dividindo em *data steward* que é mais administrativo, o *data steward* que tem aquele perfil mais de *data science*. Que é o cara que vai fazer mais a parte de rotina em R²⁷, essas coisas e vai ajudar nas análises e na parte de controle de qualidade. Tem o *data steward* mais técnico, tem o mais administrativo, que ele que vai orientar qual a licença de dados a utilizar [...] O cara que é mais da TI, mas também faz a interface com o pesquisador, que vai fazer administração do repositório, vai botar os serviços em pé (Entrevistado (a) N° 3).

Como é possível perceber nos trechos acima, a PS pode ter ainda uma variedade de perfis, os quais podem abranger desde um com caráter técnico até outro especializado. A partir desse contexto, a AC conseguiu delinear um tipo principal, que se destacou nas entrevistas. Apesar de esse arquétipo ter sido construído a partir de dados de indivíduos reais, considera-se que ele ainda demanda aprofundamento. A Persona Suporte é entendida como casualidade fortuita, uma

²⁷ A criação de um pacote do R pode ter duas motivações principais:

- Agrupar e manter uma série de funções criadas por você e que são basicamente de uso próprio ou rotineiro.
- Compartilhar funções de novos métodos e/ou implementações que você tenha criado e queira disponibilizar para outras pessoas usarem.

descoberta dessa tese. Esse perfil não foi apontado pelo quadro teórico, ele surgiu nos depoimentos que buscaram responder o problema e os objetivos dessa investigação. Por isso, ao percebê-lo considerou-se importante sua observação.

A seguir, informações sobre como essa Persona Suporte é caracterizada.

4.2.1 Perfil PS

Esta subseção apresenta aspectos que desenham o arquétipo da Persona Suporte, suas habilidades, sua rotina de trabalho e sua participação no processo de acesso aberto aos dados de pesquisa ecológicos. A PS pode pertencer ao gênero feminino ou masculino, orientação heterossexual, branca, entre 40 e 60 anos de idade, com renda mensal de aproximadamente R\$ 17.000,00, possui filhos, é casada e pode residir e trabalhar em diferentes regiões do Brasil. Ela é pesquisadora, possui formação na área da ecologia, experiência de campo e de pesquisa com dados ecológicos, demonstrando competências que corroboram com o processo de disponibilização de dados. Os excertos e inferências apresentadas na sequência demonstram essa realidade.

Sou formada em engenharia florestal, eu fiz mestrado em ecologia e fiz doutorado em ambiente e sociedade. [...] eu já tive muito mais essa rotina de ser uma pesquisadora de campo. De ir para campo, de planejar, de fazer campanha de campo, de anotar. [...] Conduzi pesquisas na região Amazônica, na região da Mata Atlântica, na região de cerrado, aqui no Brasil e também em algumas localidades da Califórnia. [...] trabalhei mais especificamente na área de conservação da biodiversidade [...] biodiversidade e serviços ecossistêmicos, que são típicas da área da ecologia [...] Tendo estado do outro lado, de quem foi para campo e fez todo esse esforço. Eu consigo entender melhor com quem eu estou dialogando e quando surgem resistências ou questionamentos eu enxergo o ponto de vista do autor do dado muito claramente [...] isso facilita a conversa, o diálogo. Mostrar o outro lado, o quanto é importante, os benefícios que a pessoa vai ter em expor esse dado. Não ter receio porque, na verdade, ela vai justamente valorizar cada vez mais o trabalho dela. Então, eu acho que ajuda muito eu ter tido essa vivência (Entrevistado (a) N° 3).

Possuo doutorado em ecologia e evolução, trabalho com consultoria em planejamento para conservação. [...] Eu corria campo, coletava fragmentos florestais, trabalhava com plantas [...] Todos são ecólogos, têm biólogos também [...] Todos têm conhecimento sobre linguagem de programação R (Entrevistado (a) N° 6).

Possuo doutorado em ecologia por uma instituição de ensino superior pública brasileira e sou voluntário do Museu da Amazônia (Entrevistado (a) N° 7).

Diferente da PE, a Persona Suporte não atua necessariamente como docente universitária ou de pós-graduação no seu dia a dia. Sua rotina envolve uma série de atividades que propiciam o debate, a formação e a articulação para o acesso aberto aos dados de pesquisa da área.

De certa forma participo de vistas internacionais relacionadas aos repositórios de dados e estou a par de como está a pesquisa nessa área hoje em dia. E aí como a internet é internacional a gente participa, mas não é tanto uma participação, é mais uma observação de vários grupos internacionais voltados a isso. (Entrevistado (a) N° 1).

Participo de resumos e atuo em pesquisas ecológicas propriamente ditas e aí sim eu faço busca em repositórios e em sistemas de informação, mas eu dedico muito mais tempo do meu dia a dia para auxiliar os pesquisadores a fazer esse trabalho, de ficar achando as fontes de informação ou então de publicar o dado que ele gera, ajudar a fazer plano de gestão de dados. (Entrevistado (a) N° 3).

A gente teve que estudar. [...] Até nos sites das bases de dados. Nos próprios Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR), *Global Biodiversity Information Facility (GBIF)* e *speciesLink* já tem algumas ferramentas *on-line*, tem alguns cursos. Eu fui fazendo cursos *on-line* que tinham na plataforma. No próprio *GBIF* tinham cursos *on-line* (Entrevistado (a) N° 6).

Hoje, ela atua internamente em sua instituição pública na gestão de dados, como analista de sistemas ou de informática. Sua função é auxiliar outros ecólogos no processo de compartilhamento e promover iniciativas, na perspectiva de garantir o reuso. Os fragmentos de texto abaixo solidificam essa inferência:

Atuo no centro de processamento de dados [...] estou participando de um projeto para implantar um repositório de dados de pesquisa institucional [...] a intenção é que ele seja de acesso aberto [...] querendo garantir o reuso de outros, mas não sentar e usar aqueles dados (Entrevistado (a) N° 1).

Desde o mestrado que eu comecei a trabalhar na parte de gestão de dados [...] continuo fazendo pesquisa na área de biodiversidade e tenho publicações no assunto, mas eu atuo muito mais como gestora de informação atualmente. Apesar de não ser formada formalmente na área de informação, comecei a trabalhar com gestão de dados com dados ecológicos de biodiversidade e aí eu venho desenvolvendo também pesquisa nessa área de gestão de dados, principalmente, voltada a área de ecologia e de agricultura. Tenho muito treinamento nisso, eu fui fazer um curso lá nos Estados Unidos. Estou me especializando nesse assunto e eu comecei a trabalhar com a parte de gestão de dados há mais de 15 anos. Eu estou super atenta para o que está acontecendo no mundo nesse assunto [...] Eu comecei a trabalhar com gestão de dados foi em 2005, mas eu continuei fazendo campo até 2010. [...] No meu mestrado e no meu doutorado eu fiz bastante análise e agora eu estou voltando mais a minha atuação para essa parte que é o Pilar I (*Interoperability*) do *FAIR* (Entrevistado (a) N° 3).

Saí da graduação de uma pessoa que coletava e gerava dados para o mestrado e doutorado, para uma pessoa que consumia dados. Essa é uma etapa fundamental. [...] Trato os dados, gero ferramentas para avaliar a qualidade desses dados [...] Apesar de ser

novo nessa área, tenho começado a dar os meus primeiros passos. [...] Eu já tinha um conhecimento, já sabia como escrever, onde buscar, mas esse mundo de qualidade de dados ninguém conhecia. [...] A gente desenvolveu um pacote R para fazer um monte de coisas, incluindo limpeza de dados, juntar esses dados, limpar os dados (Entrevistado (a) N° 6).

Eu tive a honra de participar como aluno e como técnico, talvez da primeira iniciativa brasileira de acesso aberto aos dados (Entrevistado (a) N° 7).

A PS participa ativamente no processo de compartilhamento e reuso de dados da Persona Ecologia, reconhecendo a importância dessa última para a prática.

Preencher os metadados, ver se eles estão adequados, facilitar. Como ele tem conhecimento avançado naquele setor da ecologia, talvez ele já possa sugerir alguns preenchimentos para o professor, para o pesquisador. Claro, depois isso teria que retornar ao pesquisador para ele dar anuência. [...] O certo é o dono da informação documentar, outra pessoa analisar para ver se aquela informação é o suficiente. [...] É muito difícil descrever em que ambiente o pesquisador fez aquela experiência, que aparelho ele usou para coletar a informação. Ele é que sabe. [...] Não só disponibilizar o compartilhamento, garantir o reuso (Entrevistado (a) N° 1).

Muitas vezes, as pessoas me contatam para falar: “olha eu estou escrevendo esse projeto e a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) está exigindo que eu faça esse plano de gestão de dados e eu não sei nem como começar. Eu sempre falo: “eu posso te ajudar, mas eu não faço a menor ideia do que é o seu projeto. E só quem vai saber descrever os dados é você. Eu posso te ajudar a explicar o que a FAPESP quer com isso”. É até um trabalho educacional mesmo [...] Para os pesquisadores pensarem antes que dados que eles vão coletar, eles vão armazenar, como eles vão garantir que não vai ser perdido, o que vai poder ser disponibilizado, o que não vai ser. Não vai? Porque que não vai? Qual é a restrição legal que dá suporte para que não seja compartilhado. Se são várias instituições envolvidas, em que repositório vai ficar. Qual que é a exigência da instituição de fomento. Ajudo o pessoal a escrever e às vezes eu avalio, vejo se está bom ou se precisa melhorar alguma coisa. Uma vez que o projeto é aprovado, quando começam as campanhas de campo eu vou lá ver se o que está sendo feito no plano de gestão de dados precisa ser alterado. [...] o plano de gestão de dados não é uma ferramenta estática. Ele pode ser alterado, pode ser atualizado. [...] Se os pesquisadores estão realmente tomando cuidado para não perder esses dados, se eles estão armazenando em cópias, se não está só no computador deles, que eles podem perder ou serem roubados, [...] Ou no *pendrive*. Ajudo muito os pesquisadores a entrarem com os dados no repositório, eu faço manuais, eu tiro dúvidas, eu atribuo permissão. Eu recebo as informações, às vezes, o pesquisador coloca lá no modo rascunho. Aí, eu falo: “Olha é só tabela, está misturando número com palavra, não dá para ser assim. O que é essa célula amarela?”[...] A gente vem trabalhando nessa lógica [...] de o próprio pesquisador colocar o dado no repositório, só que ele não consegue publicar. Ele passa por uma curadoria antes de ser publicado para ver se precisa aprimorar alguma descrição no metadado. [...] Essa questão da tabela de dados estar inadequada, com caracteres esquisitos, que ninguém entende, com célula colorida ou com colunas que misturam dados numéricos com textuais, esse tipo de coisa, a gente orienta a corrigir, a colocar, as práticas de gestão de dados. Está funcionando bem esse bate-bola entre o pesquisador e esse *data steward* que vai auxiliar a adequar da melhor forma aquele dado para ele ser publicado (Entrevistado (a) N° 3).

A padronização dentro dos dados tem que ser feita pelo usuário. Existe até um programa que faz isso, se chama *OpenRefine*. [...] A gente segue boas práticas de manejo de dados de biodiversidade. [...] A gente não mexe no dado da pessoa, a gente não altera nada. Adicionamos uma coluna dizendo que aquele registro tem ou não algum problema. Cada usuário tem que filtrar. Aspecto fundamental, a limpeza e avaliação da qualidade dos dados não são feitas pela plataforma, quem tem que fazer isso é o usuário final (Entrevistado (a) N° 6).

Os depoimentos que caracterizam a rotina de trabalho da Persona Suporte vão, de certo modo, ao encontro da expectativa que a Persona Ecologia possui em relação a um (a) profissional de apoio. De fato, a PS se configura como facilitadora do processo, promovendo práticas que visam à colaboração e ao compartilhamento de dados e conteúdos científicos entre os pesquisadores e o público interessado (FIOCRUZ, 2022; CIDACS, 2022). No entanto, a Persona Suporte se desenhou mais especializada na realidade do que na visão da Persona Ecologia. Apesar dessa dissimilaridade, entende-se que a PS ser conhecedora da área agrega valor e qualidade ao processo. A partir disso, compreende-se que há confluência de objetivos da PS e da PE para uma prática efetiva da ciência aberta no campo da ecologia.

A PS também pode ser considerada um *stakeholder* conforme a ocasião e a estratégia das iniciativas, pois é composta por um grupo de indivíduos que pode afetar ou é afetado por essas. Assim sendo, podem ser vistos como oportunidades ou ameaças de acordo com o contexto (FREEMAN, 1984; FRIEDMAN; MILES, 2006). O arquétipo também se desenhou como um público de *decisão*, de *comportamento* e de *consulta* perante as iniciativas, uma vez que possui poder de determinar algo sobre a existência das mesmas; executa ações, a fim de cumprir com suas finalidades e que interfere de algum modo na tomada de decisão, pois suas opiniões são consideradas no processo (SIMÕES, 1995). Trata-se de um público *ativo*, formado por indivíduos que são afetados pelo problema, o reconhecem e fazem algo a respeito (GRUNIG; HUNT, 1984). Como também, um público de *voz*, já que busca ser ativo diante das práticas diárias, influenciando políticas e resultados (HIRSCHMAN, 1970). Esse perfil varia conforme potencial de poder que pode exercer, com variada probabilidade de estimular ou desestimular ações em prol do acesso aberto aos dados de pesquisa. Portanto, esse arquétipo deve ser considerado um público estratégico no contexto de iniciativas que visam o acesso aberto aos dados de pesquisa ecológicos.

Logo, discorre-se sobre por quais mecanismos de compartilhamento a PS utiliza em sua prática.

4.2.2 Prática: meios de compartilhamento PS

Como a Persona Ecologia, a Persona Suporte utiliza alguns mecanismos de compartilhamento ligados à área da ecologia para a prática do seu trabalho. “no Brasil tem várias fontes de dados que atendem esses requisitos [...] de sistema de informação, que não são necessariamente repositórios de dados” (Entrevistado (a) N° 3). “O que existe são várias plataformas *on-line*, onde eu posso baixar os dados. [...] São agregadoras de dados. Eles não são os donos dos dados, o usuário é o dono” (Entrevistado (a) N° 6).

Os meios nacionais utilizados pela PS são a Plataforma do Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) e o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR), também mencionadas pela Persona Ecologia.

Apareceram também, o Projeto de Mapeamento Anual do Uso e Cobertura da Terra no Brasil (MapBiomas), rede colaborativa criada em 2015, formada por organizações não-governamentais, universidades e *startups* de tecnologia. Uma plataforma aberta, multiplicável e com possibilidade de aplicação em outros países e contextos. Esse mecanismo é desenhado para incorporar e acolher as contribuições da comunidade científica e demais interessados em colaborar, seu processamento é distribuído e automatizado dos dados através de parceria com *Google Earth Engine* (MapBiomas, 2022). E o Repositório do Programa de Pesquisa em Biodiversidade (PPBio) citado por Pavão *et al.* (2018a). Criado em 2004, tem como objetivos intensificar estudos sobre biodiversidade no Brasil, descentralizar a produção científica dos centros desenvolvidos academicamente, integrar atividades de pesquisa e divulgar os resultados para diferentes finalidades, como gestão ambiental e educação (PPBio, 2022).

O repositório do Programa de Pesquisas Ecológicas de Longa Duração (PELD) foi citado com a seguinte observação (BARROS *et al.*, 2005; PAVÃO *et al.*, 2018a).

A gente criou um repositório para o PELD utilizando o *Metacat* que é o *software* que implementa a *Ecological Metadata Language (EML)*. Só que esse repositório não está mais disponível, mas os dados que estavam lá foram todos para o repositório do PPBio (Entrevistado (a) N° 3).

Porém, como a Persona Ecologia, nada foi mencionado pela Persona Suporte sobre a Biblioteca Digital Georreferenciada (BDiG-PELD) (BARROS, 2005). A partir disso, é possível interpretar que a iniciativa não é reconhecida tanto pela PS, quanto pela PE.

Os meios internacionais *Dryad*, *Figshare*, *GitHub*, *Zenodo* e *speciesLink* citados pela Persona Ecologia, também foram mencionados pela PS. “*Dryad* que é um repositório que está associado com várias revistas importantes na área [...] Tem bastante dado de boa qualidade” (Entrevistado (a) N° 3). Outros utilizados pela PS são:

Atlas of Living Australia (ALA) repositório *on-line* de informações sobre plantas, animais e fungos australianos desenvolvido desde 2006.

ALA é a plataforma das bases da Austrália, talvez seja o pessoal que mais avança com qualidade de dados. Eles têm agora o que eles chamam de perfis de usuários. Por exemplo, eu estou interessado em modelar a distribuição dessa espécie, eu quero saber onde ela está. Você tendo esse perfil já tem um conjunto de filtros pré-definidos que você pode usar e seus dados vão estar com uma qualidade melhor (Entrevistado (a) N° 6).

A plataforma *Global Biodiversity Information Facility (GBIF)* citada em Vanz *et al.* (2018) apareceu nas menções das entrevistas da PS.

O Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR) é o nó nacional do *Global Biodiversity Information Facility (GBIF)* [...] qualquer um pode mandar dados para eles e lá tem milhões de registros de ocorrências de espécies (Entrevistado (a) N° 3).

A primeira grande plataforma mundial foi a *Global Biodiversity Information Facility (GBIF)* [...] *GBIF*, por exemplo, tem a opção que você pode excluir registros sem nome, sem coordenadas, registros suspeitos e isso já está disponível. Por incrível que pareça, se eu não me engano isso foi disponibilizado no *GBIF* há um ano ou dois, no máximo. É uma coisa muito recente (Entrevistado (a) N° 6).

Knowledge Network for Biocomplexity (KNB) repositório internacional lançado em 1998 destinado a facilitar a pesquisa ecológica e ambiental. O *KNB* aceita todos os dados ambientais ou ecológicos e publica conjuntos de dados com identificadores de objetos digitais (*DOI*) com o objetivo de garantir o acesso em longo prazo a esses conjuntos de dados, visando cumprir os princípios *FAIR* de compartilhamento e preservação de dados. “*KNB* um repositório para dados ecológicos que qualquer pessoa pode se cadastrar e colocar os dados” (Entrevistado (a) N° 3).

E a *Research Data Alliance (RDA)* lançada como uma iniciativa comunitária em 2013 pela Comissão Europeia, pela Fundação Nacional de Ciência do Governo dos Estados Unidos e pelo Instituto Nacional de Padrões e Tecnologia e o Departamento de Inovação do Governo

Australiano com o objetivo de construir a infraestrutura social e técnica para permitir o compartilhamento aberto e a reutilização de dados.

Os depoimentos também revelaram iniciativas específicas de desenvolvimento de meios que buscam armazenar e compartilhar dados no Brasil por parte da PS.

Biodiversity Data Cleaning (BDC) [...] A gente desenvolveu um pacote de R para fazer um monte de coisas, incluindo limpeza de dados, juntar esses dados, limpar os dados. [...] O *workflow* assume que você já tem os dados de biodiversidade. Você já coletou no campo, ou você achou esses dados em alguma plataforma *on-line* [...] O *workflow* se tornou na verdade o pacote de R, que é um conjunto de ferramentas [...] R é uma forma de escrever códigos é basicamente uma linguagem de programação [...] a gente revisou todas as ferramentas que já estavam disponíveis. Incorporou essas ferramentas no nosso e desenvolveu um monte de outras (Entrevistado (a) N° 6).

