

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

JACKSON DA SILVA MEDEIROS

**UMA INVESTIGAÇÃO SOBRE A AUTORIA DE DADOS CIENTÍFICOS: TEIAS DE
UMA REDE EM CONSTRUÇÃO**

PORTO ALEGRE

2015

JACKSON DA SILVA MEDEIROS

**UMA INVESTIGAÇÃO SOBRE A AUTORIA DE DADOS CIENTÍFICOS: TEIAS DE
UMA REDE EM CONSTRUÇÃO**

Tese apresentada ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Doutor em Comunicação e Informação.

Linha de pesquisa: Informação, redes sociais e tecnologias

Orientadora: Prof^a. Dr^a. Sônia Elisa Caregnato

PORTO ALEGRE

2015

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Alexandre Netto

Vice-Reitor: Prof. Dr. Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Prof^a. Dr^a. Ana Maria Mielniczuk de Moura

Vice-Diretor: Prof. Dr. André Iribure Rodrigues

PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

Prof^a. Dr^a. Nísia Martins do Rosário (coordenadora)

Prof. Dr. Alexandre Rocha da Silva (coordenador substituto)

CIP - Catalogação na Publicação

Medeiros, Jackson da Silva

Uma investigação sobre a autoria de dados científicos: teias de uma rede em construção / Jackson da Silva Medeiros. -- 2015.
192 f.

Orientadora: Sônia Elisa Caregnato.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, Programa de Pós-Graduação em Comunicação e Informação, Porto Alegre, BR-RS, 2015.

1. Autoria. 2. Dados científicos. 3. e-Science. 4. Repositório de dados. 5. Teoria Ator-Rede. I. Caregnato, Sônia Elisa, 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 e Informação
Rua Ramiro Barcelos, 2705, 2º andar, Santana, Porto Alegre, RS
Campus Saúde
CEP 90035-007
Telefone: (51) 3308-5116
WWW: <http://www.ppgcom.ufrgs.br>
E-mail: ppgcom@ufrgs.br

JACKSON DA SILVA MEDEIROS

Uma investigação sobre a autoria de dados científicos: teias de uma rede em construção

Tese apresentada ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Doutor em Comunicação e Informação.

Banca examinadora

Profª. Drª. Maria Inês Tomaél
Universidade Estadual de Londrina

Prof. Dr. Rodrigo Silva Caxias de Sousa
Universidade Federal do Rio Grande do Sul

Profª. Drª. Ida Regina Chittó Stumpf
Universidade Federal do Rio Grande do Sul

Prof. Dr. Alex Fernando Teixeira Primo
Universidade Federal do Rio Grande do Sul

Profª. Drª. Sônia Elisa Caregnato (orientadora)
Universidade Federal do Rio Grande do Sul

Profª. Drª. Ana Maria Mielniczuk de Moura (suplente)
Universidade Federal do Rio Grande do Sul

Porto Alegre, 06 de maio de 2015.

Para aqueles que amo. Com todo meu coração.

AGRADECIMENTOS

Raras foram as vezes que pensei poder chegar a este ponto. No entanto, diversos foram os momentos que pensei não chegar. Nesse processo há pessoas – actantes, na verdade – que merecem meu reconhecimento, pois fizeram com que eu não desistisse. Talvez tenham duvidado menos do que eu duvidei.

Assim sendo, agradeço:

À Prof^a. Sônia Caregnato, pelos ensinamentos, pela paciência, pela compreensão, pela provocação, pela parceria e pela possibilidade de desenvolver este estudo. Provavelmente eu nunca vá conseguir retribuir o que me foi dado. Terá minha eterna admiração.

Aos membros da banca de qualificação – Prof. Alex Primo e Prof^a. Maria Inês Tomaél – e da banca final de defesa – Prof. Alex Primo e Prof^a. Maria Inês Tomaél, novamente, e Prof^a. Ida Stumpf e Prof. Rodrigo Caxias. Obrigado por despenderem seu tempo lendo, avaliando e trazendo contribuições para deste trabalho.

Ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul.

Aos colegas do Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul. Em especial: Rita Laipelt, Rita Portela e Rodrigo Caxias.

Aos entrevistados. Dispuseram de seu tempo para esta pesquisa e forneceram dados e informações essenciais, além de ideias, percepções, modos de analisar e pensar sobre os objetos estudados.

To Dropbox Team. You saved my life – in fact – several times!

À CAPES, pelo auxílio financeiro durante mais de dois anos de doutorado.

À Momo. Obrigado por gritar por comida. Obrigado por sempre fugir de mim.

A meu irmão (Jeferson) e minha cunhada (Priscila). Obrigado por me mostrarem que uma meta exige esforço.

À Martiela. Tenho certeza que sabes a importância que tens para mim e na condução da minha vida. Teus conselhos, carinhos, puxadas de orelhas e brigas fizeram com que, em muitas vezes, minha razão voltasse à tona. És a emoção do pouco que tenho de razão. Amo-te. Muito.

Aos meus pais, Cleuza e Jackson. Não tenho o que dizer a vocês. Sobram-me elogios. Faltam-me palavras para expressá-los. Apenas amo vocês. Mas nem isso descreve o que realmente sinto. Obrigado por tudo.

Aos autores citados e não citados neste trabalho. Como vocês são difíceis...

Obrigado a todos. Nunca esquecerei o que me foi dado.

“Eventually everything connects – people, ideas, objects. The quality of the connections is the key to quality *per se*.” (Charles Eames)

RESUMO

Estudo qualitativo de natureza descritiva, objetivando analisar a autoria científica a partir do compartilhamento de dados de pesquisa. Para esta finalidade, foi utilizada a Teoria Ator-Rede (ANT), a qual busca, a partir de uma ideia de simetria entre humanos e não humanos, não assumir uma divisão entre essas entidades, possibilitando visualizar o estabelecimento e a dissolução de uma rede a partir das relações que se criam em diversos níveis. Esses atores-rede, actantes, estão mobilizados para a estruturação de redes temporárias a partir de um programa de ações, através da transformação e da modificação do significado ou dos elementos que veiculam. Os dados levantados, os quais serviram de base para descrição e análise a partir da literatura que também emergiu do que foi coletado, foram recolhidos a partir de observação, de entrevistas e de materiais bibliográficos. A investigação ocorreu junto ao Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio). Conclui que os objetos conceituais e pragmáticos não podem ser vistos de maneira isolada, mas sim como atores-rede afetados diretamente pela tecnologia, devendo ser considerados atuantes importantes no processo de produção, compartilhamento e uso de dados digitais de pesquisa. A construção de um fato científico e social como a autoria, é coordenada por uma série de elementos que foram levantados com o auxílio da ANT, como a coleta de dados, a curadoria, os próprios dados científicos, seus metadados, as fontes de fomento, a política de dados, o repositório e seus softwares de gerenciamento e envio de dados e metadados, o licenciamento, a propriedade e a responsabilidade, agindo como um tipo de instrumentalização de um arranjo de fenômenos, isto é, uma ferramenta que, atuante como conceito e/ou como processo, permite que seja consolidado um conjunto de ações instituídas a partir de regras explícitas e implícitas.

Palavras-chave: Autoria; Dados científicos; e-Science; Repositório de dados; Teoria Ator-Rede.

ABSTRACT

Qualitative study of descriptive nature, aimed to analyze the scientific authorship from the data sharing. For this purpose, we used the Actor-Network Theory (ANT), which seeks, from an idea of symmetry between human and non-human, do not assume a division between those entities, allowing to view the establishment and dissolution of a network from the relationships that are created at different levels. These actors-network, actants, are mobilized for structuring temporary networks from an action program, through the transformation and modification of the meaning or the elements they provide. The data collected, which were the basis for description and analysis from the literature that also emerged from what were collected by observation, interviews and bibliographic materials. The study was carried out by the Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio). Concludes that conceptual and pragmatic objects can not be seen in isolate way, but as actors directly affected by the technology, should be considered important actants in the process of production, sharing and use of digital research data. The construction of a scientific and social fact as authorship, is coordinated by a number of elements that were raised with the help of ANT, such as data collection, curation, scientific data and its metadata, funding sources, data policy, repository and its management software for data and metadata, licensing, ownership and responsibility, acting as a kind of exploitation of a phenomena arrangement, ie, a tool that can be seen as concept and/or as a process, allows it to be consolidated a set of actions established from explicit and implicit rules.

Keywords: Authorship; Scientific data; e-Science; Data repository; Actor-Network Theory.

LISTA DE FIGURAS

Figura 1 – Paradigmas da ciência	31
Figura 2 – Modelo de publicações científicas de Garvey-Griffith	33
Figura 3 – Modelo de comunicação científica para 2020	34
Figura 4 – Tela de busca do repositório	103
Figura 5 – Tela com os resultados da busca	104
Figura 6 – Excerto da tela que contém o registro completo de conjunto de dados	105
Figura 7 – Tela inicial do Morpho	111
Figura 8 – Tela de apresentação do Morpho – Assistente para criação de novo pacote de dados	112
Figura 9 – Exemplo de uma planilha de levantamento ecológico sem anotações que permitam atribuir significado aos dados	115
Figura 10 – Exemplo da planilha da Figura 9 preparada para importação em banco de dados ou pacote estatístico	115
Figura 11 – Pessoas e organizações	130
Figura 12 – Proprietário detalhes	131

LISTA DE QUADROS

Quadro 1 – Características da e-Science x Características da pesquisa	36
---	----

LISTA DE ABREVIATURAS E SIGLAS

ANT	Teoria Ator-Rede
CDB	Convenção sobre Diversidade Biológica
CERN	Organização Europeia para a Pesquisa Nuclear
DataONE	Data Observation Network for Earth
DOU	Diário Oficial da União
EML	Ecological Metadata Language
INPA	Instituto Nacional de Pesquisas da Amazônia
KNB	Knowledge Network for Biocomplexity
LNCC	Laboratório Nacional de Computação Científica
MCTI	Ministério da Ciência, Tecnologia e Inovação
NCEAS	National Center for Ecological Analysis and Synthesis
OECD	Organisation for Economic Co-operation and Development
OMPI/ONU	Organização Mundial da Propriedade Intelectual na Organização das Nações Unidas
PPBio	Programa de Pesquisa em Biodiversidade
RRI	research-related information
VoIP	Voz sobre IP
XML	eXtensible Markup Language
W3C	World Wide Web Consortium

SUMÁRIO

1 INTRODUÇÃO	16
2 RUMO AOS DADOS CIENTÍFICOS: COMUNIDADE CIENTÍFICA, COLABORAÇÃO E PARADIGMA.....	23
2.1 e-Science: dados científicos para pesquisa.....	30
2.2 Dados digitais de pesquisa e infraestrutura de compartilhamento.....	37
3 AUTORIA	46
3.1 O direito autoral como fala inicial para a autoria.....	47
3.2 A autoria: elementos de uma função	55
3.3 Ética autoral: em busca da responsabilidade	60
4 VESTÍGIOS DE UMA REDE EM CONSTRUÇÃO: AS TEIAS DA TEORIA ATOR- REDE	66
4.1 Rede de atores e a Teoria Ator-Rede.....	70
4.2 Tradução, mediação, caixa-preta	76
4.3 Organizando a narração: a Teoria Ator-Rede como método descritivo	82
4.4 E depois do passeio: anseios entre a ANT e a autoria	87
5 METODOLOGIA: TRAÇOS E LINHAS DE UM MÉTODO DESCRITIVO.....	92
5.1 O <i>locus</i> de pesquisa	93
5.2 Observação e entrevista	95
5.3 Sobre a coleta e a análise de dados.....	96
6 PESQUISA EMPÍRICA: A DESCRIÇÃO DO OBJETO	101
6.1 O Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio): a interface de acesso aos dados e aos metadados	101
6.2 O Repositório e suas ferramentas	107
6.3 O entendimento sobre os (conjuntos de) dados científicos	113
6.4 Dos (conjuntos de) dados aos metadados: curadoria e preservação.....	118
6.4 O gerenciamento político: a política de dados	123
6.6 A autoria	128
7 REUNINDO A RAZÃO: ELEMENTOS EM AÇÃO	134
8 CONCLUSÕES	145
REFERÊNCIAS.....	149
APÊNDICES.....	162
APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.....	163
APÊNDICE B – PERGUNTAS PARA ENTREVISTAS	165

ANEXOS	167
ANEXO A - PORTARIA Nº 693, DE 20 DE AGOSTO DE 2009 – Institui, no âmbito do Programa de Pesquisa em Biodiversidade - PPBio, a Política de Dados.....	168
ANEXO B – TERMO DE COMPROMISSO	178

1 INTRODUÇÃO

Autoria e comunicação científica são conceitos que, alicerçados na colaboração, entrecruzam-se. Uma vez que os esforços dos membros das comunidades científicas são coadunados com os de outros indivíduos e existe uma necessidade de comunicação, isto é, troca de informações entre pares para que novos conhecimentos sejam gerados com base em dados e/ou informações, a colaboração atua como uma ligação entre esse empenho investigativo.

Surge, nesse sentido, o conceito de autoria, isto é, uma entidade responsável por dar credibilidade à coleta de dados, às informações, às opiniões etc. No entanto, a relação estabelecida entre a autoria e os fatos científicos sofreu, e segue sofrendo, alterações durante períodos de tempo, principalmente no que se refere à tecnologia e à forma de fazer ciência.

Com o surgimento da ciência moderna no início do século XVII, as formas de fazê-la se multiplicaram, mas sua grande explosão ocorre a partir dos sistemas eletrônicos de informação produzidos após a Segunda Guerra Mundial. Neste sentido, o avanço científico depende cada vez mais da existência de infraestruturas que possibilitem aos cientistas explorar dados de forma eficaz, uma vez que o modo de fazer pesquisa também “está mudando e estamos no limiar de uma nova era da ciência dirigida por dados”¹ (HEY; HEY, 2006, p. 515).

Percebemos que, neste caminho, os avanços da ciência e da tecnologia afetam aspectos sociais, econômicos, comportamentais, entre outros, agindo de forma rápida e constante, alterando o modo de lidar com os problemas relativos à própria ciência. Dentre esses avanços está a questão dos dados digitais de pesquisa, isto é, o material que foi coletado e registrado durante uma pesquisa e que serve tanto como insumo para novas descobertas quanto para validação do que foi construído.

Estes dados devem estar alicerçados em uma infraestrutura que permita a cientistas e a pesquisadores o acesso a dados científicos primários distribuídos, utilizando acesso remoto a estes conteúdos, suportando a “capacidade de acessar,

¹ No original: “[...] is changing and we are on the threshold of a new era of data-driven science”.

mover, manipular e extrair dados[, sendo essa] a exigência central dessas novas aplicações das ciências da colaboração² (HEY; HEY, 2006, p. 517).