Em coleções de microrganismos, nós terminamos um sistema agora para minha universidade que vai disponibilizar para rede *Dado anonimizado* inteira [...] tem níveis de informações diferentes. [...] A ideia é que ele seja com código fonte aberto porque não faz sentido ser fechado. E disponível para qualquer outra instituição que esteja a fim de adotar no modelo do Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA) [...] A Fiocruz tem o seu sistema, o Instituto Nacional de Pesquisas da Amazônia (INPA) tem o seu, as Federais cada uma tem o seu (Entrevistado (a) N° 7).

A partir do exposto, compreende-se que os repositórios podem ser caracterizados como temáticos, associados a instituições governamentais, não governamentais, universidades, bibliotecas e redes de observatórios ambientais, como afirmaram Michener e Jones (2012) e Vanz *et al.* (2018). Porém, observa-se que o *Ocean Biogeographic Information System (OBIS)* também não foi citado pela Persona Suporte.

Logo, entende-se que existe certa convergência no que tange aos mecanismos utilizados pela PS e pela PE para o compartilhamento e reuso de dados de pesquisa. Eles promovem o aumento dos estoques de conhecimento público, das taxas de retornos sociais dos investimentos em ciência e tecnologia, dos índices gerais de produtividade científica e de inovação (ALBAGLI, 2015). Isso aponta determinada aproximação das realidades o que pode beneficiar efetivamente o processo.

4.2.3 Medos/Insatisfações PS

A Persona Suporte apresenta alguns descontentamentos em relação ao processo que envolve o compartilhamento e reuso de dados. Todos vão ao encontro do mencionado pela

Persona Ecologia, como problema na documentação dos dados no momento da disponibilização, o financiamento restrito, a falta de infraestrutura e de recursos humanos. Os relatos da PS reforçam mais uma vez o mencionado por Anjos (2016) e Vanz *et al.* (2018) sobre os motivos para o não compartilhamento.

Às vezes uma pessoa coloca informações na internet e falta alguma coisa que quem documentou não lembrou. [...] dados brutos, primeiro costumam estar não tão bem documentados como dados finais [...] Muitas informações que ele tem ele não vai poder disponibilizar, porque ele não documentou adequadamente durante a coleta. [...] Ele não anotou antes, se ele não planejou isso antes de começar a pesquisa e já não se atentou em anotar o aparelho que ele usou, o *software* que ele usou, a hora que ele começou. [...] Se o projeto é para uma RNP tem um pré-requisito, se é para o CNPq tem outro pré-requisito [...] Tem uma dificuldade no compartilhamento de dados volumosos que é o *upload*, demora. [...] para dados, talvez as ferramentas não estejam maduras. [...] Nossa limitação é o incentivo da pesquisa como um todo. Qual é o incentivo que você tem? Em relação a dinheiro mesmo, em relação também a suporte, por exemplo, vai contratar uma equipe, mas tem tais formatos de bolsa, tem que ter um monte de exigências e às vezes os pesquisadores têm que lidar com esse tipo de situação que atrapalha. [...] falta incentivo, no sentido assim, em tentar resolver o problema a nível nacional. [...] tem algumas instituições se organizando e muitas ficando para trás, porque não tem profissional que pode se dedicar, o profissional já tem outras atribuições (Entrevistado (a) N° 1).

[...] teoricamente você deveria conseguir acessar os dados de quem já passou por lá. Ninguém sabia onde é que estava o dado de ninguém. Quando consegui acessar era essa coisa de tabela com célula amarela. Ninguém sabe porque está amarela. Se é um problema no dado, se é só porque é um valor estranho. É um erro de medida? [...] têm características muito específicas que muitos dos sistemas de informação, que são relevantes para a área da ecologia, eles não atendem, mas eles são muito utilizados [...] Essas são basicamente as questões. A pessoa não vai entender, não vai saber interpretar, vai fazer o uso indevido, não vai respeitar a licença. [...] Teve até uns editais do PELD que exigiam que os dados fossem depositados no SIBBR. Só que daí, o SiBBR não conseguiu oferecer o local para acomodar esses dados em tempo hábil, para que o que estava lá no edital fosse cumprido. Aí, meio que queimou o filme com o pessoal porque foi prometido que ia ter uma ferramenta, um repositório, ou um sistema. Algum lugar para colocar os dados. Aí, não foi oferecido. Teve alguns treinamentos meio frustrados. [...] É um desafio imenso porque as instituições têm suas prioridades, suas demandas e essas coisas têm custo. O custo computacional, material ele não é desprezível, ele é importante. [...] o custo de recursos humanos, ele é imprescindível. Ambos os custos, eles precisam ser pensados estrategicamente em longo prazo. [...] São muitos atores, muitas universidades. Vai pensar a quantidade de atores relevantes nesse ecossistema? O Brasil tem muitas universidades super boas, tem vários institutos de pesquisa relevantes também. Têm as agências financiadoras, órgãos de governo. [...] A gente tem avançado sim, mas a complexidade de você equilibrar as demandas e as ofertas. A oferta de recurso financeiro e de recurso humano, com as demandas de cada instituição para uma solução que agrade a todos. Não sei te dizer qual é a maior dificuldade porque tudo isso junto é que faz ser difícil. É muita variável para tratar (Entrevistado (a) N° 3).

Na realidade da PS, o uso da Linguagem de Metadados Ecológicos (*EML*) varia conforme seu núcleo de atuação, possibilitando inferir que a utilização ainda é restrita por parte desse perfil.

Os depoimentos da PS complementam o contexto da Persona Ecologia, a qual não usufrui do padrão na prática do compartilhamento, ou por falta de conhecimento ou de familiaridade.

Não, nunca ouvi falar. Até vou anotar. (Entrevistado (a) N° 1)

Sim, conheço. Foi o primeiro padrão de metadados que eu trabalhei. Eu até traduzi um *software* para português que já era em *EML*. [...] Na *Dado anonimizado* eu não consegui convencer o pessoal que a *EML* é legal. Foi mais no *Dado anonimizado* que eu trabalhei com *EML*. Em alguns projetos que eu participo de biodiversidade a gente utiliza também, mas nos repositórios da *Dado anonimizado* não usamos a *EML*. Eu trabalho em outros projetos, com outras instituições que a gente usa. (Entrevistado (a) N° 3)

Desse modo, é possível perceber a necessidade factual de ações que promovam a padronização dos dados de pesquisa ecológicos e o suporte adequado para a prática da ciência aberta na área. Essa demanda pode ser mais bem compreendida no texto seguinte.

4.2.4 Expectativas/Necessidades PS

A Persona Suporte se desenha com algumas expectativas e necessidades em relação à disponibilização de recursos humanos para apoio da Persona Ecologia, informações, orientações e treinamentos, para a prática fluída do compartilhamento e do reuso de dados na ecologia.

Acho que é importante saber que tipo de dados ele poderia reutilizar [...] fazer principalmente, conscientização da importância. [...] “o que tu acha do compartilhamento dos dados, com o pesquisador? Tu temes que os teus dados sejam aproveitados de uma forma? Tu tens algum receio de perder alguma coisa por compartilhar teus dados? Tu acha que seria mais vantajoso para tua pesquisa não compartilhar?” [...] Exige uma curva de conhecimento aí para conseguir disponibilizar dados. [...] precisa ter um mínimo de infraestrutura [...] precisamos estudar [...] ter bastante treinamento [...] tutorial de como fazer, instruções de como fazer, são importantes para que o pesquisador queira começar [...] Tem que começar lá no plano de gestão de dados, muitas vezes na coleta [...] Página simples, assim, para que ele leia, tenha as instruções, como fazer, uma falta de uma infraestrutura naquela instituição, que recomendem de outra, coisas assim. [...] O pesquisador, ele precisa de um apoio para ser conscientizado da importância da gestão de dados [...] Isso precisa acontecer mesmo antes da submissão (Entrevistado (a) N° 1).

Cursos foi um ponto levantado, precisaria oferecer mais cursos sobre esse processo da gestão da informação, de onde vem esses dados, como eles são planejados, das políticas, dos padrões. [...] é um caminho longo para você percorrer, se houvessem cursos esse caminho poderia ser encurtado [...] quais eram as boas práticas de limpeza de dados e quais eram as ferramentas já disponíveis e os padrões já disponíveis. [...] Tem muita coisa para ser feita, muitos desafios e muitas discussões que vão para um campo filosófico, às vezes. O que são dados abertos, quais os princípios, que o pessoal agora gosta de chamar de FAIR (*Findable, Accessible, Interoperable e Reusable*), quais são os

padrões. [...] uma coisa básica seria “Olha a gente vai financiar a sua pesquisa e você deposita os dados em alguma plataforma que você escolher, torne os dados públicos” (Entrevistado (a) N° 6).

A padronização dos metadados também foi citada como uma demanda pela Persona Suporte, como pela Persona Ecologia. Os depoimentos abaixo possibilitam inferir sobre essa necessidade eminente do contexto de acesso aberto aos dados de pesquisa ecológicos.

[...] um padrão brasileiro, que tem que ser atendido, certos metadados tem que ser preenchidos (Entrevistado (a) N° 1).

Como que eu vou comparar os dados daí, com os daqui da região de São Paulo? Usar os mesmos termos, os mesmos descritores, as mesmas unidades de medida, o mesmo vocabulário. [...] Vamos padronizar os dados. Vamos tornar o mais *FAIR* possível. Não vamos pensar em interoperabilidade dos XMLs dos metadados, mas nos dados mesmo. [...] Aí, quando forem dados ecológicos, a gente padronizar os termos, o vocabulário, as unidades de medida sempre que possível. Acho que tem muito trabalho para ser feito nesse sentido (Entrevistado (a) N° 3).

[...] padroniza o cabeçalho de todas as bases de dados para que você junte várias bases de dados numa coisa só e aí comece a analisar. [...] Estou falando, basicamente, do cabeçalho, nem fui a fundo sobre a padronização mesmo do dado. As coordenadas estão em graus decimais ou estão em outro formato, em UTM alguma coisa do tipo? [...] padronizar as coordenadas, as datas de coleta para que sigam um formato único ano, mês e dia (Entrevistado (a) N° 6).

O arquétipo suporte visa uma prática colaborativa para a submissão e compartilhamento dados, além de considerar essa uma expectativa também da Persona Ecologia. A especialização da PS na área se caracteriza como demanda para melhoria do processo.

Eles adorariam claro, ser ajudados, se tiver pessoas da ciência da informação para ajudar, especialistas para organizar os dados deles. [...] Ele vai preferir que uma pessoa faça tudo que a pessoa puder fazer. [...] O preenchimento dos metadados ou essas informações de ambiente, se ele pudesse delegar para alguém, ele ia ficar contente [...] A cada área essa pessoa da informação vai ter que ter um conhecimento avançado naquela área [...] Eu acho que devia ser um *upload* compartilhado. Digamos assim, que o pesquisador fizesse o *upload* inicial [...] talvez ele pudesse fazer esse *upload* inicial dos arquivos e em um segundo momento [...] isso teria que retornar ao pesquisador para ele dar anuência. [...] Entrar ali e fazer uma submissão colaborativa. (Entrevistado (a) N° 1).

Eles preferem eles mesmos poderem fazer o *upload*, mas eles gostam de ter uma orientação sobre a melhor forma de fazer. [...] No futuro vai ter profissionais mais especializados para cada parte da gestão de dados que é importante (Entrevistado (a) N° 3).

Tendo em vista a publicação de Vanz *et al.* (2018) é possível interpretar que a demanda de orientações específicas sobre o processo e de recursos humanos especializado, que prepare os dados de pesquisa para serem compartilhados, se concretiza ainda atualmente, passados quatro anos da *survey*. Além disso, observa-se que 64 dos 105 respondentes da *survey* selecionados informaram conhecer pouco sobre a gestão de dados de pesquisa. Assim sendo, entende-se que as iniciativas e ações em prol do acesso aberto aos dados de pesquisa ecológicos existentes ainda não oferecem subsídios no que tange a esses aspectos. Essa realidade se configura como desafio a ser vencido no contexto.

A seguir, apresenta-se a visão da Persona Suporte sobre a Persona Ecologia.

4.2.5 Persona Ecologia na visão da Persona Suporte

Esta subcategoria explicita qual a imagem que a Persona Suporte possui da Persona Ecologia, englobando aspectos que caracterizam perfil, prática de pesquisa, de compartilhamento e reuso, tipo de dados produzidos, compartilhados e reusados, medos/insatisfações e expectativas/necessidades. A PS compreende que a PE possui formação em ecologia, atua como pesquisadora e professora de cursos de graduação e pós-graduação em instituição pública, possui doutorado, é casada ou tem união estável, possui filhos, reside e trabalha na cidade, é branca, de orientação heterossexual e possui entre 30 e 60 anos de idade.

Na visão da PS, a PE é a primeira usuária de sistemas de informação e repositórios de acesso aberto aos dados de pesquisa, configurando-se como atora principal no contexto. A PE vai depositar os seus dados e reusar os de outros pesquisadores para gerar uma nova investigação. “O primeiro que vem na minha cabeça é o usuário que vai sentar com aqueles dados e daqueles dados partir para uma nova pesquisa. [...] pesquisadores que são os que vão depositar seus dados, serão usuários.” (Entrevistado (a) N° 1). “[...] acessar uma grande quantidade de dados para fazer perguntas mais abrangentes, estudos de síntese e tudo mais” (Entrevistado (a) N° 3).

Conforme a PS, a PE é motivada à prática do compartilhamento pela exigência de agências financiadoras, dos periódicos e pelo reconhecimento que isso pode gerar entre os pares, por meio de citações e agradecimentos.

Ela quer compartilhar, porque quer receber o reconhecimento. [...] pela exigência de uma agência de fomento e/ou exigência de um periódico (Entrevistado (a) N° 1).

Como é uma tendência mundial tem pesquisadores que estão muito mais antenados no que está acontecendo no mundo. Eles mesmos são exigidos a publicar os dados. [...] As pessoas querem ser reconhecidas pelo seu esforço [...] Tem revista que não aceita que esteja em qualquer lugar. Tipo essas que o pessoal usa do *Dryad* em geral, exigem que seja no *Dryad*. Então a pessoa vai colocar no *Dryad*, senão ela não vai publicar naquela revista que é boa, que é o público que ela quer atingir (Entrevistado (a) N° 3).

Essa percepção vai ao encontro do real contexto da PE e de Anjos (2016) e Vanz *et al.* (2018) quando citam as motivações do usuário de ecologia para a disponibilização. No entanto, a PS compreende que ainda há medos presentes na PE em relação ao compartilhamento de seus dados de pesquisa, o que também condiz com o apontado por Vanz *et al.* (2018) na *survey*, que evidenciou que dos 105 respondentes selecionados, 47 concordavam total ou parcialmente que suas ideias de pesquisa poderiam ser copiadas ao compartilhar.

Pesquisadores têm receio de colocar seus dados brutos [...] não querem compartilhar porque eles acham que podem usar. (Entrevistado (a) N° 1).

Não é todo mundo que se sente confortável publicando dado, tornando público e compartilhando [...] Têm pesquisadores que passaram a vida inteira fazendo uma base de dados de, sei lá, melhoramento genético do arroz. Para ele isso é muito novo, ele acha que alguém vai pegar e fazer o uso indevido, que não vai saber interpretar. [...] Têm pesquisadores que são muito mais abertos e interessados em disponibilizar os dados e outros que são muito mais resistentes. Cada um com seus motivos. (Entrevistado (a) N° 3).

A PS acredita que o perfil da ecologia gostaria de ter uma infraestrutura nacional que viabilize o acesso aberto aos dados de pesquisa.

Eu acho que eles gostariam que fosse a instituição. [...] A tendência é que os brasileiros queiram colocar em ferramentas dentro do Brasil. [...] Acho que eles queriam encontrar, talvez já uma infraestrutura, um portal com dados já. Porque daí, vendo o que tem lá, eles meio que tem uma ideia do tipo de informação que se espera encontrar. Que seja uma coisa disponível, que sempre esteja no ar, sempre disponível com uma instrução clara de qual o motivo que aquele repositório se propõe. O propósito é a preservação por muito tempo? Coisas assim. Porque é um trabalho para ele depositar. Já pensou, ele deposita hoje e amanhã o repositório se perde ou deixa de estar no ar? [...] um lugar seguro para armazenar (Entrevistado (a) N° 1).

Em geral os pesquisadores gostariam de ter um local brasileiro para colocar os dados. [...] Ainda não tem um repositório temático nacional em ecologia que atenda o que as pessoas precisam. As instituições estão, cada vez mais, amadurecendo com os seus repositórios institucionais. Isso ajuda, mas em geral quando um pesquisador recorre para um repositório internacional é porque ele ainda não encontrou uma ferramenta adequada nacional. Senão, ele teria tido preferência por colocar no nacional (Entrevistado (a) N° 3).

Vanz *et al.* (2018) apontaram a afirmação dos respondentes da *survey*, na época as instituições de atuação dos ecologistas nunca haviam disponibilizado repositórios para o compartilhamento.

O entendimento da PS sobre os tipos de dados gerados, compartilhados e reusados pela PE também demonstra conexão com a realidade do contexto dessa. Entre as características estão: dados observacionais, públicos, disponíveis para *download*, formato tabular, *.csv*, planilhas, textos, gerados por equipamentos, imagens, medidas, entrevistas, entre outros. Essa percepção confirma a diversidade dos dados ecológicos também na visão da PS.

[...] dados que são únicos [...] Acredito que formato tabular, tipo um *.csv*. É aquele formato de planilha só que o documento de texto. Que são linhas e colunas, dados observacionais assim, com medidas e informações onde o cruzamento das linhas com as colunas é que tem um significado. [...] precisa estar disponível para *download* (Entrevistado (a) N° 1).

A ecologia realmente é um ramo de pesquisa muito amplo, dependendo da pergunta que está sendo feita pode exigir que você tenha dados dos mais variados assuntos, inclusive agrícolas. Enfim, tem climáticos e dos próprios organismos. [...] Dados geográficos são basicamente mapas de uso da terra que muitas vezes são dados obtidos por imagem de satélite [...] os dados às vezes são sensores remotos, podem ser sensores de campo, assim, *in loco*, E pode ser também entrevista, medidas. [...] Todos esses dados são super diversos e eles são relevantes para a ecologia (Entrevistado (a) N° 3).

Três informações básicas, onde ocorre um indivíduo, o espaço, estou falando de coordenadas. Qual a informação do nome desse indivíduo, estou falando de taxonomia e qual o período temporal que foi registrado, uma data de coleta, uma hora de coleta, de um ano. [...] Existem outros dados de biodiversidade talvez ligados a essas três informações básicas, que são dados genéticos, dados de fisiologia, dados de atributos funcionais das espécies, dados de filogenia [...] ferramentas epidemiológicas, registros de doenças [...] Existem muitos outros tipos de dados [...] As outras que eu falei são derivações dessas informações básicas [...] dados públicos (Entrevistado (a) N° 6).

O conhecimento da PS vai ao encontro do quadro teórico dessa tese – Zimmerman (2008), Sagarin e Pauchard (2010), Michener e Jones (2012), Anjos (2016) e Vanz *et al.* (2018) – quando confirma a diversidade dos tipos de dados de pesquisa e formatos de armazenamento desses pela comunidade da ecologia.

A partir do exposto, percebe-se que a imagem que a Persona Suporte possui da Persona Ecologia corresponde a muitos aspectos que vão ao encontro da realidade, demonstrando o entendimento por parte da PS em relação ao arquétipo da ecologia. Entende-se que essa compreensão pode facilitar a conexão entre esses dois atores para ações que visam o acesso aberto aos dados de pesquisa da área.

5 COMUNICAÇÃO DA PERSONA ECOLOGIA E DA PERSONA SUPORTE

Esta seção representa o nascimento e o amadurecimento das personas construídas. Apresenta os dois arquétipos por meio de infográficos que visam proporcionar uma comunicação dinâmica, a fim de estimular a compreensão efetiva dos perfis por parte dos interessados (PRUITT; ADLIN, 2006).

O texto vai ao encontro do objetivo específico (d) dessa pesquisa – formular a apresentação das personas construídas por meio de instrumentos que facilitem a comunicação. Primeiramente, apresenta-se a Persona Ecologia, logo a Persona Suporte.

5.1 PERSONA ECOLOGIA (PE)

Figura 17: Infográfico perfil parte 1 - Persona Ecologia

A subcategoria “Perfil” foi dividida em duas partes. Esta foi composta por fotografias e elementos gráficos para dar vida e permitir a visualização da aparência do perfil.



Fonte: elaborado pela autora.

Figura 18: Infográfico perfil parte 2 - Persona Ecologia

Descreve características, interesses e habilidades do arquétipo.



Fonte: elaborado pela autora.

Figura 19: Infográfico prática da pesquisa - Persona Ecologia

Demonstra como se caracteriza a rotina do dia a dia de pesquisa da Persona Ecologia.

Retrata a inhomogeneidade dos contextos onde a investigação ecológica é desenvolvida. Os estudos ecológicos brasileiros são de larga escala e abordam questões complexas que vão desde o gene à biosfera.



Fonte: elaborado pela autora.

Figura 20: Infográfico prática do compartilhamento e reuso - Persona Ecologia

Apresenta a prática do compartilhamento e do reuso de dados de pesquisa.

De modo geral, o arquétipo compartilha seus dados e reusa os de outros pesquisadores ao longo dos anos de atuação, sendo ator ativo no processo de acesso aberto.

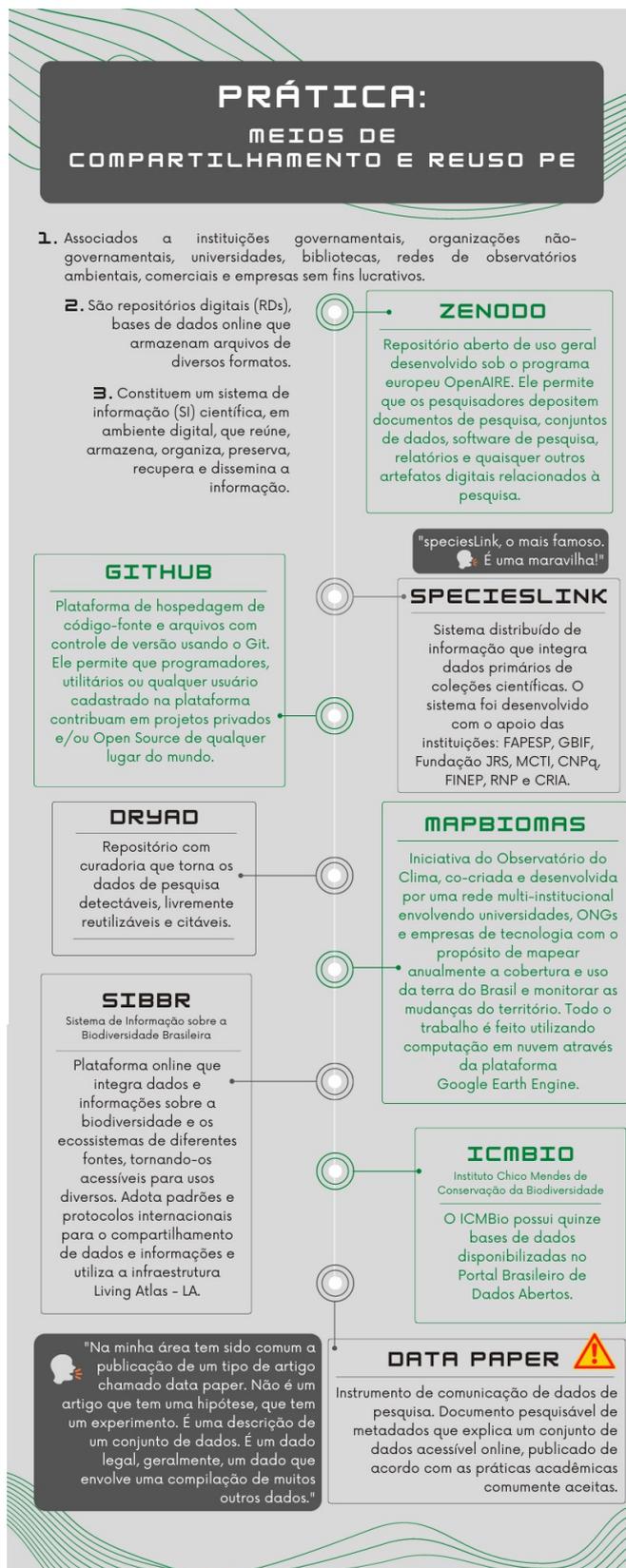


Fonte: elaborado pela autora.

Figura 21: Infográfico meios de compartilhamento e reuso - Persona Ecologia

Um meio de compartilhamento não mencionado no quadro teórico que apareceu na pesquisa qualitativa foi o *data paper*. Não se trata de um repositório ou sistema de informação, mas de um instrumento que possibilita a comunicação de dados de pesquisa.

O processo decisório sobre quais mecanismos utilizar acontece de modo espontâneo no dia a dia das investigações.



Fonte: elaborado pela autora.

Figura 22: Infográfico tipo de dado - Persona Ecologia

Ela produz e reutiliza uma grande variedade e quantidade de dados.

A abundância de tópicos estudados, dos tipos de dados coletados e das formas de armazenamento desses, afeta diretamente no processo de compartilhamento e reuso de dados de maneira integrada, fazendo do acesso aberto aos dados de pesquisa um desafio, quando tratado da ecologia.

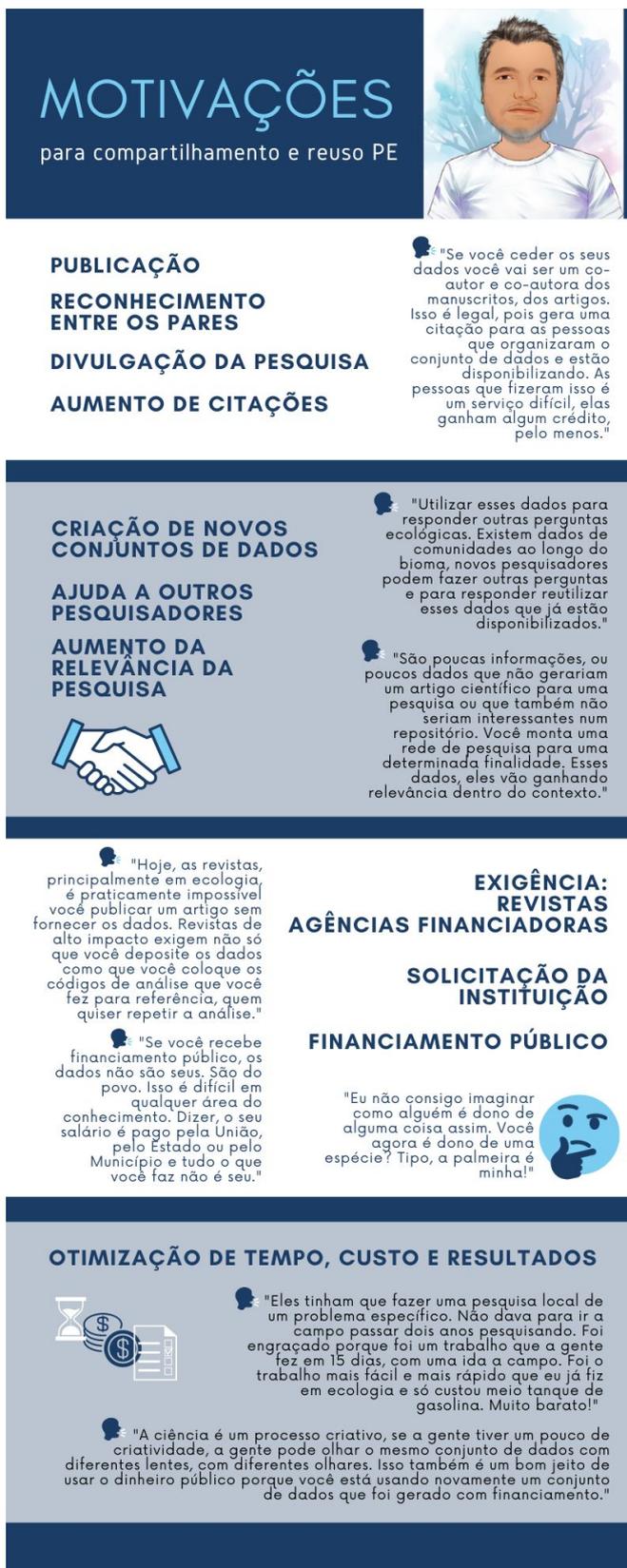


Fonte: elaborado pela autora.

Figura 23: Infográfico para motivações para compartilhamento e reuso - Persona Ecologia

Esse infográfico expressa as motivações do Personagem Ecologia em relação ao compartilhamento e reuso de dados de pesquisa.

O movimento oportuniza a reflexão sobre as atividades, processos, avaliações e produções científicas indo ao encontro da percepção sobre ciência aberta.



Fonte: elaborado pela autora.

Figura 24: Infográfico medos/insatisfações - Persona Ecologia

Explana aspectos que influenciam negativamente na prática da pesquisa, do compartilhamento e reuso de dados ecológicos.

Os diferentes tipos de dados tornam sua compreensão complexa, principalmente se reutilizados por outra pessoa, que não participou do levantamento original.



Fonte: elaborado pela autora.

Figura 25: Infográfico expectativas/necessidades - Persona Ecologia

Aponta desafios e demandas que devem ser atendidas pelas iniciativas que visam o acesso aberto aos dados de pesquisa ecológicos.



Fonte: elaborado pela autora.

5.2 PERSONA SUPORTE (PS)

Figura 26: Infográfico perfil - Persona Suporte

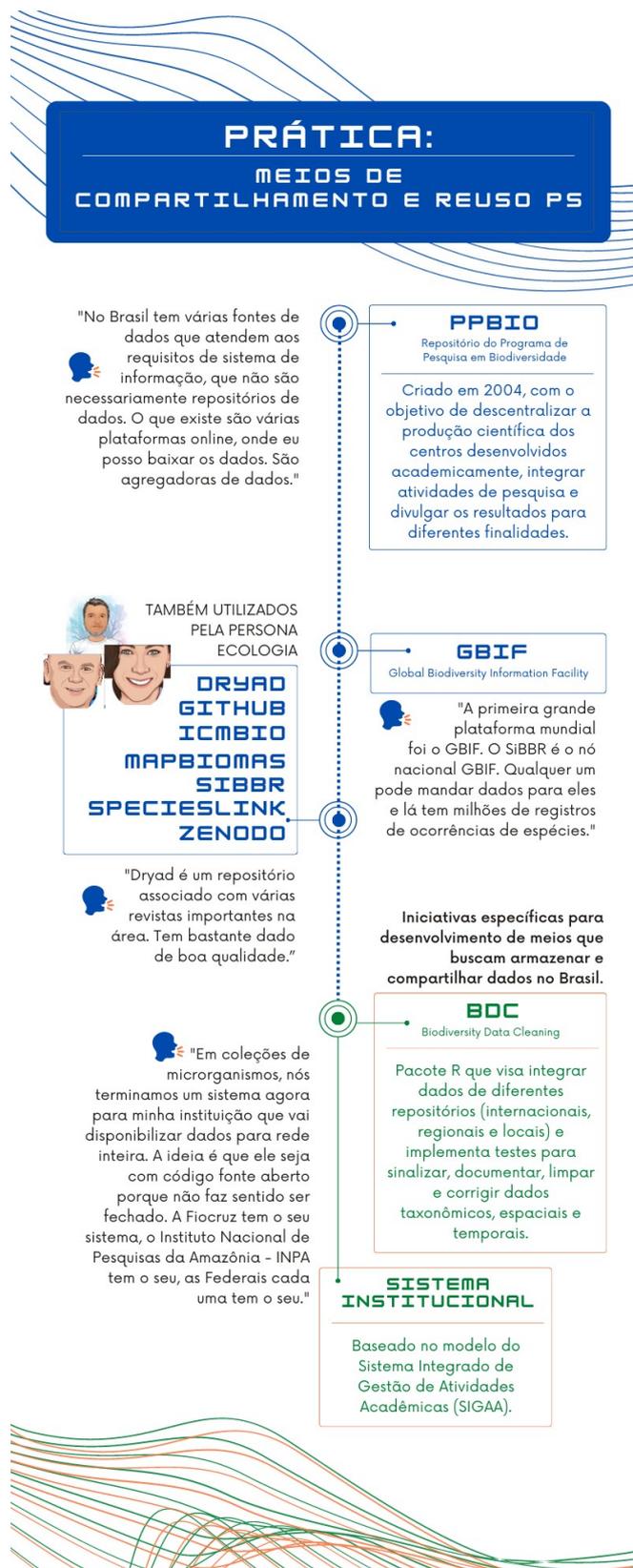
Profissional de apoio, com competência para gestão de dados ecológicos.



Fonte: elaborado pela autora.

Figura 27:
Infográfico meios de compartilhamento e reuso - Persona Suporte

Existe certa convergência no que tange aos mecanismos utilizados pela Persona Suporte e pela Persona Ecologia para o compartilhamento e reuso de dados de pesquisa.



Fonte: elaborado pela autora.

Figura 28: Infográfico comparativo: medos/insatisfações + expectativas/necessidades – Persona Suporte



Figura 29: Infográfico visão PS sobre PE

Explicita qual a imagem que a Persona Suporte possui da Persona Ecologia. A perspectiva corresponde a muitos aspectos que vão ao encontro da realidade, demonstrando o entendimento por parte da PS em relação ao arquétipo da ecologia.



Fonte: elaborado pela autora.

5.3 PERSONA ECOLOGIA E PERSONA SUPORTE ENQUANTO PÚBLICO

A partir dos dados analisados e discutidos durante a tese também foi possível reunir outros aspectos importantes sobre o contexto do acesso aberto aos dados de pesquisa que envolve a ecologia no âmbito brasileiro. Para tanto, esta subseção comunica por meio de duas imagens algumas conclusões sobre as personas enquanto público, já apresentadas em formato de texto nos subcapítulos 4.1 e 4.2. A figura 30 representa como se configura o ambiente relacional da PE, retrata os diferentes agentes de influência para a prática da pesquisa, do compartilhamento e do reuso de dados desse arquétipo.

Figura 30: Infográfico ambiente relacional Persona Ecologia



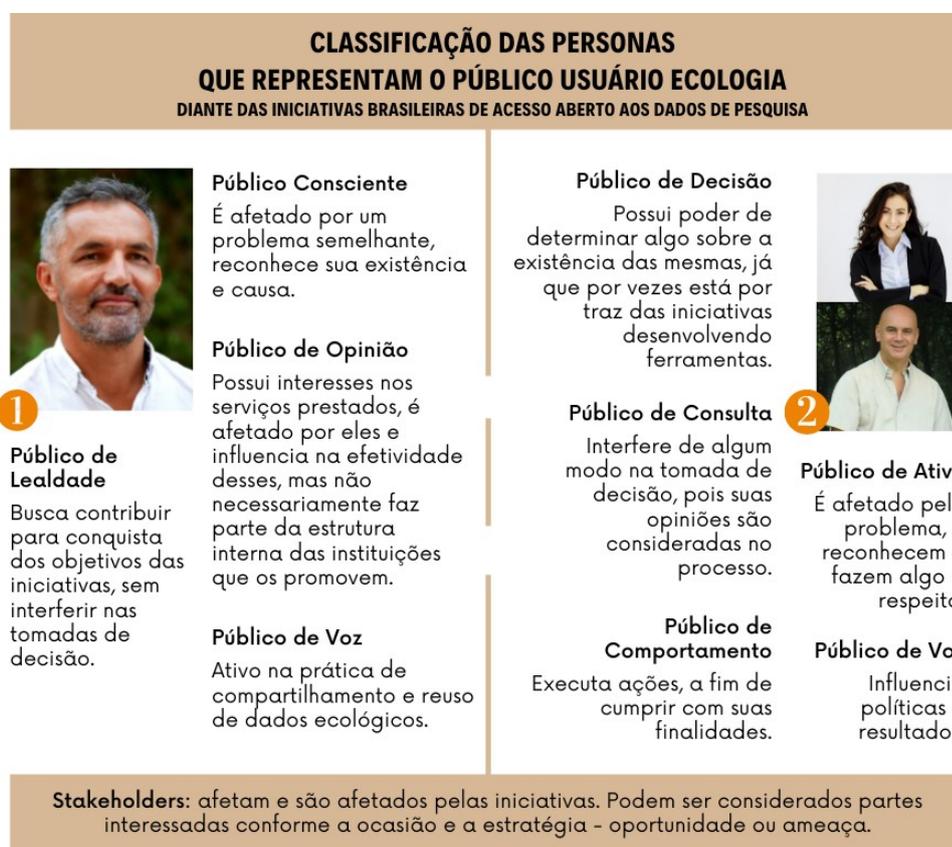
Fonte: elaborado pela autora.

Os dados indicam que o acesso aberto aos dados de pesquisa na ecologia envolve redes colaborativas que oferecem certo suporte ao processo. Apesar de ainda existirem insatisfações e necessidades, a PE percebe a importância dessas na conjuntura. Além disso, apontam outro

agente que influencia no contexto, a PS. Esse arquétipo também é um representante do público usuário, mas possui outros objetivos. Sua função e interesse é apoiar a PE na gestão, no compartilhamento e no reuso de dados.

A figura 31 comunica em quais classificações de públicos os dois arquétipos de usuário construídos são compreendidos, baseando-se nas referências apresentadas no subcapítulo 2.4 e na análise dos dados. Relembra-se que nessa pesquisa as personas foram entendidas como público usuário em função do sistema composto por diversas organizações que promovem iniciativas em defesa do acesso aberto aos dados de pesquisa no Brasil. No entanto, apesar de existirem ações concretas por parte das diferentes organizações, a articulação entre essas ainda se demonstra incipiente, como foi possível perceber durante a tese. Esse contexto abrange questões econômicas, sociais, tecnológicas, políticas e etc.. À vista disso, diferentes categorizações foram articuladas para compreender os perfis.

Figura 31: Quadro comparativo da Persona Ecologia e da Persona Suporte – enquanto público usuário de iniciativas brasileiras de acesso aberto aos dados de pesquisa



Fonte: elaborado pela autora.

A imagem acima sintetiza uma das perspectivas apontadas no início dessa investigação. Ou seja, o público usuário é representado por diferentes personas. Portanto, entender esse coletivo de modo amplo e homogêneo pode afetar diretamente no alcance de objetivos e metas das iniciativas, sejam eles de comunicação, de negócio, de tecnologia, etc.. Assim sendo, compreender como cada um dos perfis influencia no contexto, por meio de suas práticas, percepções, motivações, medos, etc. se torna essencial para a efetividade do acesso aberto aos dados de pesquisa no Brasil. Ao entendê-los é possível priorizá-los nas estratégias das diferentes etapas que envolvem a implantação das iniciativas.

A seguir são apresentadas as conclusões sobre essa investigação. Seção de encerramento que apresenta as reflexões da pesquisadora sobre o contexto, levando em consideração a pergunta norteadora, os objetivos e o quadro teórico dessa tese.