Dados científicos digitais de pesquisa demonstram, no entanto, ser mais do que apenas uma *commodity*, ou seja, insumos brutos que podem atuar na geração de novos produtos/conhecimentos, mas atuam também como ativos de pesquisa. Para refletir sobre isso, basta pensarmos nos altos custos para a geração de dados do mapeamento genético humano, por exemplo. Não há possibilidade da utilização única desses dados, isto é, seu uso deve ser feito a partir da sua reutilização de forma contínua e repetida.

Entendendo que os dados científicos são insumos para novas pesquisas e também o produto delas, consideramos que eles aumentam o poder comunicacional e informacional dos pesquisadores. Neste caso, são necessários estudos capazes de especificar como se (re)organiza a comunicação científica – e como parte disto a questão da autoria –, uma vez que, em virtude da novidade e conseqüente pouca existência de literatura referente ao tema, ainda existem preocupações incipientes sobre a utilização desse tipo de dados para o desenvolvimento científico, fazendo com que os esforços sejam alocados, em geral, em níveis computacionais, com vistas a melhorar o processamento e a transmissão de grandes quantidades dos mesmos.

Isto é, os processos que compõem o compartilhamento de dados, ou seja, o acesso e uso aos dados primários de investigação, são de importância ímpar não apenas na informatização do processo, mas também no que diz respeito à compreensão dos próprios dados, sendo, como aponta Lannon (2011), um desafio oportuno para as pesquisas científicas que ocorrem no século 21, em um mundo cada vez mais conectado digitalmente, uma vez que o fluxo de dados gerados e compartilhados é realmente grande.

Compreendemos que o compartilhamento de dados constitui-se a partir de uma infraestrutura informacional, isto é, artefatos de tecnologias da informação e da comunicação necessários para o processo dos resultados de pesquisa, bem como de sua face correspondente às relações implícitas e explícitas de regras, interesses,

² No original: “The ability to access, move, manipulate and mine data is the central requirement of these new collaborative science applications”.

anseios, decisões etc. Assim sendo, essas relações estão em níveis micro e macro, podendo ser percebido que o compartilhamento está baseado em associações dos mais diversos tipos, circunscrevendo entidades de natureza variável.

Isso está diretamente relacionado à infraestrutura do compartilhamento de dados científicos, novo campo que merece atenção e análises acuradas em todas suas camadas, uma vez que se trata de uma tentativa de conhecer mais sobre algo ainda tão incipiente:

A pergunta inicial na construção de infraestruturas de informação científica é o que construir quando sabemos tão pouco sobre como ela é usada. [...] Tecnologia e pesquisa estão interligadas e mutuamente sofrendo influência. A criação de uma infraestrutura de informação avançada exige a concepção e implantação de tecnologia com fins específicos. Identificar os usos requer uma compreensão dos contextos sociais e políticos nos quais a estrutura de informação existe.³ (BORGMAN, 2010, p. 33-35).

Tomando a ideia de que a grande movimentação científica do mundo e a geração de dados em muitos campos acarreta o crescimento contínuo e em larga escala do modo como a comunicação científica vem se apresentando, a autoria é um processo complexo que está circunscrito desde a coleta dos dados, passando pela curadoria, compartilhamento, acesso e (re)uso desses dados. Este é nosso mote de estudo.

A autoria é uma questão de pesquisa que envolve problemas ainda não discutidos, principalmente quando falamos sobre autoria de dados científicos. Isso se dá porque, além da incipiência e complexidade do tema, os trabalhos que buscam tratar essa matéria o fazem tendo como ponto de partida os documentos bibliográficos, especialmente artigos científicos e livros, mas também teatro, cinema etc. Dados primários, além de considerados, impropriamente, apenas como insumo

³ No original: "A initial question in constructing scholarly information infrastructures is what to build when we know so little about how it be used. [...] Technology and research are intertwined and mutually influencing in many respects. Creating an advanced information infrastructure requires designing and deploying technology with specific uses in mind. Identify those uses requires an understanding of the social and political contexts in which information structure exists".

de pesquisa, não despertaram, até o momento, interesse sobre o entendimento de sua autoria.

Trabalharemos aqui com um pressuposto, isto é, uma proposição que serve como ponto de apoio para esta pesquisa: a questão autoral tem sua essência alterada a partir do paradigma da e-Science e do compartilhamento de dados científicos, uma vez que o que se entende por autoria, nesse contexto, está ligada a uma nova constituição do modo de fazer ciência.

Com isso podemos sintetizar que a tese investiga em que consiste a autoria dos dados digitais de pesquisa segundo o Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade⁴ (PPBio). Para isso traz à tona e discute atores-rede envolvidos no processo autoral de dados de pesquisa. Por este viés, a pesquisa buscou relacionar o entendimento sobre o que a literatura da área entende sobre autoria, isto é, a noção de autor e autoria propriamente ditos, com a questão da e-Science e do compartilhamento de dados digitais de pesquisa, tema e demanda ainda novos no que se refere à comunicação científica.

Na busca por trabalhos que se referem à temática de autoria científica, grande parte do que recuperamos são diretrizes para organização da ordem autoral, tal como critérios de pontuação para firmar um ordenamento dos nomes que serão vistos como autores. Assim, apresentam-se, em geral, como *guidelines* (diretrizes) que tomam uma vertente prática de ordenação, sem discutir, pelo menos, qual o propósito da autoria.

A partir disso concluímos que não existem estudos com cobertura suficiente sobre o entendimento das questões referentes à autoria, principalmente no que concerne aos dados digitais de pesquisa. Deste modo, buscamos estabelecer uma explicação do fenômeno requer que nos empenhemos em uma perspectiva de análise de redes sociotécnicas para a descoberta de actantes que constituem a questão.

A pretensão do trabalho é analisar uma parte da infraestrutura de comunicação científica do compartilhamento de dados de pesquisa com vistas a

⁴ O Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio) foi escolhido como *locus* de pesquisa e será detalhado posteriormente.

compreender a autoria de dados científicos, uma vez que pode revelar estruturas explícitas e implícitas através de construtos sociais/coletivos, indo além do gerenciamento e tratamento de dados (curadoria). A importância desse tipo de abordagem é corroborada por Borgman (2010) ao anotar que a publicação de dados não possui os mesmos incentivos do que a publicação bibliográfica.

Devido a uma tradição que faz com que as publicações científicas em formato bibliográfico tenham valor acentuado para a competitividade acadêmica, os esforços científicos estão ligados à constante tentativa de divulgação e de amplo acesso. O compartilhamento e o acesso aos dados científicos primários ainda é desprovido de incentivos, podendo, inclusive, sofrer com fortes manifestações contra sua disponibilização (BORGMAN, 2010).

O foco deste trabalho está direcionado a analisar a questão da autoria a partir do compartilhamento de dados de pesquisa, compreendendo que sua construção é socialmente dada. A partir disso, podemos notar que objetos e/ou conceitos, quaisquer que sejam, mantêm relações entre si através de um conjunto de atividades, sendo suas existências objeto de análises que envolvem suas interdependências a partir de variáveis que se formam como uma rede, envolvendo questões que estão conectadas e que só podem persistir a partir da existência de outros atores, humanos ou não, que sustentam a rede.

Parte dessa rede é o repositório de dados científicos. Vemos que um repositório é o dispositivo onde há, dentro de um processo de comunicação científica formal, o estabelecimento da autoria como tal, uma vez que é através da publicação dos dados e sua consequente disponibilização com base em certos parâmetros instituídos que há o encontro da publicação (conjunto de dados) com seu público (usuários).

Também é importante considerar que as questões que envolvem o gerenciamento de dados científicos são prioritárias no que diz respeito às instituições que recebem destaque em âmbito mundial. De modo mais amplo, os países que investem em pesquisa e desenvolvimento estão cada vez mais promovendo a gestão e o acesso aos dados de pesquisa, uma vez que isso possibilita a redução do consumo de recursos humanos e financeiros, permitindo

que investimentos sejam alocados em áreas que permitam crescimento de uma nação.

Assim, é importante relatar que os dados não apenas existem e são recuperados, analisados e reutilizados. São, sim, construídos a partir de um processo de autoria, entidade com propósitos, atravessado por diversas questões sociais, políticas, institucionais, entre outras.

No que concerne ao desenvolvimento desta pesquisa, sua constituição foi dada na medida em que avançava, fazendo com que tanto os caminhos epistemológicos quanto os metodológicos fossem alterados. Isso se deu como consequência do momento em que partimos da ideia de observar e examinar o fenômeno do compartilhamento de dados científicos através da citação. No entanto, notamos que isso implicava na tentativa de compreender primeiramente a autoria no contexto abrangido, fazendo com que o trabalho deslocasse seu foco para a questão autoral.

Para tal tarefa, destacamos como objetivo geral analisar a autoria científica a partir do compartilhamento de dados de pesquisa. Os objetivos específicos são (i) identificar os actantes e seus papéis em relação à autoria de dados digitais de pesquisa; (ii) compreender as relações existentes entre os actantes; e (iii) explicitar elementos e vínculos existentes entre os atores-rede no que se refere à constituição da autoria de dados científicos.

Buscamos, pelos objetivos expostos acima, um estudo capaz de pensar a autoria de dados científicos, sendo isto caracterizado pela movimentação dos actantes envolvidos no processo de (re)organização da comunicação científica a partir do viés do compartilhamento de dados de pesquisa, percebendo o que está intimamente relacionado com as manifestações associativas dos atores-rede. Utilizamos para isso a Teoria Ator-Rede (ANT), a qual, a partir da existência de uma simetria entre humanos e não humanos, isto é, não assumir uma divisão entre humanos e não humanos, visa possibilitar que uma rede se estabeleça, se modifique e se dissolva a partir das relações que se estabelecem em diversos níveis. Esses atores-rede, actantes, estão mobilizados para a estruturação de redes temporárias a partir de um programa de ações, através da transformação e da modificação do significado ou dos elementos que veiculam.

Isso posto, chegamos a seguinte questão de pesquisa que este estudo se propõe responder: como é constituída a autoria de dados científicos?

Para responder a essa questão, o trabalho está estruturado da seguinte maneira: após a introdução, com a apresentação do tema de pesquisa, da justificativa, do problema de pesquisa e dos objetivos geral e específicos, o capítulo 2 – *Rumo aos dados científicos: comunidade científica, colaboração e paradigma* – apresenta a e-Science e o compartilhamento de dados de pesquisa, bem como sua importância, seguido do capítulo 3, responsável por tratar a autoria, evidenciando como tema é apresentado na literatura e procurando elementos que colaborem com o objeto de nosso estudo.

O capítulo 4 – *Vestígios de uma rede em construção: as teias da Teoria Ator-Rede* – tem como objetivo realizar uma introdução à ANT e compreendê-la como forma de trazer à baila os elementos que compõem a autoria, vislumbrando traços que liguem a Teoria à autoria. Após, apresentamos a metodologia descritiva – capítulo 5 – utilizada no estudo, assim como o *locus* de pesquisa e os métodos utilizados para coleta de dados.

O capítulo 6 diz respeito à pesquisa empírica realizada em relação ao objeto de pesquisa, correspondendo ao modo como a coleta de dados foi realizada, além da descrição pormenorizada sobre o Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade, desde sua interface e ferramentas, passando pelo conjunto de dados e metadados até a política de dados e a autoria.

O capítulo 7 mostra o que – pretensiosamente – chamamos de *Reunindo a razão: elementos em ação*, realizada a partir da análise dos dados coletados do Repositório do PPBio com base em observações e em entrevistas, sendo discutidos os elementos e suas relações que compõem a autoria de dados científicos de pesquisa. O capítulo 8 apresenta as conclusões do trabalho e, após, são apresentadas as referências, os anexos e os apêndices.

2 RUMO AOS DADOS CIENTÍFICOS: COMUNIDADE CIENTÍFICA, COLABORAÇÃO E PARADIGMA

“Scholarly data and documents are the most value when they are interconnected rather than independent” (BORGMAN, 2010, p. 10).

Entendendo que este estudo ressalta a importância da informação científica, em especial na forma de dados, para a comunidade científica, a possibilidade de “pensar sobre a relevância da ciência demanda reconhecer a importância da informação científica, do conhecimento científico, da comunidade científica, e, por conseguinte, da comunicação científica” (TARGINO, 2000, p. 1). Assim, este capítulo disserta sobre o compartilhamento de dados de pesquisa, abordando-o a partir do sistema de comunicação científica, resgatando elementos da própria ciência que podem contribuir para sua apreensão como uma infraestrutura que, além de tecnológica, arregimenta posições dos atores envolvidos.

A ciência, de acordo com Morin (2010), ocorre a partir do próprio juízo do que vem a ser ciência. Cria-se, neste espaço de reflexão, um processo e um produto a partir da complexidade sobre a construção e a observação de incertezas e de questões propostas pela natureza e pela cultura do ser humano, as quais são próprias do mundo habitado e modificado por ele. Assim, a ciência estaria diretamente ligada à tecnologia, sendo que suas atividades se influenciam de forma recíproca. Esse contexto que liga ciência e tecnologia pode ser robustecido pela definição proposta por Targino (2000, p. 2): “[...] ciência refere-se, sobretudo, ao conjunto de procedimentos transformadores advindos da vinculação ciência-tecnologia e de seus resultados inscritos no meio ambiente”.

A ciência, grande ou pequena, gera inúmeros resultados, principalmente em relação à publicação bibliográfica. Latour e Woolgar (1997) atestam no livro *A vida de laboratório: a produção dos fatos científicos* a importância da literatura a partir das trocas de informações e demonstrações realizadas pelos pesquisadores: “o que se deve reter [...] é essa onipresença da literatura [...] e da qual apenas uma

pequena parte irá surgir sob a forma da publicação” (LATOURE; WOOLGAR, 1997, p. 47).

A geração de literatura ocorre, sem dúvida, a partir do grande valor proporcionado pela síntese de dados, uma vez que os pesquisadores, em geral, estão interessados em resultados e valores agregados a partir de dados crus que podem ser analisados e transformados em informação para comunicação. Latour e Woolgar (1997, p. 43), em suas experiências dentro de um laboratório, descrevem essa questão de produção de uma súpula de dados que se tornam relevantes: “os tubos, que são manejados com grande cuidado durante uma semana [...], agora perderam qualquer valor. Só há interesse por essa folha cheia de números”. As bases de dados que compõem esses estoques fornecedores de dados não estocam dados bibliográficos e/ou documentos, como comumente acontece, mas, sim, dados crus (primários), como sequências genéticas, dados de transações financeiras, entre outros. Esses dados são insumos para novas pesquisas, também sendo utilizados como forma de comprovação científica de um fato descoberto.