6 CONCLUSÕES

Esta seção trata do fechamento dessa investigação, momento de finalizar as percepções da investigadora sobre o que foi desenvolvido no conteúdo e a resposta da pergunta norteadora – Quem são as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa?

Foram encontrados dois arquétipos, a Persona Ecologia e a Persona Suporte. A Persona Ecologia era esperada como resultado de pesquisa, uma vez que o quadro teórico e os dados reutilizados apontaram indícios sobre a existência desse perfil. Sendo assim, a investigação teve como um de seus objetivos aprofundar o conhecimento sobre esse. Já a Persona Suporte foi se desenhando ao longo da análise dos dados coletados durante as entrevistas realizadas. Trata-se de uma descoberta dessa pesquisa. Os resultados reúnem uma série de dados coletados com pesquisadores reais, oportunizando a compreensão sobre seus perfis, contextos de pesquisa, de compartilhamento e de reuso de dados. Isso permitiu uma visão credível da complexa realidade dessa área do conhecimento quando se fala sobre acesso aberto e dados abertos de pesquisa. Demonstrando desafios que podem ser vistos como oportunidades ou ameaças, a depender da tomada de decisão.

Confia-se que as personas construídas podem ser úteis às iniciativas brasileiras de acesso aberto por um tempo ainda não determinado, porém duradouro. Isso devido às diversas possibilidades de cruzar e interpretar os resultados a partir dos infográficos. Os instrumentos de comunicação podem ser reunidos de diversas formas, possibilitando a união de aspectos-chave que podem oferecer novas percepções sobre os perfis. Os infográficos podem ser entendidos como peças de um quebra-cabeça que pode ser montado de diferentes modos, conforme os objetivos e metas traçados.

A Persona Ecologia é representada por um homem, brasileiro, branco, heterossexual de 49 anos de idade, morador da região sudeste do Brasil, casado, doutor com conhecimentos de ecologia e biologia, que desenvolve diferentes funções na universidade pública em que atua – professor, coordenador, pesquisador. Sua rotina de pesquisa envolve a saída de campo em várias regiões do país (conforme escopo do projeto) e o trabalho em laboratório. Sua rotina pessoal compreende a família e hábitos como leitura e arte como passatempos. Ele gosta de ler sobre diversos assuntos, mas mantém seu foco na ecologia. Geralmente busca informações na internet

em veículos reconhecidos, como os jornais brasileiros O Estadão e Folha, mas não costuma utilizar redes sociais como Facebook ou Instagram.

A PE se demonstrou ativa no movimento que busca o acesso aberto aos dados de pesquisa. Ou seja, não se ausenta do sistema relacional que envolve as práticas de compartilhamento e reuso desses dados. Ela é a líder do processo, pois ocupa posição de coordenação nos projetos de pesquisa em que atua. A PE compartilha seus dados de pesquisa, reusa dados de outros pesquisadores e fomenta o acesso aberto para seus pares. Ela reconhece a importância dos metadados no processo e tenta utilizar algumas informações padrões como data, hora, coordenadas, equipamento no momento da catalogação. Porém, não reconhece a *Ecological Metadata Language (EML)*, considerada o padrão ideal para descrição de dados ecológicos.

Constata-se que existem diferenças na gestão de dados de pesquisa, nas práticas de pesquisa, de produção, compartilhamento e reuso de dados por parte da PE. Confirma-se a variedade de técnicas e equipamentos utilizados para coleta, análise de dados como também dos formatos de armazenamento desses.

Conclui-se que os periódicos científicos e as agências de financiamento têm abordado a questão de abertura dos dados de pesquisa na área. Esses agentes consideram o compartilhamento um pré-requisito para publicação e financiamento, atuando como propulsores para a prática. Outras motivações apresentadas pela PE foram: reconhecimento entre os pares, divulgação da pesquisa, aumento de citações, criação de novos conjuntos de dados, ajuda a outros pesquisadores, aumento da relevância da pesquisa, solicitação da instituição, otimização de tempo, custo e resultados.

Apesar de a PE concordar com o compartilhamento e reuso de dados, ela ainda mantém certa tutela sobre esses, pois considera importante ser a primeira a utilizá-los em uma publicação. Assim sendo, seus dados são disponibilizados de acordo com alguns níveis de acesso. Ou seja, os dados até podem ser totalmente compartilhados, no entanto, isso ocorre ao longo do tempo da pesquisa. Junto a isso, a PE considera a segurança desses dados, uma vez que não confia totalmente nos atuais mecanismos de compartilhamento e reuso existentes. Isso se deve a falta de conhecimento sobre quem gerencia os sistemas e repositórios disponibilizados, como também pela infraestrutura ainda não adequada para a prática no que tange as especificidades e diversidade dos dados da ecologia. Os depoimentos que narram a prática e as motivações para compartilhamento e reuso (subseções 4.1.3 e 4.1.7) possibilitam perceber como esse arquétipo

pensa sobre direitos de reprodução, distribuição e utilização dos dados; e as indagações existentes no que tange a direitos autorais e patrimoniais sobre os dados. Apesar de os relatos permitirem inferir que a PE tem algum entendimento sobre o assunto, eles não mencionam explicitamente esses conceitos e nem a lei nº 9.610/1998 (BRASIL, 1998) vigente no país. À vista disso, compreende-se que se trata de uma pauta importante para ser aprofundada com esse perfil em ações de relacionamento e divulgação do acesso aberto aos dados de pesquisa.

Entre os mecanismos utilizados pela PE para compartilhar e reusar dados estão: *Dryad*, *GitHub*, *ICMBio*, *MapBiomas*, *SiBBR*, *speciesLink* e *Zenodo*, cada um com suas especificidades. O instrumento *data paper* aparece como um dos meios satisfatórios e frequentes para o compartilhamento e reuso de dados de pesquisa por esse arquétipo. O *data paper* possibilita publicar os dados de uma pesquisa em andamento, que ainda não possui resultados concretos, garantindo o reconhecimento de quem publicou e gerando sentimento de confiança desse indivíduo em relação à segurança dos dados disponibilizados.

O processo de decisão da Persona Ecologia sobre quais dados compartilhar e reusar, como também qual meio utilizar para a prática ocorre de modo natural no dia a dia de pesquisa. Os aspectos levados em consideração para a escolha são: o conhecimento da própria plataforma onde o dado será disponibilizado (funcionalidades, familiaridade com a interface, confiabilidade, precisão na descrição dos dados, etc.) e acompanhamento dos pares para saber em quais locais e formatos os dados estão sendo compartilhados.

A Persona Ecologia se desenha como público de *opinião* diante das iniciativas de acesso aberto aos dados de pesquisa, uma vez que possui interesses nos serviços prestados, é afetado por eles e influencia na efetividade desses, mas não necessariamente faz parte da estrutura interna das instituições que os promovem. Trata-se de um público *consciente*, pois os indivíduos que a compõem são afetados por problemas semelhantes, reconhecendo suas existências e causas. O perfil também foi percebido como um público de *lealdade* que busca contribuir para conquista dos objetivos das iniciativas, sem interferir nas tomadas de decisão. Assim como, um público de *voz*, uma vez que é ativo na prática de compartilhamento e reuso de dados ecológicos. Por isso, é essencial compreendê-la como um público de *consulta*, considerar suas opiniões no processo de tomada de decisão sobre ações que visam o acesso aberto aos dados de pesquisa ecológicos no Brasil. Entende-se que é estratégico se comunicar com esse público *consciente* para que ele se torne *ativo* diante das iniciativas. Ou seja, não apenas sendo afetado por elas e as reconhecendo,

como também fazendo algo em seu benefício. Isso pode beneficiar os resultados obtidos pelas promotoras das iniciativas. Conforme a análise, não foi possível percebê-la como um público de *consulta* no processo atualmente.

A Persona Suporte foi uma descoberta dessa tese, um segundo perfil do público usuário, que mesmo apresentando similaridades com a Persona Ecologia, exerce função diferente no processo de compartilhamento e reuso de dados ecológicos. A PS não foi apontada pelo quadro teórico, surgiu nos depoimentos que buscaram responder o problema e os objetivos dessa investigação. Ao percebê-la se constatou sua importância e influência na prática e por isso ela foi construída. No entanto, observa-se que ainda há necessidade de aprofundar o estudo sobre esse arquétipo, pois os dados coletados e analisados fornecem um conhecimento abrangente sobre ele.

A PS se configura como um (a) profissional especializado (a) na ecologia, com doutorado. Tem como função auxiliar a PE na gestão dos dados, atuando ativamente e facilitando o processo. Os dados obtidos permitem compreender que esse perfil pode pertencer ao gênero masculino ou feminino, com orientação heterossexual, branco (a), residir em diferentes regiões do Brasil, brasileiro (a), casado (a), tem filhos e pertence à faixa etária de 40 a 60 anos de idade. Atua como analista e pesquisador (a), sendo empregado (a) público (a) ou consultor (a) em instituição pública. Além disso, a PS parece ter como objetivo e prática de trabalho o desenvolvimento de mecanismos de compartilhamento e reuso de dados de pesquisa ecológicos – visando atender algumas demandas da área. No entanto, esses meios ainda são incipientes e por isso, pouco divulgados. Ou seja, são sistemas desenvolvidos e discutidos por esse arquétipo dentro de pequenos grupos com seus pares diretos. De todo modo, esses mecanismos já demonstram ser pensados buscando a garantia da disponibilização e da reutilização dos dados de pesquisa.

Compreende-se que a Persona Suporte consegue ter uma imagem coerente sobre a Persona Ecologia. Considerada pela PS a usuária principal para iniciativas brasileiras que visam o acesso aberto. A PS consegue perceber alguns aspectos sobre o perfil da PE, seus medos, expectativas e tipos de dados produzidos, compartilhados e reusados.

A PS se desenhou como um público de *decisão*, de *comportamento* e de *consulta* perante as iniciativas. Ela possui poder de determinar algo sobre a existência das mesmas, pois às vezes as desenvolve; executa ações, a fim de cumprir com suas finalidades. Portanto, a PS interfere de algum modo na tomada de decisão, pois suas opiniões são consideradas no processo. Trata-se de um público *ativo*, formado por indivíduos que são afetados pelo problema, o reconhecem e fazem

algo a respeito. Como também, um público de voz, já que atua diante das práticas diárias, influenciando políticas e resultados. Esse perfil varia conforme potencial de poder que pode exercer, com variada probabilidade de estimular ou desestimular ações em prol do acesso aberto aos dados de pesquisa. Portanto, esse arquétipo deve ser considerado estratégico no contexto de iniciativas que visam o acesso aberto aos dados de pesquisa ecológicos.

Constata-se que existe certa convergência nos dois arquétipos construídos no que se refere a alguns aspectos. Os contextos de suas rotinas de pesquisa e de trabalho são heterogêneos, os dados produzidos, manipulados, compartilhados e reusados são diversos. Entre eles: observacionais, computacionais, experimentais, brutos, derivados e publicados em algum formato aberto. A variedade de tipos de dados também é motivada pela pluralidade de aspectos físicos, químicos e biológicos de diferentes tipos de organismos monitorados, provenientes de diversos locais/sensores/ambientes e registrados por meio de planilhas, cadernos de campo e de laboratório, entre outros. Essa realidade torna o acesso aberto um desafio para a ecologia.

De todo modo, ambas as personas colaboram com o movimento da ciência aberta, no que tange às pautas de acesso aberto e dados abertos de pesquisa. Demonstram-se disponíveis e buscam se inteirar sobre o movimento constantemente. Suas práticas estimulam a translação do conhecimento, reforçando o conceito de responsabilidade científica na área da ecologia. Junto a isso, promovem processos, métodos e produções científicas que visam à transparência, à colaboração e à abertura de dados, para que esses possam ser disseminados, acessíveis e reutilizados pelos interessados.

Tanto a Persona Ecologia, quanto a Persona Suporte praticam o compartilhamento e o reuso de dados de pesquisa, utilizando diversos mecanismos para a prática. Os mais comuns entre elas são: *Dryad*, *GitHub*, *ICMBio*, *MapBiomas*, *SiBBR*, *speciesLink* e *Zenodo*. Esses podem ser entendidos como repositórios digitais, pois são bases de dados *on-line* que armazenam arquivos de diversos formatos. Como também sistemas de informação científica em ambiente digital, uma vez que reúnem, armazenam, organizam, preservam, recuperam e disseminam dados de pesquisa. Alguns dos mecanismos apresentam especificidades para a catalogação dos dados, o que também impõe certo desafio para o compartilhamento e reuso.

Concretizam-se como medo e insatisfação da PE e da PS a falta de recursos financeiros e de infraestrutura adequada para o compartilhamento e reuso de dados de pesquisa ecológicos. Nesse contexto, os perfis têm como expectativas e necessidades mais investimentos em recursos

humanos, financeiros, treinamento e uma infraestrutura nacional – que suporte e reúna a diversidade e a quantidade de tipos de dados produzidos pela ecologia. Como também, a garantia da segurança e da qualidade no armazenamento e na disponibilização desses. A ideia de uma articulação nacional para a implantação de um sistema de informação que possibilite a conexão de várias bases de dados está ligada ao avanço científico brasileiro que a ação pode proporcionar. Isso não exclui a possibilidade de parcerias internacionais. Porém, demonstra que os arquétipos compreendem que os dados produzidos no país devem ser prioritariamente assegurados pelo próprio. Neste momento, cabe a ressalva de que essa investigação ocorreu entre os anos de 2019 e 2022, momento em que o Brasil passava por instabilidades políticas, sociais e econômicas ligadas à pesquisa, ao ensino e ao meio ambiente. Confia-se que esse contexto motivou um olhar mais cuidadoso por parte das personas em relação à segurança e à importância do compartilhamento e reuso dos dados de pesquisa.

Verifica-se a eminente demanda de padronização dos dados de pesquisa no contexto ecológico de acesso aberto aos dados de pesquisa. Esse aspecto foi considerado pelas duas personas, tanto como medo/insatisfação, quanto expectativa/necessidade. Compreende-se que não necessariamente é preciso a elaboração de um novo padrão, como disse um dos entrevistados “não precisa reinventar a roda”. No entanto, é considerada uma avaliação mais profunda dos mecanismos de compartilhamento e reuso identificados como mais comuns na prática, para um alinhamento sobre quais dados e metadados são cruciais para uma comunicação fluída. Ou seja, para que o interessado nos dados disponibilizados pelo produtor, possa compreendê-los e reutilizá-los para responder outras perguntas ecológicas de modo eficiente e efetivo.

Ambas as personas são compostas por características reais de um grupo de indivíduos que possui interesses em comum e impacto real ou potencial sobre as iniciativas. São entendidas como *stakeholders*, pois afetam e são afetadas pelas ações, conforme o contexto e a estratégia adotada. Exercendo diferentes níveis de poder, uma vez que podem reagir sobre qualquer ação – a favor ou contra ela. O reconhecimento desses arquétipos como partes influentes no processo se torna essencial para a efetividade das iniciativas. A não distinção ou entendimento das especificidades desses perfis influenciam nas soluções propostas, pois ao não valorizá-los as promotoras das iniciativas possuem uma visão equivocada sobre a realidade do público usuário.

Percebe-se então, uma determinada aproximação das conjunturas da Persona Ecologia e da Persona Suporte. Essa realidade pode beneficiar efetivamente o processo de compartilhamento

e reuso de dados de pesquisa no âmbito da ecologia brasileira. Desde que esses arquétipos sejam percebidos e compreendidos não apenas pelos promotores das iniciativas, como também por si próprios como agentes do sistema relacional que o contexto propicia. Portanto, um dos desafios para a prática da ciência aberta na ecologia está na articulação entre os diversos atores, a fim de encontrar um caminho em comum para garantir o acesso aberto aos dados. Entende-se que essa tese aprofunda o conhecimento sobre o grupo de usuários, apresentando os arquétipos pertencentes à primeira das áreas do conhecimento mais citadas na *survey* do GT – RDP Brasil em 2018. Isso agrega valor às pesquisas já realizadas pelo GT, a fim de otimizar resultados e investimentos. Além disso, essa pesquisa favorece os pesquisadores das Ciências Biológicas de instituições brasileiras, especificamente da ecologia, já que compreende quem são esses indivíduos e o contexto do compartilhamento e reuso de dados entre esses, seus pares e a sociedade.

Observa-se a importância de fomentar o debate sobre ciência aberta e as pautas que dela fazem parte junto à comunidade científica da ecologia, a fim de ampliar a abrangência das pesquisas desenvolvidas e dos dados compartilhados para o público leigo. É possível verificar que as iniciativas e os meios utilizados para o processo priorizam um público seletivo, de especialistas – o que limita a divulgação científica. De todo modo, conclui-se que existe uma trajetória de ações concretas e rotineiras dos investigadores da ecologia em prol das pautas da ciência aberta, principalmente no que tange dados abertos e acesso aberto. Essas ações visam à cooperação para o avanço da ciência no âmbito global, à comunicação da ciência e tecnologia, à difusão e à disseminação científica.

Essa tese também possibilitou refletir sobre técnicas e instrumentos de pesquisa para identificação e mapeamento dos públicos. Conclui-se que a metodologia de construção de personas pode beneficiar essa atividade de RP. A união das perspectivas demonstra ser um caminho viável para aperfeiçoamento da prática investigativa do objeto público.

Considera-se que existe literatura de diferentes áreas do conhecimento sobre a metodologia de construção de personas que indicam como a proposta pode ser avaliada sistematicamente. A união de dados quantitativos, qualitativos, com as técnicas sugeridas e com foco em indivíduos reais, permite verificar se o grupo de arquétipos representa apropriadamente a população de interesse. O conceito de arquétipo agrega valor ao entender que as personas devem ser construídas a partir de dados coletados com representantes reais ou em potencial do público –

tornando-o o centro do processo. Esse movimento permite ampliar a visão e a compreensão sobre os diversos indivíduos que compõem determinado coletivo, contemplando seus contextos, características e aspectos. Possibilita também a descoberta de mais de uma representação para um mesmo público. Nesse contexto, entende-se que não é possível obter informações reais sobre os públicos, analisando somente perspectivas centradas nas organizações, suas iniciativas e contextos. Isso pode gerar a concepção de estereótipos que representam vieses e suposições, já que considera apenas um dos lados do sistema relacional. Desse modo, a partir da perspectiva da construção de personas, percebe-se que não há um único perfil, ou número exato de perfis que representam um determinado público. Ou seja, um público específico pode abranger uma diversidade de arquétipos que possuem interesses e exercem influência distintas sobre estratégias que guiam o desenvolvimento de ações, iniciativas, produtos, serviços, etc. Essas personas podem se enquadrar em diferentes classificações de público de acordo com o contexto. Portanto, classificar um público de modo heterogêneo, sem considerar seus diferentes perfis, pode limitar sua compreensão. Como visto nessa investigação, o público usuário das Ciências Biológicas - ecologia de iniciativas em defesa do movimento brasileiro de acesso aberto aos dados de pesquisa é composto por mais de um arquétipo. Assim sendo, considera-se que ter o conhecimento detalhado sobre o público permite atingir com mais êxito objetivos e metas.

O entendimento sobre as fases do ciclo de vida de personas, que envolve desde o planejamento familiar até a aposentadoria dos perfis, também colabora com a atividade de RP. Um dos benefícios é o conceito de esqueleto – parte da fase de concepção das personas – que permite uma imagem prévia do público, já que é desenhado a partir de dados factuais preliminares obtidos na fase de planejamento familiar. Esse esboço possibilita identificar aspectos conhecidos sobre determinado público e apontar os ainda desconhecidos ou que precisam de aprofundamento, serve como guia para o conhecimento do público.