Inferimos, portanto, que não é apenas a literatura publicada diariamente que alimenta as investigações feitas por pesquisadores e cientistas do mundo inteiro. O processo passa também pelos dados coletados e que ainda podem não ter recebido qualquer tipo de tratamento e/ou interpretação para geração de documentos bibliográficos por alguém interessado. São dados científicos primários que podem ser disponibilizados e que são capazes de fundamentar e/ou alavancar pesquisas. São recursos reaproveitáveis, reusáveis, capazes de subsidiar não só uma, mas diversas pesquisas.

A experimentação a partir de dados de pesquisa parece ter se tornado um dos grandes impulsionadores das descobertas e inovações, uma vez que diversas ciências se utilizam de dados científicos para que possam desenvolver seus estudos e promover avanços em suas áreas. Morin (2010, p. 43) menciona esta questão: “o dispositivo experimental, em última instância, é a seleção de certo número de dados; é um transplante no meio artificial, que é o laboratório, e permite agir nas variações desejadas”.

Fica latente a ideia das ciências, principalmente das ciências na contemporaneidade, em que os dados têm grande valor como insumo para

pesquisa, desenvolvimento e inovação. Isto é reforçado por Morin (2010) quando, ao tratar da objetividade do conhecimento científico, diz que “os dados nos quais se baseiam as teorias científicas são objetivos, objetivos pelas falsificações, e isso é absolutamente incontestável”. Ou seja, podemos refletir que os dados por si só não exprimem teorias, mas, a partir de sua análise, é possível construí-las e/ou deduzi-las, auxiliando a interpretação do mundo, permitindo que investigadores sejam capazes de interpretá-los e desenvolver teorias, sistemas etc.

Dados científicos permitem a realização de testes, o que é uma regra do jogo científico, sendo essa testagem que possibilita desenvolver avanços científicos e tecnológicos. Poder realizar testes em dados previamente coletados e anotados com metadados, isto é, reutilizar o que já foi mapeado, possibilita maiores e melhores progressos na ciência, além de oportunizar que recursos financeiros e humanos sejam poupados.

Não restam dúvidas que “os fenômenos dependem do material, eles [os fenômenos] são *totalmente construídos* pelos instrumentos⁵ utilizados no laboratório” (LATOUR; WOOLGAR, 1997, p. 61). São os dados que devem ser dispostos para que suas interpretações/análises sejam capazes de fomentar outros trabalhos (artigos para periódicos e conferências, livros etc.) ou outras pesquisas, geradas a partir de pontos já mapeados por outros pesquisadores.

Isso leva a crer que o modo de fazer ciência nos dias de hoje não permite que se oponha aparelhagem ou qualquer outro tipo de material físico, conceitual ou teórico aportado de outro domínio a uma análise exclusivamente humana. Isto é, “os aparelhos e as habilidades de um domínio materializam os resultados finais de debates ou controvérsias ocorridas em outro espaço” (LATOUR; WOOLGAR, 1997, p. 63), o que faz com que a aceitação inter/multi/transdisciplinar seja potencializada pelos avanços de cada campo e remetem a possibilidade de utilização futura em domínios completamente diferenciados daquele em que foi produzido. Assim, como mencionam Latour e Woolgar (1997, p. 64), “todo progresso realizado em um laboratório está de certa forma ligado ao que se passa nos outros domínios”.

Em relação à importância dos dados científicos, podemos pensar, por exemplo, em maquinários que podem custar bilhões de dólares e/ou que haja

⁵ A relação existente entre dados, fenômenos, materiais, pessoas etc. será tratada no capítulo 5.

apenas um exemplar no mundo, devido a custos, dificuldade de fabricação e manutenção, estratégia de pesquisa, entre outros fatores. Podemos tomar como referência o caso do Grande Colisor de Hádrons da Organização Europeia para a Pesquisa Nuclear (CERN), situado em Genebra. Os pesquisadores necessitam ter acesso aos dados primários e mapeamentos desse tipo de equipamento, além de contarem com a “força” dos pesquisadores daquele instituto para realizarem suas pesquisas que necessitem do mais poderoso acelerador de partículas já construído, uma vez que é impossível que todo pesquisador que planeje desenvolver sua pesquisa com dados gerados pelo equipamento o utilize *in loco*.

Os progressos científicos são alcançados a partir do momento em que entendemos, à luz de Ziman (1979, p. 155), que “a ciência é uma forma de conhecimento público”, sendo que seus veículos estão “a serviço do intercâmbio de informações”. A ciência, como trabalho de uma comunidade, percorre desde o processo de sua elaboração, colaborativo em essência, até a revisão por pares, os quais avaliam o trabalho apresentado aferindo a chancela científica e o repassam a uma comunidade que valida publicamente o que foi proposto. Por fim, são necessárias instituições que atuem na função de preservadoras e disseminadoras do conhecimento gerado.

Uma comunidade científica é, de maneira geral, uma associação, em forma de grupo(s), “dos indivíduos reunidos por elementos comuns em sua educação e aprendizado e caracterizados pela relativa plenitude de sua comunicação profissional e relativa unanimidade de seu julgamento profissional” (SALVI; BATISTA, 2007). Robert Merton (1970) enxerga a comunidade científica com base na estruturação de valores e regras de conduta ligadas à sua autonomia e organização a partir de uma perspectiva informal, constituídas por universalismo, comunismo, desinteresse e ceticismo organizado, ideais que dão origem ao *ethos* da ciência. Cabe ressaltar que “a comunidade mantém relações com outros fatores não incluídos na sua organização, não sendo, assim, um ente [completamente] autônomo”, sendo relevante a análise de possíveis “desvios dos imperativos sociais de seu funcionamento” (BAUMGARTEN, 2004, p. 102).

Conceito fundamental nesse plano é o de colaboração, isto é, pesquisadores das diversas áreas do conhecimento que trabalham em cooperação. Embora o

conceito de colaboração científica seja complexo e de difícil elaboração (KATZ; MARTIN, 1997), em termos de ciência, segundo Sonnenwald (2008, p. 645), colaboração pode ser tipificada pela “interação que ocorre dentro em um contexto social entre dois ou mais cientistas, facilitando o compartilhamento de significado e de conclusão das tarefas em relação a um objetivo compartilhado mutuamente”⁶. Isso vai de encontro ao modo de fazer ciência de tempos passados, aonde o cientista, sozinho ou com ajudantes, fazia descobertas e recebia notoriedade por seus feitos (BEAVER, 2000).

Como a globalização e a internacionalização continuam no caminho para a formação de uma comunidade global, a ênfase na cooperação e na vida em grupo tornou-se um contraponto cada vez mais comum a uma existente ênfase na competição e na individualidade⁷ (BEAVER, 2000, p. 30).

Beaver pontua que, nesse sentido, o processo de colaboração entre pesquisadores traz benefícios e malefícios. No que se refere aos benefícios, podemos citar: (a) eficiência nos projetos, uma vez que diversas mãos e cabeças executam mais tarefas de forma mais rápida; (b) capacidade de fragmentação temporária do grupo, onde a coleta de dados pode ser distribuída; (c) sinergia, possibilitando diversos olhares sobre um mesmo problema; (d) redução de risco, uma vez que a execução de diversos trabalhos pode gerar lucros – não necessariamente financeiros – variados e com maior chance de acerto; (e) flexibilidade, podendo manter, inclusive, projetos de ordem especulativa; (f) precisão, possibilitando que erros sejam detectados em tempo hábil; (g) retorno, disseminação, reconhecimento e visibilidade, já que diversos colaboradores podem apresentar trabalhos em diferentes encontros, colóquios, simpósios etc., ampliando o escopo de resposta de pares e aumentando a visibilidade do grupo (BEAVER, 2000).

⁶ No original: “interaction taking place within a social context among two or more scientists that facilitates the sharing of meaning and completion of tasks with respect to a mutually shared, superordinate goal”.

⁷ No original: “As globalization and internationalization continue, on the way to the formation of a global community, emphasis on cooperation and group life become an increasingly common counterpoint to an existing emphasis on competition and individuality”.

No que se refere às desvantagens, o autor lista (a) a pouca visibilidade individual, uma vez que o reconhecimento é visível ao grupo e não ao indivíduo; (b) o investigador principal pode perder contato direto com a pesquisa, já que ela está fragmentada e devido à necessidade de busca de recursos em sistemas administrativos e burocráticos; e (c) a privatização da pesquisa pode ser prejudicial aos *ethos* da mesma, pois as ideias que se tornam pesquisas patrocinadas não podem circular de forma livre, sendo a colaboração com outros laboratórios uma questão complexa devido à competitividade (BEAVER, 2000).

O autor ainda elenca dezoito motivos para a colaboração (BEAVER, 2000, p. 37):

(1) Acesso à expertise; (2) Acesso ao equipamento, recursos ou “material” que não possui; (3) Aumentar acesso aos fomentos; (4) Obter prestígio e visibilidade; para avanço profissional; (5) Eficiência: múltiplas mãos e mentes; facilidade para aprender o conhecimento tácito que acompanha uma técnica; (6) Fazer progresso mais rapidamente; (7) Resolução de problemas “maiores” [mais importante, mais abrangente, mais difícil, global]; (8) Aumentar a produtividade; (9) Conhecer pessoas, criar uma rede, como um “colégio invisível”; (10) Reequipar, aprender novas habilidades ou técnicas, geralmente para entrar em um novo campo, subcampo, ou problema; (11) Satisfazer curiosidade, interesse intelectual; (12) Compartilhar o entusiasmo de uma área com outras pessoas; (13) Encontrar falhas de modo mais eficaz, reduzindo erros e equívocos; (14) Manter o indivíduo mais focado na pesquisa, porque os outros dependem dele; (15) Reduzir isolamento e recarregar sua energia e entusiasmo; (16) Educar [um aluno ou a si mesmo]; (17) Avanço do conhecimento e da aprendizagem; (18) Por diversão, entretenimento e prazer.⁸ (BEAVER, 2000, p. 37).

⁸ [1] Access to expertise. [2] Access to equipment, resources, or “stuff” one doesn’t have. [3] Improve access to funds. [4] To obtain prestige or visibility; for professional advancement. [5] Efficiency: multiplies hands and minds; easier to learn the tacit knowledge that goes with a technique. [6] To make progress more rapidly. [7] To tackle “bigger” problems [more important, more comprehensive, more difficult, global]. [8] To enhance productivity. [9] To get to know people, to create a network, like an “invisible college”. [10] To retool, learn new skills or techniques, usually to break into a new field, subfield, or problem. [11] To satisfy curiosity, intellectual interest. [12] To share the excitement of an area with other people. [13] To find flaws more efficiently, reduce errors and mistakes. [14] To keep one more focussed on research, because others are counting on one to do so. [15] To reduce isolation, and to recharge one’s energy and excitement. [16] To educate [a student, graduate student, or, oneself]. [17] To advance knowledge and learning. [18] For fun, amusement, and pleasure.

Notamos que a colaboração ocorre de forma a trazer benefícios a partir de um processo social entre indivíduos que coordenam atividades e que estão engajados em algum tipo de trabalho, compartilhando e cooperando entre si para a geração de conhecimento que envolve processos complexos e grandes somas de esforços financeiros e intelectuais, possibilitando, de maneira geral, acesso livre à troca de conhecimentos.

Sendo a colaboração um aspecto da ciência na atualidade, permite, a partir do trabalho em conjunto, que o conhecimento avance de maneira mais contundente se comparado aos processos individualizados. Como ressalta Ronan (1987, p. 12), “a ciência é um conjunto de conhecimentos crescentes e em expansão, até o ponto em que aquelas mudanças são motivadas por experiências ainda mais complexas”. São esses novos paradigmas e as revoluções científicas que resultam em observações e escaladas de novos rumos. De forma clara, é possível buscar apoio em Morin (2010, p. 45) para compreender as questões de paradigma e revolução científica, como propostas por Thomas Khun. Diz o autor que:

[...] o paradigma é aquilo que está no princípio da construção das teorias, é o núcleo obscuro que orienta os discursos teóricos neste ou naquele sentido. Para Kuhn, existem paradigmas que dominam o conhecimento científico numa certa época e as grandes mudanças de uma revolução científica acontecem quando um paradigma cede seu lugar a um novo paradigma, isto é, há uma ruptura das concepções do mundo de uma teoria para outra. (MORIN, 2010, p. 45).

Kuhn (1990) afirma que não há acúmulo de teorias, a não ser em casos raros; elas não andam de forma paralela; elas, sim, sobrepõem-se uma a outra, sendo inviável sua coexistência, uma vez que a antiga teoria aceita é ampliada e/ou absorvida pela mais nova. As mudanças de paradigma tendem a acontecer quando uma prática científica suplanta a outra, com consequências que extrapolam o campo científico, alterando campos sociais, institucionais etc.

As questões ditas paradigmáticas, aquelas que incitam uma virada no modo de executar pesquisas e nos rumos que elas tomam são “realizações científicas universalmente reconhecidas que, durante algum tempo, fornecem problemas e

soluções modelares para uma comunidade de praticantes de uma ciência”⁹ (KUHN, 1990, p. 13). Isto é, “cada revolução científica altera a perspectiva histórica da comunidade que a experimenta, então esta mudança de perspectiva deveria afetar a estrutura das publicações de pesquisa e dos manuais do período pós-revolucionário” (KUHN, 1990, p. 14).

Morin (2010, p. 52), por sua vez, utiliza tanto o conceito de paradigma como o de revolução científica a partir de um viés que julgamos pertinente. Para o autor, “mesmo na atividade da ciência dita normal, existe uma revolução científica permanente”. Esta concepção de paradigma ensinada por Morin deixa claro uma perspectiva um pouco mais suave, se assim é possível referir, em relação à definição proposta por Kuhn, uma vez que estabelece uma contínua movimentação da ciência em busca de novos caminhos que permitam transformação, abandonando a ideia “um pouco tola, um pouco ingênua, do progresso linear das teorias que se aperfeiçoam mutuamente” (MORIN, 2010, p. 52). Além disso, ressalta que os paradigmas não são estanques e oriundos de um estado puro, mas de um modelo conceitual plenamente em transformação aceito por uma comunidade científica. É essa movimentação da ciência que se pretende mostrar em relação à e-Science na próxima seção.

2.1 e-Science: dados científicos para pesquisa

O processo de colaboração, principalmente nos últimos anos, tem se intensificado com as transformações ocorridas na ciência e a facilidade de acesso a outros pesquisadores por meio das tecnologias de informação e comunicação.

Gray (2009) observa e descreve uma (r)evolução na ciência e o modo de fazê-la, indo ao encontro do que Ronan (1987, p. 14) relata sobre o nascimento empírico da ciência e sua evolução desde povos primitivos, os quais a faziam através da magia em uma tentativa de racionalizar “fatos observados [...] e algumas

⁹ O conceito de paradigma (a este aliadas as ideias de valores, de crenças e de técnicas) pode contribuir “bastante para proporcionar aos especialistas [...] um sentimento de pertencermos a uma comunidade” (KUHN, 1990, p. 229).

observações e explicações que seriam gradualmente reunidas, vindo a produzir, um dia, uma visão não-mágica”. Após as primeiras ideias sobre astronomia no Egito, a invenção da escrita na Mesopotâmia, a matemática grega etc., passou-se para uma ciência capaz de criar e se utilizar de modelos e generalizações para gerar teorias sobre os fatos do mundo para, posteriormente, ser apta a simular fenômenos complexos através de computadores. Chegamos, por fim, à exploração de dados, ou seja, possibilitar que “os dados sejam capturados por instrumentos ou gerados por simulações antes de serem processados pelo software e para a informação resultante ou o conhecimento ser armazenado em computadores”¹⁰ (GRAY, 2009, p. xix). Este panorama é apresentado, em forma de síntese, na Figura a seguir.

Paradigmas da ciência

- Milhares de anos atrás:
ciência era **empírica**
descrições de fenômenos naturais
- Últimos cem anos:
ramo **teórico**
utilização de modelos, generalizações
- Últimas décadas:
ramo **computacional**
 - Simulação de fenômenos complexos
- Hoje: **exploração de dados** (e-Science)
 - Unificação de teoria, experimento e simulação
 - Dados capturados por instrumentos ou gerados por simulação
 - Processados por software
 - Informação/conhecimento armazenado em computador
 - Cientista analisa base de dados/arquivos usando gerenciamento de dados e estatística



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$

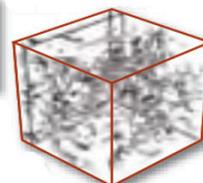


Figura 1: Paradigmas da ciência

Fonte: Traduzido de GRAY (2009, p. xviii)

¹⁰ No original: “[...] the data to be captured by instruments or generated by simulations before being processed by software and for the resulting information or knowledge to be stored in computers”.

Neste espaço que surge a e-Science como novo paradigma, isto é, a exploração de dados científicos possibilitando a sua coleta ou sua geração com armazenamento em sistemas computacionais capazes de permitir o compartilhamento para a produção de conhecimento. Podemos dizer que isto consiste em uma virada epistemológica ao observar como alguns estudiosos, como Le Coadic (2004), enfatizam a importância da informação – e não dos dados – na pesquisa científica. Diz o autor (2004, p. 26):

As atividades científicas e técnicas são o manancial de onde fluem os conhecimentos científicos e técnicos que se transformarão, depois de registrados, em informações científicas e técnicas. Mas, de modo inverso, essas atividades só existem, só se concretizam, mediante essas informações. A informação é a seiva da ciência. Sem informação, a ciência não pode se desenvolver e viver. Sem informação a pesquisa seria inútil e não haveria o conhecimento. (LE COADIC, 2004, p. 26).

Utilizamos, para maior clareza, o modelo de publicações científicas de Garvey e Griffith (1972), demonstrando como a comunicação científica se apresentava quando considerada baseada exclusivamente na produção bibliográfica:

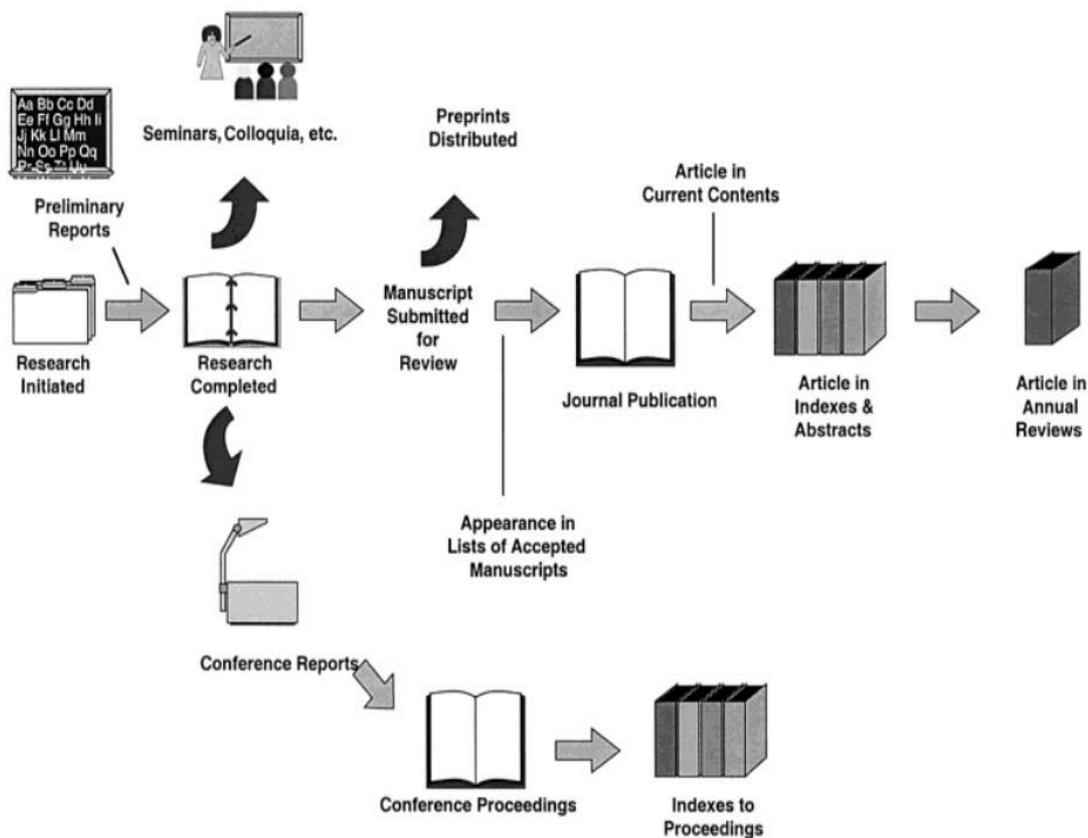


Figura 2: Modelo de publicações científicas de Garvey-Griffith

Fonte: HURD (2000)

Podemos notar que o modelo de comunicação científica mostrado na Figura 2 está voltado para a comunicação impressa e bibliográfica. Na época em que o estudo de Garvey e Griffith foi conduzido e publicado, as máquinas de cálculo e os próprios computadores recém começavam a fazer parte do cotidiano de pesquisadores e a comunicação entre eles era incipiente na utilização das tecnologias da informação e da comunicação como um intercessor. Percebemos que, deste modo, o processo é focado na comunicação científica através do papel.

Esse modelo descreve a disseminação da informação científica desde o momento de sua inicialização até a publicação dos resultados, seja em periódicos ou em anais de eventos – e posteriormente em documentos secundários do tipo

índices, resumos e *annual reviews* – revelando (e separando) os canais formais e informais da comunicação científica¹¹.

Mas a massiva entrada das tecnologias da informação e da comunicação não são contempladas pelo modelo de Garvey e Griffith. Isso fez com que Hurd propusesse, em 2000, um modelo de comunicação científica para o ano de 2020. Com a ressalva de que o modelo proposto já parece estar em vigor, assumimos como válido para os propósitos deste estudo, destacando pontos relevantes para o prosseguimento do trabalho.

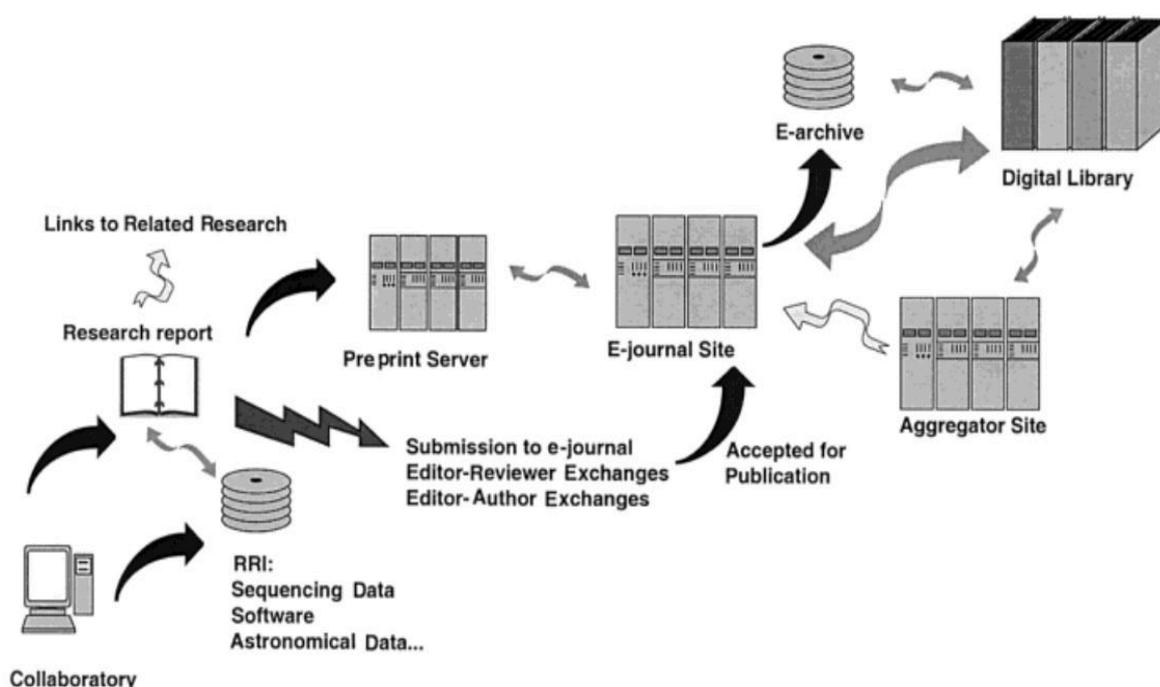


Figura 3: Modelo de comunicação científica para 2020

Fonte: HURD (2000)

Afora a importante atualização proposta pela autora, em seu modelo se destaca a RRI (*research-related information*), uma vez que este processo traz efeitos no que se refere à colaboração entre pesquisadores e entre instituições, fortalecendo os colégios invisíveis e permitindo que “repositórios de dados abertos a

¹¹ Isto está próximo do que ocorre no compartilhamento de ativos de pesquisa, como anota Borgman (2010, p. 123): “Some data in the research commons are available through repositories, while others remain with the researchers who generate them. Scholars often acquire data from each other privately, whether through contracts, research collaborations, or on an exchange basis.”

todos com conexões de rede possam contribuir para o desenvolvimento da ciência em locais que antes eram marginalizados”¹² (HURD, 2000, p. 1283).

Onde uma vez os membros de um colégio invisível poderiam ter usado os telefones e faxes [...], agora *research-related information* (RRI), como sequências genéticas ou depósitos de dados brutos são armazenados em servidores para acesso por outras pessoas que desejam construir sobre os resultados ou incluí-los em suas próprias análises.¹³ (HURD, 2000, p. 1281).

Gray (2009) destaca essa nova forma de fazer ciência, dizendo que ela ocorre porque os diferentes tipos de dados geram grandes quantidades de conteúdo. O autor ressalta que os cientistas em Astronomia, por exemplo, não mais se utilizam de telescópios, mas analisam dados estocados em grandes centros ou bases a partir de computadores pessoais, fazendo com que ciência e tecnologia atuem de forma estreita na produção, gerenciamento, acesso e uso. Ou seja, compreende que a “possibilidade de tratar dados, ao invés de ter que coletá-los novamente, possibilita que esforços sejam focados na análise dos dados já existentes e compartilhados, reduzindo custos e esforços e possibilitando avanços efetivamente mais rápidos” (MEDEIROS; CAREGNATO, 2012).

Entendemos que, de fato,

a e-Science promove algo que está ligado à constituição da ciência, isto é, maior possibilidade de cooperação, colaboração e interdisciplinaridade. Necessita-se, entre outros quesitos, de esforços e estudos na forma como os dados são liberados, podendo ser dados fortemente estruturados ou ainda que passaram por um processo de curadoria. (MEDEIROS; CAREGNATO, 2012, p. 320).

¹² No original: “data depositories open to all with network connections may contribute to the development of science in locations that were formerly marginalized”.

¹³ No original: “Where once members of an invisible college might have used telephones and faxes [...], now research-related information (RRI) such as genetic sequences or depositories of raw data are stored on servers for access by others who wish to build on the findings or include them in their own analyses”.

O Quadro 1 mostra um resumo da comparação que pode ser estabelecida entre as características da e-Science e da pesquisa tradicional, como a grande dependência entre ciência e tecnologia, bem como a mudança na atuação dos pesquisadores.

Quadro 1: Características da e-Science x Características da pesquisa tradicional

Característica	e-Science	Pesquisa tradicional
Participantes	Diversamente qualificados, equipe de pesquisa distribuída	Pesquisador individual ou pequena equipe local de pesquisa
Dados	Gerados, armazenados e acessíveis de localizações distribuídas	Gerados, armazenados e acessíveis apenas localmente
Computação e Instrumentação	Larga escala ou sob demanda ou acesso à informação compartilhada	Emprego da computação em lote ou emprego do computador ou instrumentos do próprio pesquisador
Rede	Emprego da internet e de mediadores	Independente da internet
Disseminação da pesquisa	Via websites e portais especializados	Via publicações impressas ou apresentações em conferências

Fonte: Traduzido de Appelbe e Bannon (2007, p. 84)

2.2 Dados digitais de pesquisa e infraestrutura de compartilhamento

Devemos tomar espaço inicial nesta seção para fazer uma diferenciação entre dado, informação e conhecimento. Acreditamos que esta sinalização colabora na compreensão tanto da importância quanto da função dos dados científicos nas implicações na maneira como os repositórios de dados digitais tratam seus insumos e produtos e, também, em aspectos da autoria.