Outro aspecto interessante da metodologia de construção das personas são os modelos de apresentação sugeridos pelos autores. Formados por narrativas compostas por imagem e texto, podem facilitar a comunicação dos perfis, estimulando a sapiência mais profunda por parte dos interessados. As propostas dos autores que abordam o objeto público, que contemplam uma possível apresentação, focam em tabelas e quadros. Desses, o conceito de Mapa da Persona Organizacional é o que mais contribui para a atividade de relações públicas. O Desenho da

Persona a valoriza no decorrer do processo e as demais estruturas permitem organizar os dados de um modo que facilita a construção dos perfis.

Compreende-se que as técnicas qualitativas são fundamentais para a identificação e o mapeamento dos públicos, pois proporcionam uma visão mais completa sobre o objeto investigado. Ademais, concorda-se que a revisão de literatura é essencial na etapa de levantamento de dados. Por meio da pesquisa bibliográfica é possível mapear dados significativos em relação ao objeto. As técnicas quantitativas também beneficiam o processo oferecendo precisão para análise dos dados, pois favorecem a percepção de quem são os diversos indivíduos num contexto heterogêneo.

Porém, ainda há necessidade de aprofundar os conhecimentos práticos e teóricos sobre como a metodologia de construção de personas pode beneficiar a identificação e o mapeamento dos públicos. É preciso diversificar as visões e ampliar a literatura para avaliar a união de perspectivas, a fim de promover mais confiabilidade, validade e utilidade da proposta metodológica para a atividade de RP na perspectiva de sistema organizacional. Observa-se que o fato de a conjuntura do acesso aberto aos dados de pesquisa no Brasil apresentar um sistema composto por diferentes organizações e coletividades demonstra o desafiador contexto para a identificação e o mapeamento de públicos aos profissionais da área. Compreende-se que ao centrar a atividade numa perspectiva organizacional como é atualmente, sem considerar os possíveis sistemas a serem formados, os profissionais suprimem suas competências enquanto especialistas em levantamento de perfis estratégicos para outros cenários, desencadeados por demandas sociais e mercadológicas num ambiente global competitivo e dinâmico.

Do mesmo modo, essa tese apontou contribuição do conceito de públicos em RP para o de personas. Entender os arquétipos enquanto públicos, possibilitou perceber de modo mais assertivo seus interesses em comum em relação ao serviço prestado/produto ofertado; como são afetados e afetam as ações das promotoras, na conquista de suas missões, objetivos e metas, e; o potencial de poder que podem exercer ante os interesses das iniciativas, promovendo ou dificultando suas ações. Confia-se que os resultados dessa investigação também possam agregar valor ao contexto de iniciativas brasileiras que visam o acesso aberto aos dados de pesquisa e aos resultados da *survey* do GT – RDP Brasil (2018). Desde que sejam efetivamente e estrategicamente utilizados pelas partes promotoras nos momentos de tomada de decisão. A comunicação das personas por meio dos infográficos possibilita reorganizar e cruzar as

informações sobre cada um e entre eles em outras perspectivas que não as dessa tese. Valorizando cada aspecto conforme objetivos, metas e andamento das iniciativas, otimizando resultados e investimentos. Junto a isso, entende-se que a proposta metodológica dessa pesquisa, os dados coletados e as personas apresentadas podem auxiliar no fortalecimento de redes colaborativas que visam o acesso aberto aos dados de pesquisa, como as da RNP, por exemplo. A formação de parcerias foi algo que apareceu nas práticas de pesquisa, compartilhamento e reuso da Persona Ecologia (subseções 4.1.2 e 4.1.3). Além disso, foi possível perceber os diferentes grupos envolvidos no contexto da ciência aberta (figura 2). Portanto, um trabalho de identificação e mapeamento de públicos, visando compreender os diversos coletivos envolvidos de modo mais amplo, levanta dados para um planejamento estratégico de comunicação. Através desse, pode-se potencializar o relacionamento e a interação da RNP com os coletivos e entre eles, buscando gerar novas parcerias e agregar outras redes colaborativas ao sistema.

Por fim, considera-se que mesmo com os resultados importantes dessa tese, ainda é possível destacar ideias para futuras investigações relacionadas à identificação e ao mapeamento de públicos, com base na construção de personas, para iniciativas em defesa do movimento de acesso aberto aos dados de pesquisa no Brasil. Isso pôde ser percebido no ambiente relacional que envolve a ciência aberta e as práticas de pesquisa, compartilhamento e reuso de dados na área da ecologia, demonstrados por meio das figuras 2 e 30 no decorrer do texto. Entre algumas das perguntas que ainda podem ser feitas estão: Quais outras personas podem representar o público usuário de iniciativas brasileiras de acesso aberto aos dados de pesquisa? Quem são as personas que representam o público cidadão brasileiro que busca por dados abertos de pesquisa? Quem são as personas que representam o público políticos brasileiros que fomentam leis em defesa do movimento de acesso aberto aos dados de pesquisa? Ou seja, é possível investigar mais a fundo os diversos coletivos, a fim de direcionar estratégias que beneficiam a ciência aberta brasileira como um todo.

REFERÊNCIAS

5-STAR. **5-Star OPEN DATA**. *On-line*. Disponível em: <http://5stardata.info/en>. Acesso em: 22 fev. 2023.

ALBAGLI, Sarita. *et al.* **Ciência Aberta**: correntes interpretativas e tipos de ação. *Liinc em Revista*, Rio de Janeiro, v.10, n.2, p. 434 – 450, novembro, 2014. Disponível em: <http://www.ibict.br/liinc>. Acesso em: 14 abr. 2020.

ALBAGLI, Sarita. Divulgação científica: informação científica para cidadania. **Ciência da informação**, v. 25, n. 3, 1996. Disponível em: <https://revista.ibict.br/ciinf/article/view/639/643>. Acesso em: 23 mai. 2022.

_____. Ciência aberta em questão. In: **Ciência aberta, questões abertas** / Sarita Albagli, Maria Lucia Maciel e Alexandre Hannud Abdo org. – Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p 9 – 26.

_____. 2 Ciência Aberta: movimento de movimentos. In: **Ciência Aberta Para Editores Científicos**, p. 15, 2019. Disponível em: https://www.researchgate.net/profile/Milton-Shintaku/publication/336027091_Abertura_da_ciencia_e_os_editores_cientificos/links/5ee762c8458515814a5eaaf8/Abertura-da-ciencia-e-os-editores-cientificos.pdf#page=16. Acesso em: 29 jul. 2021.

ALBANO, Claudio Sonaglio; LEAES JR, Walter; CORSO, Kathiane Benedetti. Dados abertos: um estudo em trabalhos acadêmicos no contexto brasileiro. **Navus-Revista de Gestão e Tecnologia**, v. 9, n. 2, p. 135 – 147, 2019. Disponível em: https://www.researchgate.net/publication/332128502_Dados_abertos_um_estudo_em_trabalhos_academicos_no_contexto_brasileiro. Acesso em: 16 jun. 2021.

ANJOS, Renata Lemos dos *et al.* 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** – Departamento de Ciência da Informação, Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba, 2016. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/1874>. Acesso em: 02 nov. 2022.

AQUINO, Plínio Thomaz; FILGUEIRAS, Lucia V. L. A Expressão da Diversidade de Usuários no Projeto de Interação com Padrões e Personas. **VIII Simpósio Sobre Fatores Humanos em Sistemas Computacionais**. October 21 – 24, 2008. *On-line*. Disponível em: <https://goo.gl/NQ7LzL>. Acesso em: 18 jun. 2020.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Normas para citações e referências Bibliográficas**. São Paulo: ABNT, 2003.

ATOJI, Marina. CGU realiza consulta pública para priorizar temas do 5º Plano de Governo Aberto. **Blog Transparência Brasil**, 23 de fevereiro de 2021. Disponível em: <https://blog.transparencia.org.br/cgu-realiza-consulta-publica-para-priorizar-temas-do-5o-plano-de-governo-aberto/>. Acesso em: 02 ago. 2021.

BASE DE CONHECIMENTO DA CONTROLADORIA-GERAL DA UNIÃO. 1º plano de ação nacional em governo aberto, 2011. Disponível em: <https://repositorio.cgu.gov.br/handle/1/46243>. Acesso em: 30 jul. 2022.

BARDIN, Laurence. **Análise de conteúdo**. Tradução Luís Antero reto, Augusto Pinheiro. São Paulo : Edição 70, 2016.

BARROS, Evandrino Gomes. Um ambiente baseado em arquivos abertos para integração de dados ecológicos. **Dissertação** Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais, 2005. Disponível em: <https://repositorio.ufmg.br/handle/1843/SLBS-6GUMPG>. Acesso em: 17 jun. 2021

BARROS, Evandrino Gomes. *et al.* Uma Biblioteca Digital Georreferenciada para Dados Ecológicos. In: **20º Simpósio Brasileiro de Bancos de Dados (SBBDD)**, 3 – 7 de Outubro, Uberlândia, MG, Brazil, 2005. p. 175 – 189. Disponível em: https://www.researchgate.net/profile/Alberto-Laender/publication/221535954_Uma_Biblioteca_Digital_Georreferenciada_para_Dados_Ecológicos/links/0fcfd50b7b7280c33c000000/Uma-Biblioteca-Digital-Georreferenciada-para-Dados-Ecológicos.pdf. Acesso em: 03 ago. 2021.

BARROS, Evandrino G. *et al.* A digital library environment for integrating, disseminating and exploring ecological data. **Ecological Informatics**, v. 3, n. 4 – 5, p. 295 – 308, 2008. Disponível em: <https://doi.org/10.1016/j.ecoinf.2008.09.006>. Acesso em: 04 ago. 2021.

BAYNES, Grace. **Colaboração e cooperação são fundamentais para tornar a ciência aberta uma realidade**, 6 de novembro de 2017. *On-line*. Disponível em: <https://blog.scielo.org/blog/2017/11/06/colaboracao-e-cooperacao-sao-fundamentais-para-tornar-a-ciencia-aberta-uma-realidade-publicado-originalmente-no-blog-lse-impact-of-social-sciences-em-outubro2017/#.W2iY3dVKi70>. Acesso em: 06 ago. 2018.

BIOLOGIANET. **Website**. Biologia, 2020. *On-line*. Disponível em: <https://www.biologianet.com/curiosidades-biologia/biologia.htm#:~:text=Biologia%20%20a%20ci%20%20respons%C3%A1vel,fatores%20abi%C3%B3ticos%20e%20eventos%20biol%C3%B3gicos.&text=Mais%20que%20isso%20ela%20tamb%C3%A9m,interferem%20na%20exist%C3%Aancia%20da%20vida>. Acesso em: 18 out. 2020.

BORGMAN, Christine L. **Scholarship in the digital age: information, infrastructure, and the Internet**. Massachusetts: MIT Press, 2007.

BUENO, Wilson Costa. Comunicação científica e divulgação científica: aproximações e rupturas conceituais. **Informação & Informação**, v. 15, n. 1esp, p. 1 – 12, 2010. Disponível em: <https://www.brapi.inf.br/index.php/res/v/33484>. Acesso em: 16 out. 2022

CALDEIRA, Carlos. **Definição de base de dados relacional**, 2009. *On-line*. Disponível em: <http://www.di.uevora.pt/~ccaldeira/definicao-de-base-de-dados-relacional/>. Acesso em: 20 ago. 2018.

CANALES, Manuel; PEINADO, Anselmo. *Grupos de discusión*. In: **Métodos y Técnicas Cualitativas de Investigación em Ciências Sociais**. DELGADO, Juan Manuel; GUTIÉRREZ, Juan, 4 reimpressão, 2007, p. 289 – 316.

CAREGNATO, Rita Catalina Aquino; MUTTI, Regina. Pesquisa qualitativa: análise de discurso versus análise de conteúdo. **Texto & Contexto – Enfermagem**, v. 15, p. 679 – 684, 2006. Disponível em: <https://doi.org/10.1590/S0104-07072006000400017>

CAREGNATO, Sônia Elisa et. al. Práticas e percepções dos pesquisadores brasileiros sobre serviços de acesso aberto a dados de pesquisa. **Liinc em Revista**, Rio de Janeiro, v.15, n.2, p. 121 – 141, novembro 2019. <https://doi.org/10.18617/liinc.v15i2.4771>. Disponível em: <http://www.ibict.br/liinc>. Acesso em: 03 abr. 2020.

CARVALHO, Érika Rayanne Silva. Diferenças na produção, compartilhamento e (re) uso de dados: a percepção de pesquisadores da química, antropologia e educação. **Dissertação** Programa de Pós-Graduação em Ciência da Informação da Faculdade em Ciência da Informação da Universidade de Brasília, 2018.

CAVALCANTE, Ricardo Bezerra; CALIXTO, Pedro; PINHEIRO, Marta Macedo Kerr. Análise de conteúdo: considerações gerais, relações com a pergunta de pesquisa, possibilidades e limitações do método. **Informação & sociedade: estudos**, v. 24, n. 1, p. 13 – 18, 2014.

CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE (CIDACS). **Glossário de Termos**. Disponível em: <https://cidacs.bahia.fiocruz.br/glossario-de-termos/>. Acesso em: 22 dez. 2022.

CHAPMAN, Christopher N.; MILHAM, Russell P. **The Personas' New Clothes: Methodological and Practical Arguments Against A Popular Method**. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. First Published October 1, 2006 Review Article. <https://doi.org/10.1177/154193120605000503>. p. 634 – 636. Disponível em: https://journals.sagepub.com/doi/abs/10.1177/154193120605000503?casa_token=SAf0JHiZQbsAAAAA%3AtvYUwdBfLr2-X49-DOIZ1P4BJyXdHavkixbSXrJEVNH9rTeEym7kze4n2zoUWmoqug6x6egH5Odk&. Acesso em : 09 jul. 2020.

CHAVAN, V.; PENEV, L.. *The data paper: a mechanism to incentivize data publishing in biodiversity science*. **BMC bioinformatics**, v. 12, n. 15, p. 1 – 12, 2011. Disponível em: <https://doi.org/10.1186/1471-2105-12-S15-S2>. Acesso em: 24 mai. 2022.

COLEMAN, David C. **Big ecology**, Berkeley: *University of California Press*, 2010. Disponível em: <https://doi.org/10.1525/9780520945739>. Acesso em: 24 mai. 2021.

CONRERP 2ª REG. **A profissão**. Todos os direitos reservados, 2016. *On-line*. Disponível em: <http://conrerp2.org.br/institucional/profissao-e-mercado-de-trabalho/>. Acesso em: 13 nov. 2020.

CONTROLADORIA-GERAL DA UNIÃO. Governo Aberto. *Open Government Partnership*. O que é a iniciativa. 12 de setembro de 2022. Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/o-que-e-a-iniciativa>. Acesso em: 18 dez. 2022.

_____. 5º Plano de Ação Brasileiro. Governo Aberto. Planos de Ação, 2021. Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/5o-plano-de-acao-brasileiro>. Acesso em: 02 ago. 2021.

COOPER, Alan. *The inmates are running the asylum*. Sams Publishing, 2004.

COOPER, Alan; REIMAN, Robert. *About Face 2.0: The Essentials of Interaction Design*. New York: John Wiley and Sons, 2003.

COOPER, Alan; REIMANN, Robert; CRONIN, Dave. *About Face 3. The Essentials of Interaction Design*. [S.l.] Wiley Publishing, 2007.

COOPER, Alan. *The origin of personas*. *Cooper.com Journal*, 2008. On-line. Disponível em: https://www.cooper.com/journal/2008/05/the_origin_of_personas/. Acesso em: 13 nov. 2020.

CORREIA, Marcos B. F. Dados, Informação e Conhecimento. In: **A comunicação de dados estatísticos por intermédio de infográficos: uma abordagem ergonômica**. Dissertação (Mestrado em Design) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009. Disponível em: https://www.maxwell.vrac.puc-rio.br/14038/14038_4.PDF. Acesso em: 14 ago. 2018.

DA CUNHA, Murilo Bastos. Metodologias para estudo dos usuários de informação científica e tecnológica. *Revista de Biblioteconomia de Brasília*, v. 10, n. 2, p. 5 – 19, 1982. Disponível em: https://www.brapci.inf.br/repositorio/2010/02/pdf_a7a477d359_0008278.pdf. Acesso em: 25 ago. 2022

DAMASCO, Miguel. **Conceito de Sistemas de Informação**, 2015. Disponível em: <http://www.profdamasco.site.br.com/SlidesFundamentosSI.pdf>. Acesso em: 17 ago. 2018.

DE AVELLAR BARBOSA, Carolina. Relações Públicas somente para grandes empresas? *Organicom*, v. 3, n. 5, p. 132 – 151, 2006. Disponível em: <https://www.revistas.usp.br/organicom/article/view/139996>. Acesso em: 28 jun. 2021.

DECRETO Nº 8.777, DE 11 DE MAIO DE 2016. **Institui a Política de Dados Abertos do Poder Executivo Federal**. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 15 jun. 2021

DEWES, João Osvaldo. Amostragem em bola de neve e *respondent-driven sampling*: uma descrição dos métodos. 2013. 53f. **Trabalho de Conclusão de Curso** (Monografia apresentada para obtenção do grau de Bacharel em Estatística) – Instituto de Matemática e Estatística, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/93246/000915046.pdf?sequence=1&isAllowed=y>. Acesso em: 05 nov. 2020.

DUDZIAK, Elisabeth. **Dados de Pesquisa agora devem ser armazenados e citados** [on-line], 2016. Disponível em: <http://www.sibi.usp.br/noticias/dados-materiais-metodos-revistas-exigem-dados-pesquisa-estejam-disponiveis/>. Acesso em: 31 jul. 2018.

DRUCKER, Débora Pignatari. Avanços na integração e gerenciamento de dados ecológicos. **Natureza & Conservação**, v. 9, n. 1, p. 115 – 120, 2011. Disponível em: https://www.conecte.bio.br/referencias/Drucker_2011a.pdf. Acesso em 05 ago.2021.

FAUSTO, S. Evolução do Acesso Aberto – breve histórico [on-line]. **SciELO em Perspectiva**, 2013. Disponível em: <https://blog.scielo.org/blog/2013/10/21/evolucao-do-acesso-aberto-breve-historico/>. Acesso em: 20 ago. 2018.

FIGUEIREDO, N. M. de. **Estudos de usos e usuários da informação**. Brasília: IBICT, 1994.

FONTANELLA, Bruno Jose Barcellos; LUCHESI, Bruna Moretti; SAIDEL, Maria G. Borges; RICAS, Janete; TURATO, Egberto R.; MELO, Débora Gusmão. **Amostragem em pesquisas qualitativas**: proposta de procedimentos para constatar saturação teórica. *Cad. Saúde Pública*, Rio de Janeiro, 27(2):389 – 394, fev, 2011. Disponível em: <https://www.scielo.org/pdf/csp/2011.v27n2/388-394/pt>

FORTES, Waldyr Gutierrez. **Relações Públicas**: processo, funções, tecnologia e estratégias. 2ª Ed. São Paulo: Summus, 2003.

FRANÇA, Fábio. **Públicos**: como identificá-los em nova visão estratégica. 3 ed. São Caetano do Sul – SP :Yendis Editora, 2012.

FREEMAN, R. Edward. **Strategic Management: a Stakeholder Approach**. Boston, MA: Pitman, 1984.

FREHNER, Marcel; BRÄNDLI, Martin. *Virtual database: Spatial analysis in a Web-based data management system for distributed ecological data*. **Environmental Modelling & Software**, v. 21, n. 11, p. 1544 – 1554, 2006. Disponível em: https://www.sciencedirect.com/science/article/pii/S1364815206001113?casa_token=UuUg0yDeDNYAAAAA:iE7o2y7DCjL19EuZh3MGMz1VjFICzLjN9D29I1MGMn6eUUsROEk1FanmvVab9UfGpb9yKDY. Acesso em: 16 jun. 2021.