Dados podem ser considerados como “fato' ou 'átomo', um elemento bruto, desprovido de significação imediata. Algo perceptível e de baixo teor semântico” (SEMIDÃO, 2014, p. 71), ou seja, “a menor representação convencional e fundamental de uma informação [...] sob forma analógica ou digital passível de ser submetida a processamento manual ou automático” (CUNHA, 2008, p. 112-113). Arruma e Chagas (2002) dizem que dado é um elemento obtido através de coleta adequada e, como fato bruto, pode ser moldado para ser convertido em informação. Esta ideia é corroborada e ampliada por Abbagnano (2007, p. 230), o qual teoriza que dado é

o elemento, o antecedente, a situação da qual se parte ou que serve de respaldo para formular um problema, fazer uma inferência, aventar uma hipótese. O D. tem, portanto, caráter funcional: o que se assume como D. para certo tipo ou ordem de indagação pode ser, por sua vez, tomado como problema para outro tipo ou ordem de pesquisa. (ABBAGNANO, 2007, p. 71).

Informação, em linhas gerais, “é um conhecimento inscrito (registrado) em forma escrita (impressa ou digital), oral ou audiovisual, em um suporte” (LE COADIC, 2004, p. 4). Le Coadic segue em sua elucidação dizendo que a informação carrega um componente de sentido, de significado, transmitida por um suporte a uma entidade capaz de decodificar signos em que a informação foi codificada.

A informação, como aponta McGarry (1999), depende de um elemento de surpresa, no sentido que há necessidade de haver uma novidade em algo para que os mapas cognitivos sejam modificados e para que possamos, efetivamente, sermos informados, reduzindo incertezas. Assim, informações são, necessariamente, dados ordenados e estruturados de alguma forma, fazendo que contenham sentido e se

tornem utilizáveis. Sendo assim, “o objetivo da informação [... é] a apreensão de sentidos ou seres em sua significação, ou seja, [...] o conhecimento” (LE COADIC, 2004, p. 5).

Conhecimento, por seu turno, pode ser “representado por uma estrutura de conceitos ligados por relações: nossa imagem de mundo” (LE COADIC, 2004, p. 8). Assim sendo, está ligado à experiência do indivíduo, “podendo ser tácito (não codificado) ou explícito (codificado, neste caso, tornando-se informação)” (CIANCONI, 2003, p. 29).

Setzer (1999) busca uma síntese sobre o que compreende por dado, informação e conhecimento, como reproduzimos abaixo:

Dado: seqüência de símbolos quantificados ou quantificáveis. [...] Como são símbolos quantificáveis, dados podem obviamente ser armazenados em um computador e processados por ele. [...] **Um dado é puramente objetivo - não depende do seu usuário.**

Informação: abstração informal (isto é, não pode ser formalizada através de uma teoria lógica ou matemática), que representa algo significativo para alguém através de textos, imagens, sons ou animação. [...] Não é possível processar informação diretamente em um computador. Para isso é necessário reduzi-la a dados. A representação da informação pode eventualmente ser feita por meio de dados. Nesse caso, pode ser armazenada em um computador. Mas, atenção, o que é armazenado na máquina não é a informação, mas a sua representação em forma de dados. Essa representação pode ser transformada pela máquina - como na formatação de um texto - mas não o seu significado, já que este depende de quem está entrando em contato com a informação. Por outro lado, dados, desde que inteligíveis, são sempre incorporados por alguém como informação, porque os seres humanos (adultos) buscam constantemente por significação e entendimento. [...] **A informação é objetiva-subjetiva no sentido que é descrita de uma forma objetiva (textos, figuras, etc.), mas seu significado é subjetivo, dependente do usuário.**

Conhecimento: abstração interior, pessoal, de alguma coisa que foi experimentada por alguém. [...] não pode ser descrito inteiramente - de outro modo seria apenas dado (se descrito formalmente e não tivesse significado) ou informação (se descrito informalmente e tivesse significado). Também não depende apenas de uma interpretação pessoal, como a informação, pois requer uma vivência do objeto do conhecimento. Assim, quando falamos sobre conhecimento, estamos no âmbito puramente subjetivo do homem ou do animal. [...] **O conhecimento é puramente subjetivo - cada um**

tem a experiência de algo de uma forma diferente. (SETZER, 1999, grifo nosso).

Se pudermos considerar o conhecimento, como diz McGarry (1999, p. 30), “como algo teórico e mais generalizado, e informação como conhecimento potencial”, podemos levar adiante a ideia de que os dados científicos também se revestem de conhecimento. Ou seja, compreendendo que existe clara relação em cadeia entre dado, informação e conhecimento, podemos assumir que não apenas a informação gera novos conhecimentos, mas também os dados científicos os geram.

A compreensão de que os dados são matérias-primas para a formulação de teorias a partir de análise é aqui de suma importância, já que sua investigação permite a constatação e criação de novas hipóteses, uma vez que parte da coleta de determinada porção de mundo, servindo como insumo para que novos cenários sejam criados e desvendados cientificamente. Dados de pesquisa são um tipo de capital científico, advindo de um novo modelo de publicação eletrônica que surge nesse espectro.

Dada a importância e dificuldade da definição, caracterizamos dados científicos, de maneira ampla, recorrendo à definição de Torres-Salinas, Robinson-García e Cabezas-Clavijo (2012):

[...] todo aquele material registrado durante uma pesquisa, reconhecidos pela comunidade científica e que servem para certificar os resultados alcançados. Além destas características, também se especificam que devem provir de uma fonte única e devem ser difíceis ou impossíveis de se obter novamente por serem próprios de um momento ou circunstâncias irreplicáveis de uma forma exatamente igual.¹⁴ (TORRES-SALINAS; ROBINSON-GARCÍA; CABEZAS-CLAVIJO, 2012, p. 175).

¹⁴ No original: “todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza. Además de estas características también se especifica que debe provenir de una fuente única y deben ser difíciles o imposibles de obtener de nuevo por ser propios de un momento o circunstancias irrepetibles de una forma exactamente igual”.

Outra definição, apresentada pela Organisation for Economic Co-operation and Development (OECD), diz que dados de pesquisa são:

registros factuais (números, registros textuais, imagens e sons) utilizados como fontes primárias para a pesquisa científica, e que são geralmente aceitos na comunidade científica como necessários para validar os resultados da pesquisa. Um conjunto de dados de pesquisa constitui uma representação sistemática, parcial do assunto que está sendo investigado.¹⁵ (OECD, 2007, p. 13).

De acordo com Borgman (2010), os próprios pesquisadores apresentam dificuldades em manifestar o que são dados científicos, uma vez que sua tipologia é variada. Talvez esta distinção seja o primeiro passo – e um dos mais importantes – para que se possa proporcionar uma infraestrutura de dados científicos. Por exemplo, em Medicina podem ser raios-x, em Química, as estruturas de proteínas, e em Biologia, os espécimes. Pelas tipologias apresentadas, parece-nos claro que a maioria dos dados são criados por e para propósitos de pesquisa, nascidos, atualmente, em formato digital, o que faz com que sejam vistos como parte de um aglomerado que inclui software, equipamento, documentação e conhecimento requeridos para usá-los.

Cabe ressaltar que, neste sentido,

[...] decidir o que/quais são os dados é ainda mais difícil nas ciências sociais do que nas ciências puras, devido ao volume e à diversidade de fontes de dados e a falta de controle sobre as fontes criadas por terceiros. [...] Os cientistas sociais sempre estão à procura de novos dados e novos métodos para coletar e analisar dados.¹⁶ (BORGMAN, 2010, p. 204).

¹⁵ No original: “factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated”.

¹⁶ No original: “deciding what are data is even more difficult in the social sciences that in the sciences due the volume and diversity of data sources and the lack of control over sources created by third parties. [...] Social scientists always are looking for new data, and new methods to collect and analyze data”.

A National Science Foundation (2007), em seu relatório *Cyberinfrastructure Vision for 21st Century Discovery*, aponta a necessidade de planos de ação para que a e-Science seja viável. São eles:

- (i) computação de alta performance, com sistemas computacionais provedores e capazes de lidar com quantidades gigantescas de dados;
- (ii) dados, análise de dados e visualização, com investimentos nos próprios dados e no processo de sua geração e compartilhamento, bem como em sistemas de representação capazes de promover esse acesso, como metadados e ontologias;
- (iii) organizações virtuais para comunidades distribuídas, necessitando da aplicação de ferramentas e tecnologias que suportem colaboração entre pesquisadores;
- (iv) desenvolvimento de aprendizado e força de trabalho, onde a capacitação profissional para essa nova infraestrutura deve gerar trabalho para suporte, desenvolvimento, estudos de usuários, classificação, design etc.

Isso nos permite argumentar que “agora é o momento para determinar o que deve ser construído”¹⁷, uma vez que “estes são tempos excitantes e confusos para a comunicação científica”¹⁸ (BORGMAN, 2012, p. 2). Estas citações expressam, de modo amplo, a forma como podemos entender o momento da e-Science e do compartilhamento de dados para com a comunicação científica: são desafios que vão desde o entendimento do compartilhamento por parte dos cientistas, passando pela (não) utilização de padrões de metadados já estabelecidos em determinadas áreas do conhecimento (BORGMAN, 2010), bem como a compreensão da validade e importância dos dados como insumo de pesquisa¹⁹ e a implicação e a relevância da autoria de dados de pesquisa.

Em suma, o compartilhamento dos dados de pesquisa diz respeito ao ato de liberar dados para o uso de outros pesquisadores, podendo ser realizado de forma

¹⁷ No original: “now is the time to determine what we should be building”.

¹⁸ No original: “these are exciting and confusing times for scholarship”.

¹⁹ Prova da importância dos dados científicos para a ciência é a criação, pela Thomson Reuters, do Data Citation Index (DCI) em novembro de 2012. Utilizado como ferramenta para descoberta e acesso a uma variedade de repositórios de três grandes áreas (Ciência e Tecnologia; Ciências Sociais; Artes e Humanidades), passa por um crivo qualitativo para que a base possa ser incluída no DCI. Os registros existentes na base DCI estão linkados às publicações que os utilizaram – quando informado – permitindo informações sobre a citação do conjunto de dados a que fazem referência (TORRES-SALINAS; MARTÍN-MARTÍN; FUENTE-GUTIÉRREZ, 2013).

privada ou através de grandes repositórios. De forma privada ocorre com o envio de dados através de serviços que restringem o acesso aos dados, fazendo com que estes circulem dentro de uma comunidade específica, utilizando-se do serviço de e-mail ou outro serviço pessoal. A disponibilização de dados em repositórios possibilita que uma grande variedade de pesquisadores com os mais diversos níveis de especialização e que estejam espalhados pelo mundo possam se utilizar destes ativos, fazendo reuso em suas pesquisas.

Como vantagens do compartilhamento público de ativos digitais, estão: (a) maior aproveitamento de recursos financeiros investidos em ciência, já que o reuso em novas análises faz com que não haja necessidade de que novas coletas sejam financiadas para alcançar objetivos semelhantes; (b) prevenção de fraudes, possibilitando a execução de experimentos e verificação de hipóteses; e (c) crescimento das citações dos autores desses dados (TORRES-SALINAS, ROBINSON-GARCÍA E CABEZAS-CLAVIJO, 2012).

Apesar de contar com as vantagens mostradas acima, o compartilhamento de dados de pesquisa enfrenta questões ainda não delimitadas. A primeira é o próprio conjunto de dados, uma vez que ele passa a ser tanto o insumo como o produto da pesquisa, isto é, a utilização de dados faz com que novos dados sejam gerados, e estes devem ser compartilhados da mesma maneira que os primeiros. Isto significa dizer que os dados estão ligados aos documentos bibliográficos publicados, criando certa tensão entre dados científicos e publicações bibliográficas, mas delegando aos pesquisadores a compreensão de que “simulações, repositórios de dados e outros conteúdos complexos, com links interativos podem ser considerados publicações, especialmente em campos que utilizam dados intensivamente”²⁰ (BORGMAN, 2010, p. 99).

Outra questão que podemos assumir é em relação à curadoria de dados científicos. A noção de curadoria é ainda imprecisa na literatura, mas é possível admitir que “tanto a gestão quanto o estabelecimento de políticas de preservação fazem parte de um leque de atividades que são englobadas pelo conceito de curadoria digital” (SALES, 2014, p. 54). Mas como Borgman (2010, p. 117) afirma

²⁰ No original: “simulations, data repositories, and other “complex content with interactive links may be considered publications, especially in data-intensive fields”.

com propriedade, “dependendo da área e do projeto, uma publicação pode apresentar vários conjuntos de dados, ou um conjunto de dados pode ser usado em várias publicações”²¹ (BORGMAN, 2010, p. 117), o que faz com que pensemos a curadoria dos dados como uma rede complexa de elementos que interfere de forma significativa na (re)utilização dos dados, gerindo e possibilitando o processo de compartilhamento de dados de início a fim.

A curadoria digital, em síntese, garante a sustentabilidade dos dados para o futuro, não deixando, entretanto, de conferir valor imediato a eles. Os recursos estratégicos, metodológicos e as tecnologias envolvidas nas práticas da curadoria digital facilitam o acesso persistente a dados digitais confiáveis por meio da melhoria da qualidade desses dados, do seu contexto de pesquisa e da checagem de autenticidade. Dessa forma, a curadoria contribui para assegurar a esses dados validade como registro arquivístico, garantindo que eles possam ser usados no futuro como evidência legal. (SALES, 2014, p. 56).

Outro ponto que sobre o qual devemos buscar clareza, refere-se ao acesso aos dados de pesquisa. Sabendo que os recursos que fomentam pesquisas, especialmente no Brasil, “são oriundos dos cofres públicos, permitir que os resultados fiquem apenas nas mãos de editoras comerciais é não oferecer a quem financiou – em última instância, os cidadãos – os direitos de acesso a esses resultados” (SILVA; TOMAÉL, 2008, p. 124). Ou seja, a criação de repositórios de informação científica – em qualquer formato – ocorrida principalmente a partir dos anos 2000, visa compreender um deslocamento do sistema de comunicação científica, uma vez que, oriundo de financiamentos públicos, o conhecimento científico deve ser comunicado de forma livre para acesso universal e geração de novos conhecimentos.

A atenção neste caso está voltada ao propósito do repositório. Enquanto o repositório de informação visa a disseminação de obras textuais, imagéticas, entre outras, o repositório de dados, como o próprio nome diz, visa o compartilhamento de dados, o que faz com que ele possua características próprias advindas, inclusive, do

²¹ No original: “The lack of a direct relationship between documents and data sets also hampers linking. Depending on the field and the project, one publication can draw on multiple data sets, or one data set can be used in multiple publications”.

tipo de tratamento que esses dados e metadados recebem, ou seja, pelo próprio tipo de gestão, de curadoria.

Assim sendo, é importante que existam repositórios de acesso público, uma vez que

[...] são importantes ferramentas que facilitam o compartilhamento das informações de natureza científica e acadêmica. Portanto, a implementação de repositórios digitais, ou repositórios de informação é uma das formas que as universidades – ou comunidades científicas das áreas temáticas – dispõem para minimizar a falta de visibilidade de sua produção intelectual, porque favorecem o equilíbrio de forças no contexto da comunicação científica. (SILVA; TOMAÉL, 2008, p. 128).

Podemos elencar, com base em Sales (2014, p. 63-64), as vantagens da utilização de repositórios de dados digitais de pesquisa:

- (a) amplia a visibilidade dos resultados de pesquisa, posto que, via de regra, só é formalmente disseminada a fração desses resultados que está registrada nas publicações acadêmicas;
- (b) proporciona mecanismos de preservação de longo prazo, em termos de preservação digital, de arquivamento seguro e de curadoria digital;
- (c) permite que o material depositado esteja disponível on-line continuamente para ser consultado e citado mais frequentemente;
- (d) é um instrumento chave para os processos de reformatação e recriação de dados proporcionados pela curadoria digital;
- (e) abre a possibilidade de criação de novos serviços de informação para pesquisadores e gestores a partir da análise dos dados arquivados e a integração de dados e publicações acadêmicas;
- (f) permite a criação de redes de repositórios interoperáveis;
- (g) aumenta o grau de reuso dos dados minimizando a duplicação de esforços e otimizando os investimentos na geração de dados.

A existência de uma infraestrutura tecnológica – essencial para que o compartilhamento de dados aconteça – mantém, também, uma infraestrutura social,

uma vez que a “e-Research não surgiu de um vácuo político”²² (BORGMAN, 2010, p. 45). Deste modo, “para construir uma infraestrutura eficaz, precisamos entender mais sobre as práticas, comportamentos, políticas e estruturas institucionais dos campos”²³ (BORGMAN, 2010, p. 43-44), provendo um estudo que envolva tanto os atores humanos quanto não-humanos que atuam no processo de autoria e têm suas características ligadas ao valor dos dados para a pesquisa científica.

Devemos refletir, além disso, sobre a relação que existe entre o documento bibliográfico como resultado e a publicação de pesquisas e os dados científicos. Sabendo que os dados são essenciais para o trabalho científico, mesmo que sua publicação possa ser feita separadamente de artigos, há grande disparidade de estudos no que se refere a esta relação. Esta observação pode ser constatada pela literatura disponível para ambos os casos. Enquanto a literatura que trata sobre material bibliográfico enfoca questões sociais, comportamentais, políticas etc. a partir de centenas de anos de estudos, os trabalhos voltados a dados são recentes e estão focados em perspectivas etnográficas, técnicas e políticas. Além disso, os tipos de dados e suas variações de uso entre os diferentes campos científicos fazem com que apenas aspectos gerais sejam identificados (BORGMAN, 2010).

Ao lado disto, devemos tomar em consideração também a validação dos dados. Neste aspecto, a publicação parece referendar o conjunto de dados, isto é, sua validade está ligada, em geral, à sua apresentação em material bibliográfico (BORGMAN, 2010). Esta é outra questão que é afetada pela autoria de dados de pesquisa e material bibliográfico produzido com base nos dados.

O que se pretende deixar claro é a necessidade de que os dados sejam compartilhados em repositórios próprios para este fim, buscando o máximo de alcance possível no seu reuso. São os estudos provenientes da reutilização destes dados que poderão estabelecer novos cenários científicos. É neste sentido que este estudo foi desenvolvido, buscando as teias que sejam tecidas com os elementos que compõem a autoria no compartilhamento de dados científicos.

²² No original: “e-Research did not emerge from a political vacuum”.

²³ No original: “to build an effective infrastructure, we need to understand more about the practices, behaviors, policies, and institutional frameworks of the fields”.

3 AUTORIA

“A autoria é uma instituição em crise” (MIRANDA; SIMEÃO; MUELLER, 2007, p. 35).

Este capítulo trata a questão da autoria e os aspectos relacionados a essa noção, compreendendo que o arcabouço teórico, conceitual e pragmático que a compõe apresenta dimensões que tornariam impossível abordar de forma separada o autor como figura única e central em um processo que é vívido e capaz de aglutinar tantos atores-rede. Isto é, sendo a autoria “uma questão subjacente à ciência, à ética e à legislação, e que afeta e envolve produtores, comunicadores, cientometristas e consumidores” (WITTER, 2010, p. 131-132), torna-se complexa e muitas vezes difusa. Assim sendo, antes de prosseguir nesta seção, cabe reconhecer pontos que tornam dificultosa a tarefa de lidar com a noção de autoria.

Em primeiro lugar no que se refere à própria noção, em relação à sua complexidade, foi possível constatar que a literatura busca explorá-la e explicá-la de maneira conturbada e plural, o que acarreta em diversas possibilidades de apresentar o tema, bem como certa incipiência. Isso se reflete nas propostas de textos que, ao buscarem tratar a autoria, fixam-se no autor como ente responsável por todo o processo. Pretendemos adiante tornar clara a separação que entendemos existir entre autor e autoria.

Em segundo lugar, a dificuldade está colocada em relação ao próprio objeto de nossa pesquisa. O que encontramos em trabalhos sobre autoria está relacionado, invariavelmente, a documentos textuais, gráficos, ficcionais, cinema, teatro etc., ou seja, a literatura busca o entendimento da autoria através do seu discurso, além do indivíduo autor, fazendo com que abordagens para uma explicação a partir de dados científicos digitais ainda não sejam relatadas. Advertências realizadas, seguimos a empreitada.

Em um amplo apanhado, as definições sobre autoria dispostas pela literatura buscam aclarar o entendimento sobre autor e autoria, basicamente, destacando o autor, ou seja, um indivíduo ou entidade que se responsabiliza pela criação

intelectual de determinado conteúdo explícito que constitui uma obra. Embora essa opacidade seja insuficiente para nosso objetivo, vemos nela um modo de iniciar o pensar sobre esta parte de nosso trabalho, já que, em uma tentativa de expandir essas definições, podemos compreender o autor como indivíduo/entidade e a autoria como a ação do autor.

Esperamos fazer ecoar neste ponto os argumentos levantados pela Teoria Ator-Rede, uma vez que o “processo de 'publicação' dos textos implica sempre uma pluralidade de espaços, de técnicas, de máquinas e de indivíduos” (CHARTIER, 2002, p. 64). Isto é, uma noção de que autoria está ligada não exclusivamente a uma entidade, mas a um conjunto de associações que se ligam e desligam de forma a construir uma rede (ator-rede).

Iniciamos, assim, uma tentativa de compreender a questão epistemológica da autoria – se assim podemos nos referir – visando trazer a lume a perspectiva levantada por diversos autores, bem como isso pode confluir na capacidade de articular fundamentos que julgamos essenciais a esse entendimento.

3.1 O direito autoral como fala inicial para a autoria

Historicamente, a ideia de uma ligação direta entre autor e originalidade não existiu entre a Idade Média e a modernidade, “seja porque era inspirada por Deus: o escritor não era senão o escriba de uma Palavra que vinha de outro lugar. Seja porque era inscrita numa tradição, e não tinha valor a não ser o de desenvolver, comentar, glosar aquilo que já estava ali” (CHARTIER, 1998, p. 31). É a partir do século XVII que há uma modificação de como se compreende a autoria no meio científico.

Nessa época, os discursos científicos que possuíam verdades sistematicamente demonstradas, podendo incluir sistemas teóricos e metodológicos organizados, detinham reconhecimento pela comunidade científica mesmo sem a indicação de autoria. Atualmente, a questão autoral na ciência envolve, em muitos casos, diversos autores (uma equipe) que trabalham em conjunto com a função de

produzir determinado conhecimento que pode servir de base para exploração, análise e produção de conhecimento, possibilitando verificar e validar métodos e resultados (ANTONIO, 1998).

Mas é importante relatar que não é a partir de uma visão do autor, proprietário de uma obra ou do vislumbre de uma classe burguesa que nasce uma definição de propriedade literária, mas, sim, através dos interesses de defesa dos direitos adquiridos pelos editores livreiros. Estes, em Londres no século XVI, tinham o direito de exploração das obras, e se utilizavam dessa arma para lutar contra os livreiros oriundos das províncias. Até essa época era utilizado o termo *right in copies*, direito registrado perpétuo e exclusivo sobre o manuscrito (CHARTIER, 2012a, 2012b).

Somente a partir de 1710 a *Lei para o Encorajamento do aprendizado por meio da concessão de direitos de cópia de livros impressos ao autor, ou compra de tais cópias*, fez com que os autores pudessem registrar suas próprias obras, obtendo *copyright*, por um período de 14 anos²⁴ renováveis por mais 14 anos em caso de autor vivo, e ainda editá-las. O contra-ataque dos livreiros de Londres foi inventar a propriedade literária, a qual buscava impor a ideia de que embora um autor fosse proprietário de um trabalho, a partir da cedência da obra a outra pessoa (no caso ao livreiro londrino), o indivíduo transmitia, juntamente com o trabalho, um direito que não prescrevia, além de ceder sua perpetuidade. Embora o estatuto tenha proporcionado uma nova definição posição para o autor, isso consistia, basicamente, em uma forma de reiteração de privilégios já estabelecidos (CHARTIER, 2012a, 2012b).

Isso nos permite compreender que um conceito fundamental no que concerne à autoria é o direito autoral sobre a propriedade intelectual. A criação da imprensa de Gutenberg parece ter alavancado a questão do direito autoral, embora, como diz Campos (1994), esse processo tenha apenas se instituído no século XVIII e contasse com certa vergonha do autor em vender suas ideias e reflexões a um editor.

²⁴ A limitação do *copyright* compreendia um período de 14 anos “porque era o limite da duração que cobria a exploração exclusiva de uma invenção técnica, de um procedimento, de uma máquina. [...] Tratava-se, assim, de equiparar o *copyright* a uma patente ou a um *brevet*, como se diz em francês, e, ao mesmo tempo, por parte dos livreiros londrinos e seus advogados, de mostrar que a obra literária não era identificável a uma invenção técnica porque ela era caracterizada como criação irredutivelmente singular e original, separada de toda materialidade” (CHARTIER, 2012a p. 45).

Assim sendo, a forma de exercer o direito sobre algo decorre, necessária e obviamente, da existência desse algo. Isto é, a propriedade intelectual advém da existência de algum bem, fruto de “transformação da natureza, (re)descobrimo-a ou esgotando-a, [aonde] o homem inicia o trabalho idealizando na mente a forma e o processo para ser realizado o objeto, e, em seguida, o abstrato é concretizado²⁵” (BARBOSA, 1999, p. 23).

O tema da propriedade intelectual vem de longa data, “longe de se originar de uma aplicação particular da definição burguesa do direito de propriedade, [estando ...] profundamente enraizada na defesa de privilégios dos livresiros” (CHARTIER, 2012a, p. 45). Isto é, está voltada à questão econômica, fazendo com que as entidades envolvidas pensem sobre os dividendos que podem gerar.

Os autores das criações intelectuais adquirem direitos exclusivos sobre suas produções, independente do valor técnico ou artístico que apresentem. Ao conjunto desses direitos resultantes do trabalho da inteligência e da criatividade dá-se o nome genérico de propriedade intelectual, ou ainda, *direitos intelectuais, direitos imateriais ou direitos sobre bens imateriais*.²⁶ (CARBONI, 2007, p. 60).

Como aponta Antonio (1998, p. 189), direito autoral diz respeito, essencialmente, “à propriedade intelectual ou artística sobre obras ou produtos. Basicamente, é o direito legal do autor ou criador de uma obra a controlar a reprodução e a distribuição dessa obra”. Santos (2009, p. 2) atua na mesma direção ao dizer que “a invenção e a expressão criativa, mais a proteção, são iguais à propriedade intelectual”.

Dessas assertivas podemos tirar duas proposições: (i) o direito autoral está ligado à propriedade de um indivíduo/entidade, isto é, fundamenta-se do direito natural, aonde sendo o homem o proprietário do seu corpo, qualquer trabalho por ele

²⁵ “Entretanto, deve ser ressaltado que o trabalho intelectual nem sempre é, em essência, idêntico entre si. Ele pode ser um fim em si mesmo, essencialmente subjetivo e intangível, ainda que concretizado; mas, também, pode ser uma parte do processo que culmina no bem tangível” (BARBOSA, 1999, p. 23).

²⁶ Carboni (2007, p. 60) alerta para a divisão de grupos na propriedade intelectual, uma vez que elas podem se dar no campo da técnica ou da estética: “As criações intelectuais no campo da técnica são protegidas por propriedade industrial e as manifestações do espírito no campo estético, pela propriedade literária, artística e científica, constituindo o chamado direito autoral”.

produzido implica na sua propriedade; e (ii) o direito autoral está ligado à questão intelectual, não tangível. Com isso, podemos notar que há um tipo de preservação do interesse na continuidade da informação e da evolução, uma vez que formular essas questões “visa não apenas garantir o respeito à propriedade imaterial presente, mas, sobretudo, incentivar a continuidade do trabalho de criação e o conseqüente desenvolvimento social que esta provoca” (CURIA, 2009, p. xv).

Podemos ainda colocar o direito autoral como fundamentalmente instaurado a partir de dois aspectos: o moral e o patrimonial. No que se refere ao primeiro, está ligado à garantia dada ao criador para “o controle à menção de seu nome na divulgação de sua obra e o respeito à sua integridade, além dos direitos de modificá-la ou retirá-la de circulação” (GANDELMAN, 2007, p. 33), isto é, está ligado ao direito natural, e, em tempos de tecnologia digital avançada, desconsidera o direito moral como “uma mera proteção de um valor individualista do sujeito-autor para passar a ser concebido como um interesse social no reconhecimento da identidade do emissor da mensagem” (CARBONI, 2007, p. 75).

Nesse sentido, os direitos morais permitem que o autor tome decisões, por exemplo, sobre “retirar sua obra de circulação, limitando os direitos dos editores” (BARBOSA, 1999, p. 27). São características dos direitos morais, por estes serem classificados como direitos de personalidade: (a) a pessoalidade (são irrenunciáveis); (b) a perpetuidade (sobrevivem ao próprio autor); (c) inalienabilidade (não são negociáveis); (d) a imprescritibilidade (não é possível perder o direito para outrem); e (e) a impenhorabilidade (se não podem ser negociados, não há possibilidade de penhora) (CARBONI, 2007; FRAGOSO, 2009).