FRIEDMAN, Andrew L.; MILES, Samantha. **Stakeholders: Theory and Practice**, OUP Oxford, 2006.

FRIESIKE, Sascha; FECHER, Benedikt. **Open Science: one term, Five schools of thought**. In: BARTLING, Sonke; FRIESIKE, Sascha (Ed.). *Opening Science: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing*. London: Springer Open, 2014, p. 17 – 47.

FUNDAÇÃO OSWALDO CRUZ (FIOCRUZ). Repositório Institucional – Arca. [on-line] Disponível em: <https://portal.fiocruz.br/repositorio-institucional-arca>. Aceso em: 23 jun. 2022.

_____. **Ciência Aberta na Fiocruz**. Comunicação e informação » Ciência Aberta, 2022. Disponível em: <https://portal.fiocruz.br/ciencia-aberta-na-fiocruz>. Acesso em: 29 jul. 2022.

_____. **Glossário do Acesso Aberto**. Disponível em: <https://portal.fiocruz.br/glossario-acesso-aberto/G>. Acesso em: 22 dez. 2022.

GALEANO, Eduardo. *Las palabras andantes*. Espanha: Siglo XXI, 1995.

Global Biodiversity Information Facility (GBIF). Website, 2021. Disponível em: <http://www.gbif.org>. Acesso em: 14 jun. 2021.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4^a ed. Editora Atlas, 2002.

GOODWIN, Kim. *Getting from Research to Personas: Harnessing The Power of Data*, 2002. *On-line*. Disponível em: https://articles.ue.com/research_to_personas/. Acesso em: 13 nov. 2020.

GREEN, Ann; MACDONALD, Stuart; RICE, Robin. *Policy-making for Research Data in Repositories: A Guide*. EDINA, 2009. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.218.467&rep=rep1&type=pdf>. Acesso em: 21 jul. 2021.

GRUDIN, Jonathan; PRUITT, John. *Personas, Participatory Design and Product Development: na infraestrutura for engagement*. In: **Proc. PDC**. Palo Alto, 23 – 25 Jun, 2002. *On-line*. Disponível em: <https://www.ece.uvic.ca/~aalbu/CENG%20412%202009/scenarios%20and%20personas.pdf>. Acesso em: 13 nov. 2020.

GRUNIG, James E.; HUNT, Tood. *Managing Public Relations*. Holt, Rinehart and Winston. ISBN: 0-03-058337-3, 1984. Disponível em: https://www.researchgate.net/publication/322802009_Managing_Public_Relations. Acesso em: 09 nov. 2022.

GRUPO DE TRABALHO DA REDE DE DADOS DE PESQUISA BRASILEIRA (GT – RDP Brasil), 2018. **Planilha de respostas do Questionário Práticas e percepções sobre acesso aberto a dados de pesquisa**. *Survey* – Questionário Primeira Fase. Rede de Dados de Pesquisa Brasileira (RDP Brasil). Disponível em: <https://cedap.ufrgs.br/xmlui/handle/20.500.11959/161>. Acesso em: 22 dez. 2022.

GUÐJÓNSDÓTTIR, Rosá. *Personas and Scenarios in Use*. Tese de Doutorado apresentada à Kungliga Tekniska Högskolan em cumprimento do requisito para o Grau de Doutor em Interação Humano-Computador. Estocolmo, Suécia, 2010. Disponível em: https://www.researchgate.net/publication/282326654_Personas_and_scenarios_in_use. Acesso em: 03 out. 2020.

HAEFFNER, Cristina; ZANOTTO, Sônia Regina; GUIMARÃES, Jorge Almeida. *Internacionalização da Universidade Brasileira. Desafios e Perspectivas na Busca pelo Padrão de*

Universidade de Classe Mundial. **Revista Brasileira de Pós-Graduação**, v. 17, n. 37, p. 1 – 28, 2021. Disponível em: <https://doi.org/10.21713/rbpg.v17i37.1797>. Acesso em: 21 set. 2021.

HAMPTON, Stephanie E. *et al.* **Big data and the future of ecology**. *Frontiers in Ecology and the Environment*, v. 11, n. 3, p. 156 – 162, 2013. Disponível em: <https://doi.org/10.1890/120103>. Acesso em: 05 ago. 2021.

HENRIQUES, Márcio Simeone. Dimensões dos públicos nos processos de comunicação pública: formas de conhecimento, ação e organização. In: SCROFERNEKER, Cleusa Maria Andrade; AMORIM, Lidiane Ramirez de. **(Re) leituras contemporâneas sobre comunicação organizacional e relações públicas** [recurso eletrônico]. Porto Alegre : EDIPUCRS, 2017. Disponível em: <http://www.pucrs.br/edipucrs/>. Acesso em: 03 nov. 2022

HERNÁNDEZ SALAZAR, Patricia (coord.). *et al.* **Seminario Latino-Americano sobre formación de usuarios de La información y los estudios de usuarios**. México: UNAM, Centro de Investigaciones Bibliotecológicas, 1997.

HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (orgs). **O quarto paradigma: descobertas científicas na era da eScience**. 1947. Disponível em <http://ofitexto.arquivos.s3.amazonaws.com/Deg-4Paradigma.pdf>

HIRSCHMAN, Albert. **Exit, Voice and Loyalty: responses to decline in firms, organizations and states**. Cambridge, Mass: Harvard University Press, 1970.

HUNT, James R.; BALDOCCHI, Dennis D.; VAN INGEN, Catharine. **Redefinição da Ciência Ecológica com o uso de dados**. In: HEY, T.; TANSLEY, S.; TOLLE, K. (Orgs.). O quarto paradigma: descobertas científicas na era da e-Science. São Paulo: Oficina de textos, 2011. p. 47 – 51.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Quem somos** [on-line], 2021. Disponível em: <https://www.ibict.br/sobre-o-ibict-1>. Acesso em: 21 jul.2021.

_____. **Repositórios Digitais**, [on-line], 2018a. Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais>. Acesso em: 16 ago. 2018.

_____. **Ibict lança Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã**, [on-line], 2016. Sala de Imprensa. 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: 16 ago. 2018.

_____. **Plano de Trabalho**, 2018b. Disponível em: <http://www.ibict.br/cooperacao-tecnico-cientifica/cooperacao-nacional/PlanodeTrabalho assinado.pdf>. Acesso em: 21 ago. 2018.

ISO. ISO 13407:1999 **Human-centered design processes for interactive systems**. Geneve: International Organization for Standardization, 1999.

KOHLER, Robert E. **Landscape and labscapes: Exploring the lab-field border in biology**. Chicago, IL: University of Chicago Press, 2010.

KREBS, Luciana; CAREGNATO, Sônia. Implicações do Acesso Aberto aos Dados de Pesquisa: questões em debate. **Pôster** – XVIII Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB 2017. Disponível em:

<http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/viewFile/191/572>.

Acesso em: 20 ago. 2018.

KUNSCH, Margarida Maria Krohling. **Planejamento de relações públicas na comunicação integrada**. Summus editorial, 1986.

LAKATOS, Eva; MARCONI, Marina. **Metodologia do Trabalho Científico**: procedimentos básicos, pesquisa bibliográfica, projeto, relatório, publicações e trabalhos científicos. 4ª ed. São Paulo : Atlas, 1992.

LARA, Marilda; CONTI, Vivaldo. **Disseminação da Informação e Usuários**. São Paulo em Perspectiva, 17(3 – 4): 26 – 34, 2003.

LEITE, Fernando C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto, Brasília : Ibict, 2009.

LEROUGE, Cynthia *et al.* *User profiles and personas in the design and development consumer health technologies*. *International Journal of Medical Informatics*. v. 82, n. 11, p. 251 – 268, Nov. 2013. *On-line*. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S1386505611000724?via%3Dihub>. Acesso em: 13 nov. 2020.

LIMA, Clóvis R. M. de. **Estudos de Usuários de Sistemas de Informação: contribuição metodológica da epidemiologia**. Revista Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) 165 – 173, jul./dez, 1989. Disponível em:

<http://www.brapci.inf.br/index.php/article/download/1555>. Acesso em: 09 nov. 2022.

LIMA, Fábيا Pereira. Públicos organizacionais pela perspectiva do acontecimento: algumas implicações teóricas e práticas para as Relações Públicas. Trabalho apresentado no GP Relações Públicas e Comunicação Organizacional do XV Encontro dos Grupos de Pesquisas em Comunicação, evento componente do XXXVIII Congresso Brasileiro de Ciências da Comunicação. **Intercom**, Rio de Janeiro, RJ – 4 a 7 de setembro de 2015.

LISBOA, Ândlei. **Por que criar Personas?** 2017. Disponível em: <https://brasil.uxdesign.cc/por-que-criar-personas-bc796a1ffc7e>. Acesso em: 24 ago. 2020.

LONG-TERM ECOLOGICAL RESEARCH (LTER). **Website**. *Managed by LTER Network Office, NCEAS, UCSB*, Santa Barbara, 2021. Disponível em: <https://lternet.edu/about/>. Acesso em: 12 ago. 2021.

LUIZÃO, Flávio *et al.* Impactos antrópicos no Ecossistema de Floresta Tropical, 2013, p. 57 – 86. In: TABARELLI, Marcelo *et al.* PELD-CNPq dez anos do Programa de Pesquisas Ecológicas de Longa Duração do Brasil: achados, lições e perspectivas. Ed. Universidade da UFPE, 2013.

MACDONALD, Steven. *How to Create a Customer Centric Strategy For Your Business*, 2018. Disponível em: <https://www.superoffice.com/blog/how-to-create-a-customer-centric-strategy/>. Acesso em: 08 ago. 2018.

MADSEN, Adrian *et al.* *Personas as a powerful methodology to design targeted professional development resources*, 2014. *On-line*. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1408/1408.1125.pdf>. Acesso em: 22 dez. 2022.

MALACARNE, Vilmar; CHRISTOFOLETTI, João F.; STRIEDER Dulce M. **Ética, ciência e educação**: caminhos para uma sociedade humana e tecnologicada e o papel da escola. *Revista Educação e Cultura Contemporânea*, v. 16, n. 42, 2019. DOI 10.5935/2238-1279.20190036. p. 252 – 271.

MAPBIOMAS. Projeto de Mapeamento Anual do Uso e Cobertura da Terra no Brasil. **Website, on-line**, 2022. Disponível em: <https://mapbiomas.org/>. Acesso em: 24 out. 2022.

MARQUES, Joana Brás Varanda; FREITAS, Denise de. **Método DELPHI**: caracterização e potencialidades na pesquisa em Educação. V. 29, N. 2 (87) | maio/ago. 2018, p. 389 – 415. *Revista pro. posições*. <http://dx.doi.org/10.1590/1980-6248-2015-0140>. e-ISSN 1980-6248. Disponível em: <http://www.scielo.br/pdf/pp/v29n2/0103-7307-pp-29-2-0389.pdf>. Acesso em: 06 abr. 2020.

MARTINS, Marina Rodrigues. Mapeamento de Públicos Estratégicos para iniciativas brasileiras universitárias de arquivamento da web no âmbito acadêmico: uma projeção para Universidade Federal do Rio Grande do Sul e seu Programa de Pós-Graduação em Comunicação, 2019. 125 f. Orientador: Moisés Rockembach. **Dissertação** – Programa de Pós-Graduação em Comunicação e Informação, Universidade Federal do rio Grande do Sul. Disponível em: <https://lume.ufrgs.br/handle/10183/189170> . Acesso em: 22 dez. 2022.

MARTINS, Marina Rodrigues; ROCKEMBACH, Moisés. **Mapeamento de públicos para iniciativas acadêmicas de arquivamento da web**. *Intercom, Rev. Bras. Ciênc. Comun.*, São Paulo, v. 43, n. 1, p. 71 – 88, Jan. 2020 . Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1809-58442020000100071&lng=en&nrm=iso. Acesso em: 16 jun. 2020. Epub Apr 27, 2020. <https://doi.org/10.1590/1809-5844202014>.

MARTINS, Marina R.; VANZ, Samile Andrea de S.. **Construção de Personas: mapeamento de estudos e métodos**. In: **Trajetórias de pesquisa em comunicação**: temas, heurísticas, objetos. Ana Taís Martins (org.). São Paulo: Pimenta Cultural, 2021. 268p.

MARTINS, Marina R.; VANZ, Samile Andréa de Souza. A prática da comunicação científica por meio do compartilhamento e reuso de dados de pesquisa na Ecologia, 2022. **Artigo**. Grupo de pesquisa Comunicação, Divulgação Científica, Saúde e Meio ambiente. 45º Congresso Brasileiro de Ciências da Comunicação. Disponível em: <https://comsaudeambiente.wixsite.com/intercom/2022>. Acesso em: 16 out. 2022.

MATTOS, Alexandre M. **Organização**: uma visão global - introdução-ciência-arte. 2.

ed. Rio de Janeiro: FGV, 1978.

MATRAT, Lucien; CARIN, Alex. **As relações públicas: motor da produtividade**. Lisboa: Livraria San Pedro Editora, s.d.

MARSHALL, Russell *et al.* *Design and evaluation: End users, user datasets and personas*. *Applied Ergonomics*, v. 46, p. 311 – 317, 2015. *On-line*. Disponível em: shorturl.at/dryBX. Acesso em: 13 nov. 2020.

MATTOS, CLG. **A abordagem etnográfica na investigação científica**. In MATTOS, CLG., and CASTRO, PA., orgs. *Etnografia e educação: conceitos e usos [on-line]*. Campina Grande: EDUEPB, 2011. p. 49 – 83. ISBN 978-85-7879-190-2. Disponível em: <http://books.scielo.org>. Acesso em: 17 ago. 2018.

MEDEIROS, Claudia. B. *Gestão de Dados Científicos – da coleta à preservação [on-line]*. *SciELO em Perspectiva*, 2018. Disponível em: <https://blog.scielo.org/blog/2018/06/22/gestao-de-dados-cientificos-da-coleta-a-preservacao/>. Acesso em: 06 ago. 2021.

MIASKIEWICZ, Tomasz; SUMNER, Tamara; KOZAR, Kenneth A. *A latent semantic analysis methodology for the identification and creation of personas*. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2008. p. 1501 – 1510. *On-line*. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/1357054.1357290>. Acesso em: 13 nov. 2020.

MICHENER, William K. *et al.* *Nongeospatial metadata for the ecological sciences*. *Ecological Applications*, v. 7, n. 1, p. 330 – 342, 1997. Disponível em: [https://doi.org/10.1890/1051-0761\(1997\)007\[0330:NMFTEJ\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1997)007[0330:NMFTEJ]2.0.CO;2). Acesso em: 16 jun. 2021.

MICHENER, William K. *Metadata*. In: *Ecological Data: Design, Management and Processing*. Blackwell Science, Oxford, 2000, p. 92–116.

_____. *Meta-information concepts for ecological data management*. *Ecological informatics*, v. 1, n. 1, p. 3 – 7, 2006. Disponível em:

https://www.sciencedirect.com/science/article/pii/S157495410500004X?casa_token=MgfsbbPhdxMAAAAA:-RJ8ozNnF36wBAz7QFrFQizGUOQYq-g_dZaqMXqYtiRfSt5xLpt6ZveC6JG1vjYk8rmwiaU. Acesso em: 16 jun. 2021.

_____. *Ecological data sharing*. *Ecological informatics*, v. 29, p. 33 – 44, 2015. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574954115001004>. Acesso em: 14 jun. 2021.

MICHENER, William K.; JONES, Matthew B. *Ecoinformatics: supporting ecology as a data-intensive science*. *Trends in ecology & evolution*, v. 27, n. 2, p. 85 – 93, 2012. Disponível em: <https://doi.org/10.1016/j.tree.2011.11.016>. Acesso em: 09 ago. 2021.

MÍGUEZ-GONZÁLEZ, María Isabel. *Teoría situacional de los públicos: las nuevas aportaciones desde la década de los noventa*. *Communication & Society*, v. 19, n. 2, p. 133 – 162, 2006. Disponível em: <https://revistas.unav.edu/index.php/communication-and-society/article/view/36309/30763>. Acesso em: 09 nov. 2022.

MÓDOLO, Cristiane Machado. Infográficos: características, conceitos e princípios básicos. In: **Congresso Brasileiro de Ciências da Comunicação da Região Sudeste**, 2007. Disponível em: <http://www.intercom.org.br/papers/regionais/sudeste2007/resumos/r0586-1.pdf>. Acesso em: 10 nov. 2022.

MONFASANI, Rosa Emma; CURZEL, Marcela Fabiana. *Usuários de La información: formación y desafíos*. 1ª ed, Buenos Aires: Alfagrama, 2006.

MOYA, Juan A. Gaitán; RAIGADA, José L. Piñuel. *Técnicas de investigación em comunicación social: elaboración y registro de datos*. Editorial Síntesis, 1998.

MURRAY-RUST, Peter. *Open data in science*. *Nature Precedings*, p. 1, 2008. Disponível em: <https://www.nature.com/articles/npre.2008.1526.1>. Acesso em: 26aou. 2020.

NATIONAL CENTER FOR ECOLOGICAL ANALYSIS AND SYNTHESIS (NCEAS). *Our history*, 2021. **Website**. Disponível em: <https://www.nceas.ucsb.edu/our-history>. Acesso em: 11 ago. 2021.

NATIONAL ECOLOGICAL OBSERVATORY NETWORK (NEON). **Website**, 2021. Disponível em: <https://www.neonscience.org/>. Acesso em: 14 jun. 2021

NATIONAL SCIENCE BOARD (NSB). *Digital Research Data Sharing and Management*. *National Science Foundation (NSF)*, 2011. Disponível em: <https://www.nsf.gov/nsb/publications/2011/nsb01211.pdf>. Acesso em: 15 jun. 2021.

NIELSEN, Lene. *Engaging personas and narrative scenarios*. *PhD series*, v. 17, 2004.

NIELSEN, Lene; STORGAARD HANSEN, Kira. *Personas is applicable: a study on the use of personas in Denmark*. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014. p. 1665 – 1674. *On-line*. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/2556288.2557080>. Acesso em: 13 nov. 2020.

NIELSEN, Lene. 30. *Personas*. In: *The Encyclopedia of Human-Computer Interaction*, 2 ed. [S.l.] *Interaction Design Foundation*, 2018. *On-line*. Disponível em: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/personas>. Acesso em: 13 nov. 2020.

O'BRIEN, James A. **Sistemas de informação e as decisões gerenciais na era da internet**. 2. ed. São Paulo: Saraiva, 2004. 431 p.

OCEAN BIODIVERSITY INFORMATION SYSTEM (OBIS). **Website**, 2020. Disponível em: <https://www.obis.org/>. Acesso em: 14 jun. 2021

OCEAN OBSERVATORIES INITIATIVE (OOI). **Website**, 2021. Disponível em: <https://oceanobservatories.org/>. Acesso em: 14 jun. 2021

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **OECD principles guidelines for access to research data from public funding**. França: OECD, 2007. Disponível em: <https://www.oecd.org/sti/sci-tech/38500813.pdf>. Acesso em: 12 abr. 2019.

OSAWA, Takeshi.. *Perspectives on biodiversity informatics for ecology*. **Ecological Research**, v. 34, n. 4, p. 446 – 456, 2019. Disponível em: <https://doi.org/10.1111/1440-1703.12023>. Acesso em: 24 mai. 2022.