No que se refere ao direito patrimonial, basicamente está ligado à utilização econômica da obra criada. Carboni (2007) e Fragoso (2009) elencam como características dos direitos patrimoniais: (a) o cunho real; (b) o caráter de bem móvel (para que seja disposto pelos meios possíveis); (c) a alienabilidade (para permitir o seu ingresso no comércio jurídico); (d) a temporalidade (são protegidos por um determinado prazo legal); (e) a penhorabilidade (a possibilidade de sofrer constrição judicial); e (f) a prescritibilidade (a perda da ação por inércia no prazo previsto em lei).

Embora o registro do direito autoral seja facultativo, Santos (2009, p. 4) leva em consideração que

a partir do momento em que o objeto autoral ganha expressão no mundo, ainda que por meio de uma simples execução, recitação ou rascunho, então será considerado autor aquele que deu ensejo a essa exteriorização, recaindo sobre sua obra toda a proteção da lei. (SANTOS, 2009, p. 4).

Nesse sentido, a Lei 9610²⁷, de 19 de fevereiro de 1998, normatiza, em seu artigo 1º, a regulação dos “direitos autorais, entendendo-se sob esta denominação os direitos de autor e os que lhes são conexos” (BRASIL, 1998). Esses ditos direitos conexos podem ser compreendidos como aqueles que têm alguma relação com a criação ou a disseminação da obra, estando, assim, também protegidos pelo direito autoral.

Ainda em relação à Lei 9610, o artigo 7º considera como obras intelectuais protegidas pela lei “as criações do espírito, expressas por qualquer meio ou fixadas em qualquer suporte, tangível ou intangível, conhecido ou que se invente no futuro”. Seu inciso XIII coloca as bases de dados como parte dessa criação intelectual: “coletâneas ou compilações, antologias, enciclopédias, dicionários, bases de dados e outras obras, que, por sua seleção, organização ou disposição de seu conteúdo, constituam uma criação intelectual”.

Após esses apontamentos sobre os elementos que correspondem minimamente ao direito autoral, apresentamos a visão de Manso (1987, p. 7-8) sobre direito autoral:

[...] direito autoral é o conjunto de prerrogativas de ordem *patrimonial* e de ordem *não patrimonial* atribuídas ao autor de uma obra intelectual que, de alguma maneira, satisfaça algum interesse cultural de natureza artística, científica, didática, religiosa, ou de mero

²⁷ A Lei 9.610 foi alterada pela Lei 12.853, de 14 de agosto de 2013, a qual altera os arts. 5º, 68, 97, 98, 99 e 100, acrescenta arts. 98-A, 98-B, 98-C, 99-A, 99-B, 100-A, 100-B e 109-A e revoga o art. 94 da Lei nº 9.610, de 19 de fevereiro de 1998, para dispor sobre a gestão coletiva de direitos autorais, e dá outras providências. Não altera, no entanto, as seções importantes para este trabalho.

entretenimento; que tais prerrogativas lhe são conferidas pelo simples fato de ser o *criador* daquele bem, independentemente, até, da existência de leis especiais que as proclamem; que tais prerrogativas consistem, em suma, num poder de utilização do seu produto intelectual, cabendo-lhe decidir se ele deve ou pode ser levado ao conhecimento do público em geral, ou de um público particular; de que maneira essa publicação será feita, para que fim pode dar-se sua publicação, e se a utilização autorizada deve ou não ser remunerada; ou, enfim, se jamais essa criatura haverá de ver a luz do dia, permanecendo inédita, não nascida, ignorada de todas, conservando-se coisa estritamente pessoal de seu autor. (MANSO, 1987, p. 7-8).

Assim, entendemos, à luz de Carboni (2007) e Fragoso (2009), que a expressão direitos autorais é mais abrangente e cumpre a função de englobar tanto os direitos dos autores quanto os que lhes são conexos.

Buscando colocar essa situação em termos acadêmicos, o pesquisador que lida com a tecnologia entra em um jogo de negociações com a sua comunidade e a sociedade como um todo para ser reconhecido como o criador de fato e de direito daqueles dados e/ou informações (BARBOSA, 1999). Isso permite refletir que os meios tecnológicos possibilitam uma reprodução acelerada de dados e/ou informações e que isso pode, ao longo de um processo de publicação acelerado e pressionado, levar em consideração as posições assumidas pelos atores para um sistema de atribuição de crédito sobre a propriedade.

Uma vez que “a ciência não é de imediato economicamente produtiva, embora sempre seja útil[, ...] a medida que cresce a acumulação científica, o seu estoque de conhecimentos aumenta o potencial de apropriação pela esfera econômica” (BARBOSA, 1999, p. 29). Isso é facilmente visualizado em empresas de diversos ramos que mantêm setores e desenvolvimento tecnológico que possibilitam registros de patentes.

Em outros termos, o âmbito acadêmico tem, há algum tempo, um espaço fundamental que comprova que a ciência e a tecnologia andam juntas para um processo de apropriação do ambiente econômico, já que decorre de uma série de proteções a capacidade do pesquisador ser reconhecido e responsabilizado sobre aquilo que trabalhou. Tal reconhecimento é de vital importância que, segundo Barbosa (1999), a Organização Mundial da Propriedade Intelectual na Organização

das Nações Unidas (OMPI/ONU) elaborou o Tratado Internacional de Registro de Descobertas Científicas em 1978²⁸.

Em tempos de mudança tecnológica, como a criação da imprensa de Gutenberg, a revolução industrial e o desenvolvimento da internet, novas formas de lidar com o direito autoral têm surgido (BRANCO JÚNIOR, 2011), como o caso das licenças públicas. Este tipo de licença surge a partir do momento em que percebemos que “modelos de proteção construídos a partir do final do século XIX são hoje insuficientes para dar conta da revolução tecnológica que vivemos” (BRANCO JUNIOR, 2011, p. 234), e novos modelos se tornaram necessários para lidar com uma (r)evolução ainda desconhecida, fazendo com que, para além do interesse dos autores e editores, a sociedade possa se beneficiar do conhecimento existente.

Esses novos modelos são oriundos, basicamente, de ordens sociais, econômicas e jurídicas. No primeiro caso, a questão é prover o acesso da sociedade à informação que foi gerada pela própria sociedade, garantindo-se a matéria-prima para geração de novos conhecimentos. No aspecto econômico, embora seja completamente compreensível a necessidade do autor e seus entes deterem os direitos sobre a obra durante certo período de tempo para fins de sustento, o excesso no período de restrição faz com que o público seja destituído de uma concorrência entre produtores de novos conhecimentos e/ou de melhores traduções, por exemplo, sem contar o fator financeiro que tem impacto direto na quantidade de opções que o mercado oferece. Do ponto de vista jurídico, compreende-se que durante o prazo de proteção do direito agem as limitações impostas pela lei. Transcorrido o prazo, a obra entra em domínio público, ou seja, o domínio público é o destino de qualquer obra criada, sendo esse domínio a regra e a limitação a exceção (BRANCO JÚNIOR, 2011).

[...] um autor, por um motivo qualquer (porque entende que é inútil proteger sua obra – que carece de importância econômica, ou porque prefere vê-la difundida para auferir lucros por meio de outras modalidades de negócio), deseja que sua obra seja simplesmente copiada pelos usuários da internet, não basta apenas não coibir a reprodução. Quem copia obra na íntegra, ainda que o autor nada

²⁸ Após elaborado, o Tratado não entrou em vigor devido a assinatura de apenas cinco países.

faça para impedir a cópia, viola direitos autorais. Por isso, tornou-se necessário que o autor consinta expressamente com a reprodução de sua obra. (BRANCO, 2011, p. 234).

Esse tipo de licença, considerada pública, atua a partir da ideia de que a comunidade em que determinado conteúdo está inserido deve ter acesso irrestrito a algo sem qualquer necessidade de permissão para tal. Assim sendo, quando falamos de licenças públicas gerais, estamos argumentando que elas possibilitam e facilitam a cópia integral ou parcial de um trabalho sem a necessidade de autorização prévia do autor.

Dentre essas licenças estão, por exemplo, o *copyleft* e o Creative Commons. A primeira utiliza um tipo de sistema parecido com o *copyright* para

garantir que todos que recebam sua versão da obra possam usar, modificar e distribuir tanto a obra original quanto as suas versões derivadas[, ...] requer[endo] que suas modificações ou extensões do mesmo, sejam livres, passando adiante a liberdade de copiá-lo e modificá-lo novamente. (SANTOS, 2009, p. 137).

No que se refere ao Creative Commons, a ideia foi desenvolvida por Lawrence Lessig na Universidade de Stanford com a finalidade de “expandir a quantidade de obras criativas disponíveis ao público, permitindo criar outras obras sobre elas, compartilhando-as” (SANTOS, 2009, p. 138), viabilizando que licenças flexíveis, a partir da tríade “todos direitos reservados”, alguns direitos reservados” ou “nenhum direito reservado”, sejam assumidas e/ou mescladas para cada objeto a ser disponibilizado.

Procuramos com essa introdução ao direito autoral mostrar que a discussão e o entendimento sobre o que vem a ser a noção de autoria nascem da contenda dos autores sobre os direitos de suas criações. Permitimo-nos estabelecer, desse modo, que a luta dos indivíduos por seus direitos autorais ou propriedades intelectuais derivam de duas vertentes, sendo a primeira mais óbvia, ou seja, os lucros relativos aos direitos sobre as vendas exclusivas de suas obras, e a segunda tendência nasce do reconhecimento pela produção intelectual, confeccionada através de

esforço, trabalho, criatividade etc. Esses aspectos fazem entender que, dependendo da natureza do trabalho e da percepção do autor sobre a possível utilidade da sua obra, não somente os lucros financeiros sejam almejados. Na próxima seção buscamos aclarar o que a literatura nos fornece para trabalharmos com a ideia de autoria.

3.2 A autoria: elementos de uma função

Nesta seção buscamos evidenciar como a autoria é apresentada na literatura, procurando indícios, traços e/ou elementos que possam colaborar com o objeto de nosso estudo. Para tal, trazemos para nosso lado da investigação, principalmente, dois autores: Michel Foucault e Roger Chartier. Ainda que reconheçamos que nenhum dos dois esgote a questão autor/autoria em seus trabalhos, entendemos que eles disponibilizam grande aporte para compreensão dos elementos que constituem a autoria.

O próprio Michel Foucault assume que trabalhar a questão da autoria é por demais complexo. Neste sentido, busca compreender o que está por trás deste processo, aquilo que não foi explicitamente dito. Ora, o que se busca aqui é a compreensão do que está sendo fechado em caixas-pretas a partir de uma maneira racionalmente relatada, como destaca Bruno Latour, através das associações que estão postas e formam o coletivo autoria. “Trata-se, portanto, de considerar o autor como função variável e complexa” (CHARTIER, 2012b, p. 27).

Na introdução à obra *O que é um autor? – A lição de Foucault* –, José Bragança de Miranda e António Fernando Cascais relatam a importância de construir o que vem a ser um autor, isto é, não estabelecer uma categoria autor, uma vez que criar categorias implica em construir definições, mas explicitar sua constituição a partir da reunião de elementos. Assim, o que se prospecta é entender a composição do que vem a ser esse elemento, o qual exige explicação, uma vez que as ligações socialmente (ou coletivamente) construídas por esse fenômeno não perpassam apenas ao autor como indivíduo, mas a autoria como processo que envolve atores (atores-rede, na verdade) variados. Estes são, acima de tudo,

actantes que interferem diretamente na construção do processo e, por conseguinte, no estabelecimento de “seres inesperados que são necessários para qualquer entidade existir” (LATOURE, 2013, p. 25).

Convém dizer que Foucault produz seu trabalho sobre autoria de forma mais incipiente que outros temas por ele estudados. Embora sua ideia inicial em *O que é um autor?* (1997), de acordo com Souza (2011), seja a correção de algumas ideias expressas em *As palavras e as coisas*, o autor também está expressando

[...] seu descontentamento com a teorização da escritura na década de 60, através da qual escritores lutavam apenas pela preservação de seus privilégios políticos, o que segundo o filósofo deu origem a obras literárias medíocres envoltas em epistemologia, lingüística e ciência, a fim de terem mais possibilidade de sobrevivência em uma cultura tecnocrática. (SOUZA, 2011, p. 129).

Assim sendo, Foucault, nessa obra, parte de uma perspectiva onde o sujeito é visto a partir das formas de representação históricas, com os textos admitindo autoria no momento em que passam a transgredir uma cultura já estabelecida e que, até esse momento, não aceitava contestações. A partir dessa época, quando feitas, acarretavam em punições para quem violasse o sistema estabelecido. Antiguidade para textos literários (narrativas, tragédias, contos etc.) e autoria identificável para textos científicos eram uma padronização praticamente imposta pela época (séculos XIV ao XVIII). Nas palavras de Chartier (2012b, p. 39): “enquanto permaneceu o regime de anonimato para os textos que chamaríamos literários, a presença do nome próprio autorizava, certificava como verdadeiros, os enunciados que chamaríamos científicos”.

Ainda que depois haja uma inversão sobre a necessidade de autoria, já que os discursos científicos eram vistos por si mesmos, sem necessidade de verificação de autoria, uma vez que sua credibilidade é vista ligada a “um conjunto sistemático que lhes confere garantias e não a referência ao indivíduo que os produziu” (FOUCAULT, 1997, p. 49), possibilita colocar a função-autor “não apenas ligada à

propriedade literária, não somente ligada ao mecanismo e aos dispositivos²⁹ de censura, mas também à certificação, à atribuição de verdade” (CHARTIER, 2012b, p. 40).

Foucault nos remete, primordialmente, à questão da responsabilidade, uma vez que, para ele, o escrito necessita de um autor para que haja punição – isto é, responsabilização – pelo seu dito, fazendo com que o autor seja

[...] aquilo que permite explicar tanto a presença de certos acontecimentos numa obra como as suas transformações, as suas deformações, as suas modificações diversas (e isto através da biografia do autor, da delimitação da sua perspectiva individual, da análise da sua origem social ou da sua posição de classe, da revelação do seu projeto fundamental). O autor é igualmente o princípio de uma certa unidade de escrita, pelo que todas as diferenças são reduzidas pelos princípios da evolução, da maturação ou da influência. O autor é ainda aquilo que permite ultrapassar as contradições que podem manifestar-se numa série de textos: deve haver – a um certo nível do seu pensamento e do seu desejo, da sua consciência ou do seu inconsciente – um ponto a partir do qual as contradições se resolvem, os elementos incompatíveis encaixam finalmente uns nos outros ou se organizam em torno de uma contradição fundamental ou originária. Em suma, o autor é uma espécie de foco de expressão, que, sob formas mais ou menos acabadas, se manifesta da mesma maneira, e com o mesmo valor, nas obras, nos rascunhos, nas cartas, nos fragmentos, etc. (FOUCAULT, 1997, p. 53-54).