PARHAM, S. W.; BODNAR, J.; FUCHS, S.. **Supporting Tomorrow's Research: assessing faculty data curation needs at Georgia Tech**. *College Research Libraries News*, p. 10 – 13, jan. 2012. Disponível em: <https://crln.acrl.org/index.php/crlnews/article/view/8686/9159>. Acesso em: 20 ago. 2018.

PAVÃO, Caterina Marta Groposo; GABRIEL JÚNIOR, Rene Faustino; ROCHA, Rafael Port da. Proposta de criação de uma rede de dados abertos da pesquisa brasileira. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 16, n. 2, p. 329–343, 2018. DOI: 10.20396/rdbci.v16i2.8651180. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8651180>. Acesso em: 23 dez. 2022.

PAVÃO, Caterina Groposo *et al.* **Acesso aberto a dados de pesquisa no Brasil: repositórios brasileiros de dados de pesquisa: relatório 2018**. - Porto Alegre, RS : UFRGS, 2018a. 19 p. Disponível em: <http://hdl.handle.net/10183/185138>. Acesso em: 22 dez. 2022

PENNOCK, Maureen. *Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information*. **Library & Archives**, Janeiro 2007. Disponível em: http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf. Acesso em: 27 mar. 2020.

PETERS, Debra PC *et al.* *Taking the pulse of a continent: expanding site-based research infra structure for regional-to continental-scale ecology*. **Ecosphere**, v. 5, n. 3, p. 1 – 23, 2014. Disponível em: <https://esajournals.on-linelibrary.wiley.com/doi/full/10.1890/ES13-00295.1>. Acesso em: 12 jun. 2021.

PINHEIRO, Thiago Moreira; SZANIECKI, Barbara; MONAT, André. **O uso da metodologia de personas na produção de hiperídia adaptativa para visitas guiadas a museus**. *Da Pesquisa – UDESC*, v. 11, n. 15, 2016. DOI: <https://doi.org/10.5965/1808312911152016255>. Disponível em: https://www.researchgate.net/publication/301830408_O_USO_DA_METODOLOGIA_DE_PERSONAS_NA_PRODUCAO_DE_HIPERMIDIA_ADAPTATIVA_PARA_VISITAS_GUIADAS_A_MUSEUS. Acesso em: 23 mar. 2020.

POLIVANOV, Beatriz. *Etnografia Virtual, Netnografia ou Apenas Etnografia? Implicações dos Termos em Pesquisas Qualitativas na Internet*. Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação XXXVI Congresso Brasileiro de Ciências da Comunicação – Manaus, AM – 4 a 7/9/2013.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBio). **Website**, *on-line*, 2022. Disponível em: <https://ppbio.inpa.gov.br/>. Acesso em: 24 out. 2022.

PRUITT, John; ADLIN, Tamara. *The persona lifecycle: keeping people in mind through out product design*. San Francisco, CA. Elsevier, 2006.

PRUITT, John; GRUDIN, Jonathan. *Personas: practice and theory. Proceedings of the 2003 conference on Designing for user experiences*. 2003. p. 1 – 15. *On-line*. Disponível em: <https://www.microsoft.com/en-us/research/wp-content/uploads/2017/03/pruitt-grudinold.pdf>. Acesso em: 13 nov. 2020.

RAUTENBERG, Sandro; BURDA, Alessandra Cassiana; SOUZA, Lucélia de Souza. Um *workflow* para compartilhamento de dados científicos primários baseado em dados abertos conectados. **Encontros Bibli**, 01 Setembro 2018, Vol.23(53), pp.110 – 123. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p110/37293>. Acesso em: 20 mar. 2020.

REDE DE DADOS DE PESQUISA (RDP BRASIL). Citação de dados em poucas palavras: identificar, acessar, reutilizar, citar. **Infográfico**. Aprendendo sobre Dados de Pesquisa, n. 2, 2023.

REDE NACIONAL DE ENSINO E PESQUISA (RNP). **Website**.

Quem somos. [*on-line*]. Disponível em: <https://www.rnp.br/sobre>. Acesso em: 21 jul. 2021.

_____. Participe da consulta para mapear o uso de dados abertos para pesquisa, 16 de março de 2018. Disponível em: <https://www.rnp.br/noticias/participe-consulta-mapear-uso-dados-abertos-pesquisa>. Acesso em: 31 jul. 2018.

_____. RNP e Ibict lançam chamada para promover o acesso aberto a dados de pesquisa, 15 de março de 2017. Disponível em: <https://www.rnp.br/noticias/rnp-e-ibict-lancam-chamada-promover-acesso-aberto-dados-pesquisa>. Acesso em: 31 jul. 2018.

_____. Repositórios da Rede Nacional de Ensino e Pesquisa – RNP Brasil. Questionários. Disponível em: <https://dadosabertos.rnp.br/dataverse/rdp>. Acesso em: 22 ago. 2020.

_____. RNP participa de rede nacional para compartilhamento de dados e ciência aberta Pesquisa e Desenvolvimento – 01 de abril de 2021a. Disponível em: <https://www.rnp.br/noticias/rnp-participa-de-rede-nacional-para-compartilhamento-de-dados-e-ciencia-aberta>. Acesso em: 28 jul. 2021.

_____. RNP, CNPq e Ibict celebram acordo tripartite para ciência aberta. *On-line*. 15 de julho de 2021b. Disponível em: <https://www.rnp.br/noticias/rnp-cnpq-e-ibict-celebram-acordo-tripartite-para-ciencia-aberta>. Acesso em: 28 jul.2021.

REICHMAN, O. James; JONES, Matthew B.; SCHILDHAUER, Mark P. *Challenges and opportunities of open data in ecology*. **Science**, v. 331, n. 6018, p. 703 – 705, 2011.

REISSWITZ, Flávia. **Análise de Sistemas** Vol 1: tecnologias e sistemas de informações, 2008.

RESEARCH DATA ALLIANCE (RDA). *About RDA*. **Website**, 2022. Disponível em: <https://www.rd-alliance.org/about-rda>. Acesso em: 23 dez. 2022.

ROCHA, Rafael Port da *et al.* **Acesso aberto a dados de pesquisa no Brasil : soluções tecnológicas**: relatório 2018. - Porto Alegre, RS : UFRGS, 2018. 75p. Disponível em: <https://lume.ufrgs.br/handle/10183/185126>. Acesso em: 23 dez. 2022.

RODRIGUES, Márcio Urel. Curso Avançado de Análise de Conteúdo em Pesquisas Qualitativas. **Vídeoaulas**, 2021. Grupo de Estudos e Pesquisas em Educação Matemática nas Escolas (GPEME). Universidade do Estado de Mato Grosso (UNEMAT). Disponível em <https://youtu.be/XtF75xtvByY>. Acesso em: 30 ago. 2022.

RODRIGUES, Virginia. **Estudo de Usuários da Informação como Instrumento para Avaliação de Serviços em Unidades de Informação**: estudo de caso. Dissertação – Mestrado Profissional em Sistemas de Informação e Gestão do Conhecimento, Universidade FUMEC, 2017.

ROSSI, Gabriel. **Sociografia: um novo conceito de análise para as marcas na internet**. 20 de março de 2012. Disponível em: <https://goo.gl/jV9FCF>. Acesso em: 06 ago. 2018.

ROTH, Wolff-Michael; BOWEN, G. Michael. *Creative Solutions and Fibbing Results Enculturation in Field Ecology*. *Social Studies of Science* 31, no. 4, p. 533 – 556, 2001a. _____. *Of disciplined minds and disciplined bodies: On becoming an ecologist*. *Qualitative Sociology* 24 (4), p. 459 – 81, 2001b.

ROZADOS, Helen; PIFFER, Bárbara. **Pesquisa de Marketing e estudos de usuário**: um paralelo entre os dois processos. Em *Questão*, Porto Alegre, v. 15, n. 2, p. 169 – 182, jul./dez. 2009.

SAGARIN, Raphael; PAUCHARD, Aníbal. *Observational approaches in ecology open new ground in a changing world*. *Frontiers in Ecology and the Environment*, v. 8, n. 7, p. 379 – 386, 2010. Disponível em: <https://doi.org/10.1890/090001>. Acesso em: 05 ago.2021

SANT’ANA, Ricardo César Gonçalves. Ciclo de vida dos dados e o papel da ciência da informação. **Encontro Nacional de Pesquisa em Ciência da Informação**, GT8 Informação e Tecnologia, 14, 2013, Florianópolis. Anais [...]. Florianópolis, 2013. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/185240>. Acesso em: 22 dez. 2022

SANT’ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, Londrina, v. 21, n. 2, p. 116–142, maio/ago. 2016. Disponível em: <https://brapci.inf.br/index.php/res/v/35252>. Acesso em: 22 dez. 2022.

SANTOS, Paulo J. Parreira; FREITAS, Leandro; MELO, Adriano Sanches. **Documento de Área: Área 07: Biodiversidade**. Ministério da Educação, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Diretoria de Avaliação, 2019. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/biodiversidade-pdf>. Acesso em: 28 out. 2022.

SANTOS, Paula Xavier *et al.* **Livro Verde – Ciência aberta e dados abertos**: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. 2017. Disponível em: <https://www.arca.fiocruz.br/handle/icict/24117>. Acesso em: 17 jun. 2022.

SANTOS, Paulo J. Parreira; FREITAS, Loreta Brandão de; FREITAS, Leandro. **Documento de Área: Biodiversidade**. Avaliação Quadrienal. Ministério da Educação, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Diretoria de Avaliação, 2016. Disponível em: https://posecologia.ib.usp.br/images/CAPES/07_BIOD_docarea_2016.pdf. Acesso em: 28 out. 2022

SANTOS, Gildenir C; RIBEIRO, Célia M. **Acrônimos, Siglas e Termos Técnicos**: arquivística, biblioteconomia, documentação, informática. Campinas, SP: Editora Átomo, 2003.

SAYÃO, Luis Fernando; SALES, Luana Farias. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **RECIIS – Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 8, n. 2, p. 76 – 92, jun. 2014.

_____. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://www.icb.usp.br/~sbibicb/images/guia%20gestaoPDF/Guia%20de%20gestao%20dados%20de%20pesquisa.pdf>. Acesso em: 20 jul. 2021.

SHEDROFF, Nathan. **Information Interaction design: a unified Field theory of design**, 1994. Disponível em:

http://www.nomads.usp.br/documentos/textos/design_interfaces_computacionais/info_interac_de_sign_unified_nathan.pdf. Acesso em: 14 ago. 2018.

SIDORKO, Nadezhda. *Personas in user-centered design of university websites*. 2019. **Master's Degree (Programme in Human Technology Interaction) – Faculty of Communication Sciences, University of Tampere**, Finlândia, 2019. Disponível em:

<https://trepo.tuni.fi/bitstream/handle/10024/105804/SidorkoNadezhda2019.pdf?sequence=1&isAllowed=y>. Acesso em: 13 nov. 2020.

SILVA, Fabiano Couto Corrêa da Silva. O que são os dados científicos. In: SILVA, Fabiano Couto Corrêa da Silva. **Gestão de dados científicos**. Rio de Janeiro: Editora Interciência, 2019. p. 21 – 25.

SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia da. O ecossistema da Ciência Aberta. **Transinformação**, v. 31, 2019. Disponível em: <http://dx.doi.org/10.1590/2318-0889201931e190001>. Acesso em: 30 jul. 2021.

SILVA, Maria. Helena. F. X. da; SALES, Luana. F.; SAYÃO, Luis F.; DRUMOND, Geisa M.; MARANHÃO, Ana Maria N. Competências dos bibliotecários na gestão dos dados de pesquisa. **Ciência Da Informação**, Brasília, DF, v.48 n.3 (Supl.), p. 303 – 313, set./dez. 2019, 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/4973>. Acesso em: 26 jul. 2021.

SIMÕES, Roberto Porto. **Relações públicas**: função política. 3. ed. São Paulo: Summus, 1995.

SIMÕES, Roberto Porto. **Relações Públicas e Micropolítica**. São Paulo: Summus, 2001.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA (SiBBr).

Conjunto de Dados. Disponível em:

https://collectory.sibbr.gov.br/collectory/datasets?lang=pt_BR. Acesso em: 18 dez. 2022.

SOARES, Marisa de Carvalho; LOMANDO, Nadege Gonçalves. Públicos: o diferencial estratégico dos planejamentos de Relações públicas. In: S. Dornelles (org). **Relações públicas: construindo relacionamentos estratégicos**, p. 31 – 42, 2010. Porto Alegre: Edipucrs.

SOUSA, Milene R. de; BERTOMEU, João V. C. *UX Design* na criação e desenvolvimento de aplicativos digitais. **Informática na educação: teoria & prática**. v. 18, n. 2, 2015. e-ISSN: 1982 - 1654. DOI: https://doi.org/10.22456/1982_1654.54897. Disponível em:

<https://seer.ufrgs.br/InfEducTeoriaPratica/article/view/54897>. Acesso em: 17 abr. 2020.

SPECIES LINK. **Website**. Disponível em: <https://specieslink.net/>. Acesso em: 18 dez. 2022.

STM ASSOC.ORG. **Website**. *Text and Data Mining for non-commercial scientific research*.

Disponível em: <https://www.stm->

[assoc.org/2013_11_11_Text_and_Data_Mining_Declaration.pdf](https://www.stm-assoc.org/2013_11_11_Text_and_Data_Mining_Declaration.pdf). Acesso em 31 jul. 2018.

STURMER, Adriana *et al.* Comunicação Estratégica: Públicos de interesse e Persona Organizacional. **Revista Visão: Gestão Organizacional**, p. 25 – 43, 2019a. Disponível em:

<http://dx.doi.org/10.33362/visao.v8i1.1728>. Acesso em: 13 nov. 2020.

STURMER, Adriana *et al.* Mapa da Persona Organizacional como estratégia de mapeamento de públicos em relações públicas. **Temática**, Paraíba, Ano XV, n. 6, p. 131. Junho, 2019b.

Disponível em: <http://periodicos.ufpb.br/ojs2/index.php/tematica>, 2019b. Acesso em: 13 nov. 2020.

TABARELLI, Marcelo *et al.* **PELD-CNPq dez anos do Programa de Pesquisas Ecológicas de Longa Duração do Brasil: achados, lições e perspectivas**. Ed. Universidade da UFPE, 2013.

Disponível em: <http://hdl.handle.net/10183/104869>. Acesso em: 03 ago. 2021.

TARAPANOFF, Kira. *et al.* **Técnicas para tomada de decisão nos Sistemas de Informação**. Brasília : Thesaurus, 1995.

THE GLOBAL VOICE OF SCHOLARLY PUBLISHING. **Website**. *Brussels Declaration*.

Disponível em: <https://www.stm-assoc.org/public-affairs/resources/brussels-declaration/>. Acesso em: 31 jul. 2020.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS). Guia de normalização de trabalhos acadêmicos da Faculdade de Biblioteconomia e Comunicação da UFRGS

[**manuscrito**], 2019. Disponível em: <https://www.ufrgs.br/bibfbc/ferramentas/manual-de-normalizacao/>. Acesso em: 14 dez. 2022.

VANDERBILT, Kristin L. *et al.* *Fostering ecological data sharing: collaborations in the International Long Term Ecological Research Network*. **Ecosphere**, v. 6, n. 10, p. 1 – 18, 2015.

VANZ, Samile Andrea de Souza *et al.* **Acesso aberto a dados de pesquisa no Brasil: práticas e percepções dos pesquisadores: relatório 2018.** Disponível em: <http://hdl.handle.net/10183/185195>. Acesso em: 23 dez. 2022.

VASARA, Kristoffer. **Introducing personas in a software project.** 2003. *Master's thesis (Department of Computer Science and Engineering) – Helsinki University of Technology, Helsinki*, 2003. Disponível em: <https://pdfs.semanticscholar.org/bccd/0ae4162c9be64b532f87d0386c00b2f7e399.pdf> . Acesso em: 13 nov. 2020.

VEIGA, Viviane; MACHADO, Rejane; JORGE, Vanessa de Arruda; PENEDO, Erick. Compartilhamento de dados de pesquisa na Fiocruz: diagnóstico e percepção do pesquisador. **Ciência da Informação**, Brasília, DF, v.48 n.3 (Supl.), p.319 – 327, set./dez. 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/136386>. Acesso em: 23 jun. 2021.

VINUTO, Juliana. A amostragem em bola de neve na pesquisa qualitativa: um debate em aberto. **Temáticas**, v. 22, n. 44, 2014. Disponível em: <https://econtents.bc.unicamp.br/inpec/index.php/tematicas/article/view/10977>. Acesso em: 05 nov. 2020.

ZIMMERMAN, Ann S. *New knowledge from old data: The role of standards in the sharing and reuse of ecological data.* **Science, technology, & human values**, v. 33, n. 5, p. 631 – 652, 2008. Disponível em: <https://doi.org/10.1177/0162243907306704>. Acesso em: 04 ago. 2021.

APÊNDICE A – Roteiro de Entrevista Semiestruturada Partes Interessadas

Olá,

Como vai?

Sou Marina Rodrigues Martins, doutoranda em Comunicação pelo Programa de Pós-Graduação em Comunicação, da Universidade Federal do Rio Grande do Sul. Obrigada por colaborar com a pesquisa “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa”.

Nossa entrevista tem duração prevista de 30 minutos a 1 hora e será gravada (vídeo e áudio) pela plataforma “X”. Temos um roteiro prévio, dividido em dois blocos, que visa levantar informações sobre as variáveis que influenciam no objetivo principal do estudo.

Você tem alguma dúvida ou podemos prosseguir?

O primeiro bloco é bem objetivo e rápido. Visa descobrir sobre os aspectos técnicos e administrativos que influenciam na efetividade e no aperfeiçoamento das iniciativas brasileiras de acesso aberto aos dados de pesquisa. Para tanto, pedimos para que você expresse suas ideias sobre as seguintes variáveis:

1. Nome completo:
2. Instituição em que atua:
3. Qual sua função/cargo na iniciativa brasileira de AADD?
4. Qual cidade reside atualmente:
5. Idade:
6. Gênero:
7. Em que regiões do Brasil ou do mundo que desenvolve sua pesquisa:
8. Para ti, enquanto parte interessada em promover a Iniciativa Brasileira de Acesso Aberto a Dados de Pesquisa, o que é fundamental saber sobre o usuário?
9. Qual imagem que você tem da iniciativa?
10. O usuário influencia de algum modo na iniciativa? Caso sim, como? Caso não, você acredita que deveria influenciar? Por quê?

11. O que você pode falar sobre expectativas, objetivos, capacidades, limitações do contexto técnico e orçamentário em torno da iniciativa (forças, fraquezas, oportunidades e ameaças)?

Obrigada, encerramos o primeiro bloco. Podemos prosseguir?

No segundo bloco utilizaremos um pouco mais de tempo. Nesse momento queremos mapear qual a imagem que você possui do usuário (real ou em potencial) para iniciativas brasileiras de acesso aberto aos dados de pesquisa. Para tanto pedimos que discorra sobre as seguintes variáveis:

1. Considerando que o usuário é da área de Ciências Biológicas – ecologia, se você fosse elaborar uma lista de palavras-chave utilizada pelo usuário real ou em potencial no dia-a-dia de pesquisa dele. Quais seriam estas palavras? (De 3 a 5 palavras);
2. Na sua visão, o que esses usuários compreendem por acesso aberto aos dados de pesquisa?
3. Será que eles promovem a ideia e como ocorre a prática do uso, compartilhamento e reuso de dados científicos na área desse usuário?
4. Como/por que meios praticam o compartilhamento e o acesso aberto? (Repositórios indicados na survey GT – RDP Brasil (2018): *Ocean Biogeographic Information System (OBIS)*, a *Global Biodiversity Information Facility (GBIF)* e a *speciesLink*)
5. Conforme sua experiência quais são as características dos dados compartilhados pelos pesquisadores brasileiros das Ciências Sociais Biológicas – ecologia (Que tipo de dado – observacionais, experimentais, computacionais? O dado é volumoso? São arquivos pesados? Quais são os formatos? São digitais, com origem em sensores, radares ou satélites? Dados históricos? Dados de campo? Dados textuais? altitude, longitude, temperatura, níveis de relevo, tipos de terreno, imagens geográfica? Georreferenciados – conjuntos de dados, informações biológicas e espécies de seres vivos, a locais geográficos através de nome de lugares ou de coordenadas de latitude e longitude?)
6. Você conhece ou já ouviu falar da *Ecological Metadata Language (EML)*, linguagem própria para descrever dados ecológicos? O padrão *EML* foi criado por pesquisadores da ecologia, do *National Center for Ecological Analysis and Synthesis (NCEAS)* e da rede *International Long-Term Ecological Research (ILTER)* no início dos anos 80. Ele visa à catalogação histórica de dados de natureza ecológica com ênfase nos aspectos: geográfico, data e hora, taxonomia, metodologia, dados gerais e específicos

7. Quais as expectativas, objetivos, capacidades, limitações do contexto técnico e orçamentário em torno das iniciativas brasileiras de acesso aberto aos dados de pesquisa que esse usuário possui?
8. No seu ponto de vista qual imagem que o usuário tem do serviço das iniciativas brasileiras de acesso aberto aos dados de pesquisa?
9. Onde, quando, por quê e como o serviço é ou será usado por esse usuário?
10. Na sua visão quais funções a iniciativa brasileira deve exercer na vida do usuário (pessoal e profissional)?
11. Quais são as metas e motivações para uso/consumo por parte desse usuário?
12. No seu ponto de vista existem problemas, insatisfações, frustrações com serviços atuais de dados de pesquisa (similares à proposta de acesso aberto aos dados de pesquisa);
13. Qual o processo de decisão do usuário em relação à utilização de dados abertos, na sua visão?
14. Como você imagina que é a estrutura das relações sociais do usuário, será que ela influencia no uso do serviço?
15. Como você imagina que é a rotina do dia a dia de pesquisa, locais e ambientes de uso do serviço e de dados de pesquisa por parte do usuário das Ciências Biológicas – ecologia?
16. Será que as revistas que esses pesquisadores costumam enviar artigos para publicação exigem armazenamento de dados em repositórios de dados de pesquisa? Caso sim, quais repositórios?
17. Na sua opinião, os pesquisadores brasileiros das Ciências Biológicas – ecologia têm expectativa que suas próprias instituições ofereçam o serviço de acesso aberto ou preferem utilizar um repositório internacional?
18. A partir da imagem que você possui de pesquisadores das Ciências Biológicas – ecologia, peça para que descrevas as necessidades e características da personalidade desse usuário;
19. Como você acha que a iniciativa brasileira de dados abertos de pesquisa afeta ou afetará a vida desse usuário;
20. Você imagina que esse usuário pesquisador fará o *upload* dos seus dados na plataforma de acesso aberto ou ele prefere que a instituição/iniciativa tenha um funcionário para executar esse processo? Justifique?

21. Será que esse usuário necessita de apoio de profissionais para gestão dos dados de pesquisa ou ele se dispõe e está preparado para executar essa tarefa sem dificuldades (técnicas e de tempo)?
22. Conforme seu conhecimento sobre pesquisadores da ecologia, quais são as principais necessidades e características da personalidade desse usuário?
23. Na sua opinião, quais são os tipos de conteúdo de interesse do usuário (além do assunto dados abertos de pesquisa), onde e em qual formato?
24. Quais outros interesses em geral esse usuário possui? (além da pesquisa)
25. Quais as variáveis demográficas que descrevem esse usuário na sua ideia? Cite idade; etnia; naturalidade; sexo; gênero; estado civil; renda; nível de escolaridade; onde reside, trabalha, estuda; profissão/cargo/ocupação/emprego; habilidades.

Obrigada! Chegamos ao fim de nossa entrevista. Você gostaria de acrescentar algo mais em suas colocações sobre o assunto?

Mais uma vez agradecemos sua colaboração. Assim que tivermos resultados publicados encaminhamos para conhecimento, caso for de seu interesse.

APÊNDICE B – Termo de Consentimento Livre e Esclarecido Partes Interessadas

Observação: o presente TCLE foi enviado para os contatos selecionados via e-mail e ferramenta Google Formulários (*Google Forms*).

Prezado(a) Senhor(a),

Dirigimo-nos ao (à) Senhor (a) para convidá-lo (a) a participar da pesquisa intitulada “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa”. Seu contato foi selecionado a partir da sua atuação como técnico (a)/pesquisador (a) do/indicado pelo Grupo de Trabalho – Rede de Dados de Pesquisa Brasileira. Desse modo compreende-se que você possui algum tipo de autoridade e/ou responsabilidade sobre a iniciativa brasileira de dados abertos de pesquisa.

Sua colaboração no presente estudo é muito importante, mas a decisão em participar é sua. Para tanto, leia atentamente as informações abaixo e não se apresse em decidir. Se você concordar em participar basta preencher esse termo. A coleta de dados ocorrerá por meio de entrevista semiestruturada com roteiro prévio que observa variáveis que visam atingir o objetivo principal do estudo. As entrevistas serão aplicadas de modo síncrono em horário e data pré-agendados entre pesquisador e pesquisado, por intermédio de plataforma de videoconferência MS Teams UFRGS com a possibilidade de gravar áudio e vídeo. As entrevistas estão programadas para durar de 30 minutos até 1 hora. As respostas serão transcritas, organizadas e analisadas pela pesquisadora do projeto. Todos os procedimentos para a garantia da confidencialidade aos participantes serão observados, procurando evitar descrever informações que possam lhe comprometer. O benefício esperado ao compreender “Quem são as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa?” é contribuir para a promoção da ciência aberta no Brasil. Ao participar dessa pesquisa, você não terá nenhum benefício direto; entretanto, esperamos que futuramente os resultados desse estudo sejam usados em benefício de outras pessoas e à ciência brasileira. Todas as informações coletadas nessa investigação são estritamente confidenciais. Acima de tudo interessam os dados coletivos e não aspectos particulares de cada entrevistado. A participação nessa pesquisa não traz complicações legais de nenhuma ordem e os

procedimentos utilizados obedecem aos critérios da ética na Pesquisa com Seres Humanos conforme a Resolução nº 510, de 07 de abril de 2016, do Conselho Nacional de Saúde. Nenhum dos procedimentos utilizados oferece riscos à sua dignidade. O risco que você pode correr ao realizar a pesquisa é de ser identificado mesmo com todos os cuidados de sigilo adotados. Se você não concordar em participar ou quiser desistir em qualquer momento, isso não causará nenhum prejuízo a você. Você não terá nenhum tipo de despesa por participar desse estudo, bem como não receberá nenhum tipo de pagamento por sua participação.

Após as explicações, solicitamos seu consentimento de forma livre e esclarecida para ser um (a) participante da pesquisa. Para tanto, acesse e preencha o formulário Consentimento Livre e Esclarecido. Desde já agradecemos a disponibilidade.

Pergunta 1 – *Você acha que está suficientemente informado(a) a respeito da pesquisa “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa” e concorda de livre e espontânea vontade em participar como informante?*

Sim Não

Pergunta 2 – *Para você, qual o horário adequado para realizar a entrevista?*

Das 9h às 10h Das 10h às 11h Das 11h às 12h

Das 14h às 15h Das 15h às 16h Das 16h às 17h

Pergunta 3 – *Para você, qual o dia da semana adequado para realizar a entrevista?*

segunda-feira terça-feira quarta-feira quinta-feira sexta-feira sábado

Pergunta 4 – *Nome completo:*

Pergunta 5 – *Confirmação de endereço de e-mail para contato:*

Pergunta 6 – *Telefone para contato (whatsapp) com DDD:*

Mensagem de confirmação do Google Forms

Suas respostas foram registradas com sucesso! Agradecemos sua colaboração e colocamo-nos à disposição para explicações adicionais. Caso você possua perguntas sobre o estudo ou se pensar

que pode haver algum prejuízo por sua participação, você pode contatar a pesquisadora responsável Mestra Marina Rodrigues Martins (PPGCOM/UFRGS) pelo e-mail mrodriguesmartins@gmail.com, bem como a orientadora Profa. Dra. Samile Andréa de Souza Vanz pelo e-mail samile.vanz@ufrgs.br ou o Comitê de Ética em Pesquisa – CEP/UFRGS (51) 3308-3738. Horário de atendimento do CEP, das 08:00 às 12:00 e das 13:00 às 17:00.

APÊNDICE C – Roteiro de Entrevista Semiestruturada Usuários

Olá,

Como vai?

Sou Marina Rodrigues Martins, doutoranda em Comunicação pelo Programa de Pós-Graduação em Comunicação, da Universidade Federal do Rio Grande do Sul. Obrigada por colaborar com a pesquisa “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa”.

Nossa entrevista tem duração prevista de 30 minutos até 1 hora e será gravada (vídeo e áudio) pela plataforma “X”. Temos um roteiro prévio que visa levantar informações sobre as variáveis que influenciam no objetivo principal do estudo.

Você tem alguma dúvida ou podemos prosseguir?

Antes de iniciar é importante esclarecemos que essa pesquisa compreende que o acesso aberto aos dados de pesquisa é um movimento que aspira a cooperação e o avanço da ciência no âmbito global, promovendo a ideia e a prática do uso e do compartilhamento de dados científicos dentro de repositórios digitais. E que os repositórios digitais são mais do que simples bases de determinado assunto e finalidade estruturados hierarquicamente por dados. Ou seja, podem ser uma coleção de dados integrados num Sistema de Informação (SI), que são mecanismos facilitadores que por meio de recursos humanos, hardwares, softwares e tecnologias de rede recebem, processam, armazenam, controlam e convertem dados em informação a fim de disseminá-las.

Prosseguindo, neste momento queremos mapear o máximo possível das suas características e da rotina do seu dia a dia a partir de dados qualitativos e quantitativos. Para tanto pedimos que discorresse sobre as seguintes variáveis:

1. Nome completo:
2. Se você fosse eleger uma lista de palavras-chave para seu dia a dia de pesquisa, quais seriam? (Entre 3 e 5 palavras);
3. O que você compreende por acesso aberto aos dados de pesquisa?

4. Você promove a ideia e a prática do uso, compartilhamento e reuso de dados científicos na sua área?
5. Como/por que meios pratica o compartilhamento e o acesso aberto? (Repositórios indicados na survey GT – RDP Brasil (2018): *Ocean Biogeographic Information System (OBIS)*, a *Global Biodiversity Information Facility (GBIF)* e a *speciesLink*)
6. Fale sobre as características dos dados que você compartilha (Que tipo de dado – observacionais, experimentais, computacionais? O dado é volumoso? São arquivos pesados? Quais são os formatos? São digitais, com origem em sensores, radares ou satélites? Dados históricos? Dados de campo? Dados textuais? altitude, longitude, temperatura, níveis de relevo, tipos de terreno, imagens geográfica? Georreferenciados – conjuntos de dados, informações biológicas e espécies de seres vivos, a locais geográficos através de nome de lugares ou de coordenadas de latitude e longitude?)
7. Ao compartilhar os dados você usa a *Ecological Metadata Language (EML)*, linguagem própria para descrever dados ecológicos? O padrão *EML* foi criado por pesquisadores da ecologia, do *National Center for Ecological Analysis and Synthesis (NCEAS)* e da rede *International Long-Term Ecological Research (ILTER)* no início dos anos 80. Ele visa à catalogação histórica de dados de natureza ecológica com ênfase nos aspectos: geográfico, data e hora, taxonomia, metodologia, dados gerais e específicos.
8. Quais são suas expectativas, objetivos, capacidades, limitações do contexto técnico e orçamentário em torno do serviço de uma iniciativa brasileira de dados abertos de pesquisa?
9. Você já ouviu o termo iniciativa brasileira de acesso aberto aos dados de pesquisa?
10. Caso sim – Qual a imagem que você tem das iniciativas brasileiras de acesso aberto aos dados de pesquisa?
11. Caso não – O que você imagina quando pensa no termo?
12. Conforme o seu dia a dia onde, quando, por que e como serviços de acesso aberto aos dados de pesquisa são/serão usados?
13. Quais funções a iniciativa deve exercer ou possui na sua vida (pessoal e profissional)?
14. Quais são as metas e motivações para uso/consumo de um serviço de acesso aberto aos dados de pesquisa?
15. Você possui problemas, insatisfações, frustrações com serviços atuais de dados de pesquisa (similares à proposta de acesso aberto aos dados de pesquisa)? Se sim, pode falar sobre suas experiências? (Peters *et al.* (2014) observou que os ecologistas estão despreparados para lidar com questões de escala regional a continental devido à falta de cultura de compartilhamento de

dados, métodos de coleta não padronizados e desatenção para documentar a proveniência.) (Os diferentes tipos de dados ecológicos tornam sua compreensão complexa, principalmente se reutilizados por outra pessoa, que não participou do levantamento original (DRUCKER, 2011).

16. Discorra brevemente sobre o seu processo de decisão em relação à utilização de um serviço de acesso aberto aos dados de pesquisa;
17. Como é a estrutura das suas relações sociais que influencia no uso do serviço?
18. Relate como é sua rotina do dia a dia de pesquisa, locais e ambientes de uso do serviço e de dados de pesquisa;
19. As revistas que tu costumava enviar artigos para publicação exigem depósito? Caso sim, quais repositórios tu utilizas?
20. Você tem expectativa que sua própria instituição ofereça o serviço de acesso aberto ou prefere utilizar um repositório nacional ou internacional?
21. Você imagina que fará o *upload* dos seus dados na plataforma de acesso aberto ou você prefere que a instituição/iniciativa tenha um funcionário para executar esse processo? Justifique?
22. Você considera o apoio de profissionais para gestão dos dados de pesquisa ou você se dispõe e está preparado para executar essa tarefa sem dificuldades (técnicas e de tempo)?
23. Como a iniciativa brasileira de dados abertos de pesquisa afeta ou afetará a sua vida?
24. Você se imagina fazendo o *upload* dos seus dados numa plataforma de acesso aberto ou imagina que a instituição/iniciativa tenha um funcionário para executar esse processo? Justifique?
25. Quais são suas principais necessidades e características da personalidade?
26. Que tipo de conteúdo é de seu interesse (além do assunto dados abertos de pesquisa), onde e em qual formato?
27. Quais são seus interesses em geral (além da pesquisa)?
28. Quais as variáveis demográficas que o (a) descrevem? Idade; etnia; naturalidade; sexo; gênero; estado civil; renda; nível de escolaridade; onde reside, trabalha, estuda; profissão/cargo/ocupação/emprego; habilidades.

Obrigada! Chegamos ao fim de nossa entrevista. Você gostaria de acrescentar algo mais em suas colocações sobre o assunto?

Mais uma vez agradecemos sua colaboração. Assim que tivermos resultados publicados encaminhamos para conhecimento, caso for de seu interesse.

APÊNDICE D – Termo de Consentimento Livre e Esclarecido Usuários

Observação: o presente TCLE foi enviado para os contatos selecionados via e-mail e ferramenta Google Formulários (*Google Forms*).

Prezado (a) Senhor(a),

Dirigimo-nos ao (ã) Senhor (a) para convidá-lo (a) a participar da pesquisa intitulada “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa”. Seu contato foi selecionado a partir da sua participação na *survey* realizada pelo Grupo de Trabalho – Rede de Dados de Pesquisa Brasileira em 2018. Na qual, vossa senhoria demonstrou interesse em participar de uma entrevista para explorar as questões abordadas na *survey* em mais profundidade.

Sua colaboração no presente estudo é muito importante, mas a decisão em participar é sua. Para tanto, leia atentamente as informações abaixo e não se apresse em decidir. Se você concordar em participar basta preencher esse termo. A coleta de dados ocorrerá por meio de entrevista semiestruturada com roteiro prévio que observa variáveis que visam atingir o objetivo principal do estudo. As entrevistas serão aplicadas de modo síncrono em horário e data pré-agendados entre pesquisador e pesquisado, por intermédio de plataforma de videoconferência *MS Teams* UFRGS com a possibilidade de gravar áudio e vídeo. As entrevistas estão programadas para durar de 30 minutos a 1 hora. As respostas serão transcritas, organizadas e analisadas pela pesquisadora do projeto. Todos os procedimentos para a garantia da confidencialidade aos participantes serão observados, procurando evitar descrever informações que possam lhe comprometer. O benefício esperado ao compreender “Quem são as personas que representam o público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa?” é contribuir para a promoção da ciência aberta no Brasil. Ao participar dessa pesquisa, você não terá nenhum benefício direto; entretanto, esperamos que futuramente os resultados desse estudo sejam usados em benefício de outras pessoas e à ciência brasileira. Todas as informações coletadas nessa investigação são estritamente confidenciais. Acima de tudo interessam os dados coletivos e não aspectos particulares de cada entrevistado. A participação nessa pesquisa não traz complicações legais de nenhuma ordem e os

procedimentos utilizados obedecem aos critérios da ética na Pesquisa com Seres Humanos conforme a Resolução nº 510, de 07 de abril de 2016, do Conselho Nacional de Saúde. Nenhum dos procedimentos utilizados oferece riscos à sua dignidade. O risco que você pode correr ao realizar a pesquisa é de ser identificado mesmo com todos os cuidados de sigilo adotados. Se você não concordar em participar ou quiser desistir em qualquer momento, isso não causará nenhum prejuízo a você. Você não terá nenhum tipo de despesa por participar desse estudo, bem como não receberá nenhum tipo de pagamento por sua participação.

Após as explicações, solicitamos seu consentimento de forma livre e esclarecida para ser um(a) participante da pesquisa. Para tanto, acesse e preencha o formulário Consentimento Livre e Esclarecido. Desde já agradecemos a disponibilidade.

Pergunta 1 – *Você acha que está suficientemente informado(a) a respeito da pesquisa “Personas: os perfis do público usuário das Ciências Biológicas - ecologia para iniciativas em defesa do movimento brasileiro para o acesso aberto aos dados de pesquisa” e concorda de livre e espontânea vontade em participar como informante?*

Sim Não

Pergunta 2 – *Para você, qual o horário adequado para realizar a entrevista?*

Das 9h às 10h Das 10h às 11h Das 11h às 12h

Das 14h às 15h Das 15h às 16h Das 16h às 17h

Pergunta 3 – *Para você, qual o dia da semana adequado para realizar a entrevista?*

segunda-feira terça-feira quarta-feira quinta-feira sexta-feira sábado

Pergunta 4 – *Nome completo:*

Pergunta 5 – *Endereço de e-mail para contato:*

Pergunta 6 – *Telefone para contato (whatsapp):*

Mensagem final do Google Forms

Suas respostas foram registradas com sucesso! Agradecemos sua colaboração e colocamo-nos à

disposição para esclarecimentos adicionais. Caso você possua perguntas sobre o estudo ou se pensar que pode haver algum prejuízo por sua participação, você pode contatar a pesquisadora responsável Mestra Marina Rodrigues Martins (PPGCOM/UFRGS) pelo e-mail mrodriguesmartins@gmail.com, bem como a orientadora Profa. Dra. Samile Andréa de Souza Vanz pelo e-mail samile.vanz@ufrgs.br ou o Comitê de Ética em Pesquisa – CEP/UFRGS (51) 3308-3738. Horário de atendimento do CEP, das 08:00 às 12:00 e das 13:00 às 17:00.