Seguindo com Foucault, são apresentadas características do que o filósofo apresenta como a função-autor. A primeira característica está ligada à questão punitiva, uma vez que os escritos ultrapassam a divisão entre o sagrado e o profano. Ou seja, há atribuição de autoria a indivíduos, não mais a seres intocáveis ou imaculados ou sagrados, fazendo com que os indivíduos fossem responsabilizados e punidos pelos seus discursos contra os costumes impostos (FOUCAULT, 1997). A

²⁹ Para fins de clareza, compreendemos dispositivo no sentido que Foucault (2012, p. 364) o estabelece: “conjunto decididamente heterogêneo que engloba discursos, instituições, organizações arquitetônicas, decisões regulamentares, leis, medidas administrativas, enunciados científicos, proposições filosóficas, morais, filantrópicas. Em suma, o dito e o não dito são elementos do dispositivo. O dispositivo é a rede que se pode estabelecer entre esses elementos”. O autor segue sua explicação do que entende por dispositivo, alertando para as relações que os elementos podem estabelecer entre si, independentemente de sua natureza, agindo como *players* que podem fazer com que a rede e, conseqüentemente, seus movimentos sejam alterados funcionalmente e suas articulações ajam como respostas a uma inquirição julgada necessária para aquele momento.

segunda característica da função-autor está ligada à colocação do discurso no tempo e na cultura, ou seja, em relação ao seu contexto ou, ainda como relata Foucault (1997, p. 48), “a função autor não se exerce de forma universal e constante sobre todos os discursos”. A terceira é a complexidade que envolve essa função, uma vez que “ela não se forma espontaneamente como a atribuição de um discurso a um indivíduo. É antes o resultado de uma operação complexa que constrói um certo ser racional a que chamamos de autor” (FOUCAULT, 1997, p. 50). A última característica diz que o autor é uma pluralidade de “eus”, onde cada indivíduo pode apresentar posições e discursos diferenciados para cada momento.

O entendimento de Foucault é que existe uma pluralidade de posições que tomam o indivíduo como autor, isto é, o autor não é um ente solitário, mas, sim, um arranjo de posicionamentos, formas de pensar e criar, advindas de antecedentes culturais, bem como por indivíduos que colaboram ou que concorrem com ele (CHARTIER, 2012b).

A função-autor é assim resumida por Foucault (1997, p. 56-57):

Resumi-los-ei assim: a função autor está ligada ao sistema jurídico e institucional que encerra, determina, articula o universo dos discursos; não se exerce uniformemente e da mesma maneira sobre todos os discursos, em todas as épocas e em todas as formas de civilização; não se define pela atribuição espontânea de um discurso ao seu produtor, mas através de uma série de operações específicas e complexas; não reenvia pura e simplesmente para um indivíduo real, podendo dar lugar a vários 'eus' em, simultâneo, a várias posições-sujeitos que classes diferentes de indivíduos podem ocupar. (FOUCAULT, 1997, p. 56-57).

Dessa forma, para Foucault, a função-autor é algo relacionado à posição do sujeito, isto é, um aspecto do indivíduo que é atravessado por questões culturais, políticas, sociais, econômicas, temporais, entre outras, enfim, a construção e colocação do sujeito como ator social de um mundo complexo. Entre o autor e a pessoa do autor, a entidade autor só existe do corte entre as entidades, a partir dos múltiplos “eus”, ou seja, numerosos egos.

O que enxergamos é que tanto Foucault (1997) quanto Chartier (2012a, 2012b) estão preocupados com a função-autor pela “maneira como o texto aponta para essa figura que lhe é exterior e anterior, pelo menos aparentemente” (FOUCAULT, 1997, p. 34). Isso reflete no que viemos argumentando, ou seja, a análise que deve ser levada em conta, para além do autor, é uma argumentação de autoria, de ligações desse conceito/processo que se estabelece em forma de rede, podendo ser considerado um ator-rede.

O que é possível notar, corroborados por Souza (2011, p. 128), é a complexidade com que Foucault apresenta o tema, sendo esse “um caminho repleto de obstáculos e limitações reconhecidas pelo próprio filósofo”, pois assinar um texto, como uma carta ou um documento legal, por exemplo, não faz de um indivíduo um autor. A função-autor, como destaca Chartier (2012b, p. 28), “resulta [...] de operações específicas, complexas, que relacionam a unidade e a coerência de alguns discursos a um dado sujeito”.

Uma das dificuldades para compreender o conceito/processo, de acordo com Foucault (1997, p. 37), está ligada a definição de obra: “O que é uma obra? Em que consiste essa curiosa unidade que designamos por obra? Que elementos a compõem? Uma obra não é o que escreveu aquele que se designa por autor?”. É construída, a partir disso, uma dúvida que, segundo o filósofo, é de cunho teórico e técnico: “será que tudo o que ele [o autor] escreveu ou disse, tudo que ele deixou atrás de si, faz parte de sua obra? (FOUCAULT, 1997, p. 38).

Outras dificuldades incluem também a escrita e o nome do autor. Se Shakespeare não nasceu onde temos notícia, por exemplo, em nada afetaria seu nome de autor. Já caso fosse descoberto que Shakespeare não existisse ou que não escreveu *Os Sonetos*, o nome do autor já passa a ser diferente em relação ao seu funcionamento (FOUCAULT, 1997).

Essas averiguações ficam mais nebulosas quando necessitam explicitar e entender que “o fato de que aquele que é proprietário do objeto escrito não é mais o proprietário do texto e que o proprietário do texto é aquele que, eventualmente, se desfez da propriedade do objeto” (CHARTIER, 2012b, p. 50). Isso está diretamente ligado ao que tratamos na seção anterior, ratificando a dupla face da propriedade,

uma econômica, possibilitando transformar a obra em um negócio lucrativo, e outra moral, ligada à reputação, à honra do autor.

Assim sendo, advogamos que o nosso interesse não recai no autor, mas sim na autoria. Embora entendamos que o termo autor destaca os elementos que estão ao redor do conceito/entidade, parece-nos estar ligado à pessoa, fazendo com que a análise recaia sobre o indivíduo como forma humanizada, sendo alvo de um estudo quase psicológico. Dessa forma, nossa preferência está colocada no termo autoria, o qual compõe o processo como um todo e comporta, à luz da Teoria Autor-Rede, um modo de entender os objetos, conceituais ou pragmáticos, atuantes na construção do elemento³⁰.

3.3 Ética autoral: em busca da responsabilidade

A noção de ética aqui exposta está ligada ao seu acontecimento na ciência, principalmente do que chamamos de ciência moderna a partir do século XVII. A isso se soma a grande dependência do ser humano a diversos objetos tecnológicos, alterando a forma de atuação do homem com o mundo e da relação do mundo com o homem, bem como suas questões éticas e morais. Assim, podemos compreender, à luz de Droit (2012, p. 10), que as questões éticas “referem-se ao sentido da vida, aos limites das nossas intervenções sobre a matéria viva, ou ainda à dignidade humana”. Isto nos remete ao que está no cerne da ética, ou seja, a liberdade, as escolhas feitas pelos indivíduos.

A liberdade, no entanto, compreende uma questão restritiva: a preocupação com o outro, sendo os problemas éticos colocados na relação humana com outras entidades. “A ética é em primeiro lugar a problemática dos outros, a preocupação que temos com sua existência, com sua presença, com seus desejos, com suas esperanças, com sua dignidade e com sua liberdade” (DROIT, 2012, p. 38).

³⁰ Em alguns casos, especialmente no que se refere às citações, seguiremos utilizando o termo “autor” em vez de “autoria”, porque, como apontamos, a literatura, de forma geral, utiliza os dois termos de maneira indiscriminada e não nos sentimos no direito de interferir no discurso de um autor e nos termos utilizados em seus trabalhos.

Colocado de outra maneira, podemos dizer, à luz de Fourez (1995), que a ética está ligada àquilo que é desejável em um contexto histórico-social envolto em perspectivas de valores, ideologias etc. na relação do ser humano. Há sempre uma posição de comportamento, um modo de ver e colocar as coisas na discussão, uma apropriação que se baseia em uma ótica de vida e que nunca é neutra.

Em decorrência, se comportamento está por trás da moral e, por consequência, da ética, compreendê-la é entender “uma forma de conhecimento que diz respeito aos comportamentos” (DROIT, 2012, p. 15), sendo possível assimilar que existem comportamentos melhores e piores. Isso ocorre porque

na vida real, defrontamo-nos com problemas práticos [...] dos quais ninguém pode eximir-se. E, para resolvê-los, os indivíduos recorrem a normas, cumprem determinados atos, formulam juízos, e às vezes, se servem de determinados argumentos ou razões para justificar a decisão adotada ou os passos dados. (SANCHEZ VÁZQUEZ, 1987, p. 7).

É oportuno realizar uma diferenciação entre moral e ética, segundo Droit: “'moral' seria referente às normas herdadas, 'ética', às normas em construção. 'Moral' designaria principalmente os valores existentes e transmitidos; 'ética', o trabalho de elaboração ou de ajuste necessário em face das mudanças em curso” (DROIT, 2012, p. 19). Assim, como anota Sánchez Vásquez (1987), além de agir moralmente em atitudes cotidianas, como problemas relacionais e tomada de decisões, necessitamos de reflexão sobre esses comportamentos, isto é, na passagem da prática para o pensamento reflexivo sobre ela passamos à ética.

A ética é a teoria, a investigação ou a explicação de um tipo de experiência humana ou forma de comportamento dos homens – a experiência da moral –, considerando, porém, na sua totalidade, diversidade e variedade. O que nela se afirma sobre a natureza ou fundamento das normas morais deve valer para a moral da sociedade grega ou para a moral que vigora de fato numa sociedade moderna. É isso que assegura seu caráter teórico e evita sua redução a uma disciplina normativa ou pragmática. O valor da ética como teoria está naquilo que explica, e

não no fato de prescrever ou recomendar com vistas à ação em situações concretas (SÁNCHEZ VÁSQUEZ, 1987).

A partir disso seguimos nos apoiando em Sánchez Vásquez (1987, p. 12) para definirmos o que entendemos por ética: “[...] teoria ou ciência do comportamento moral dos homens em sociedade. Ou seja, é ciência de uma forma específica de comportamento humano”. Assim sendo, além de um caráter científico que visa abordar características universais de seu objeto de estudo, descreve conceitos, teorias, hipóteses etc. que expliquem o fato moral de forma sistemática, metódica e comprovável.

Aportando na contemporaneidade e percebendo os avanços da tecnologia que tendem a nos colocar em frente a questões nem sempre claras, as respostas podem não ser únicas, uma vez que a moral de uma comunidade pode não ser compartilhada por outra (DROIT, 2012). Assim, trabalhar a ética permite pensar os atos diretos e indiretos das entidades envolvidas no processo de autoria de dados científicos. Por este motivo, e para o contexto deste trabalho, a questão da ética está ligada à ciência e a forma como os dados científicos se relacionam com a autoria.

Regras morais se apresentam de forma variada, as quais são modificadas entre determinadas sociedades, épocas, crenças etc. Por outro lado, o ser humano quer chegar ao universal, às regras gerais, àquilo que é entendido e aceito de forma geral e, para isso, não pode depender da temporalidade ou de algum mandatário de um povo ou seita. Então não são apenas costumes que devem ser observados, mas, principalmente, “a preocupação com o outro. Porque a existência dos outros, as múltiplas relações entre eles e eu constituem o ponto de partida mais universal de todas as formas de ética” (DROIT, 2012, p. 37).

Isto nos leva a inferir que os problemas éticos existem apenas na relação com os outros. Assim, uma vez que “a ética é em primeiro lugar a problemática sobre os outros, a preocupação que temos com sua existência, com sua presença, com seus desejos, com suas esperanças, com sua dignidade e com sua liberdade” (DROIT, 2012, p. 38), significa ter consciência sobre a existência e plenitude dos outros. Em outras palavras, ser responsável com as atitudes e como elas afetam as entidades que se envolvem em determinada ação.

Não se deve, no entanto, utilizar a ética como parâmetro de comportamento, uma vez que trata da consciência de saber a razão e as consequências de uma decisão baseada em determinados valores, isto é, o estabelecimento de critérios para a escolha, para a decisão. Dessa forma, é deveras importante conhecer o ambiente e as possibilidades de critérios a serem analisados com vistas a melhor decisão para a ocasião, deixando de lado o despontar de uma interpretação ética/moral da ação e da racionalidade humana, atrelando critérios e estabelecendo o que é bom ou que é mau, o que é correto ou o que é errado, aprovando ou desaprovando atitudes.

A ideia é compreender, como propõe Silva Filho (2009, p. 9), a razão dos acontecimentos, ou seja, “conceber por razão aquilo que, num quadro discursivo, explica uma ação”. Isto é, exercitarmos a capacidade descritiva da vida humana – em nosso caso a questão autoral – e suas escolhas/atitudes em determinado contexto.

Silva Filho (2005, p. 88) busca clarificar esta ideia:

uma ação moral, nesta perspectiva, não apenas se é uma 'boa ação', mas se é expressão de uma decisão livre e racional: se numa alternativa posso julgar que a ação A é preferível, é mais justa do que a ação B, logo ajo racionalmente se faço A, pois, na ordem das razões, eu não poderia fazer outra coisa senão A. Se o que eu faço pode ser reconhecido como algo nobre ou reprovável, isto é outro problema: isto está associado a 'seguir ou não uma lei'. Se A era preferível a B, se eu tinha razões para agir como em A, mesmo que isto infrinja uma lei, A permanece sendo uma ação moral e, como tal, posso ser inclusive punido por tal. (SILVA FILHO, 2005, p. 88).

Podemos utilizar como exemplo o relato de Lipinski e Britz (2000) ao tratarem sobre as implicações éticas da propriedade da informação, uma vez que a responsabilidade sobre a informação e/ou dados pode ter impacto no controle de acesso a eles, podendo restringir a criatividade e a disseminação de conhecimento. “A posse de informações refere-se a qualquer matéria tangível (propriedade) sobre o

qual se estabeleceu direitos, a saber: a usá-lo, para se divertir, para alienar e de excluir os outros a partir dele”³¹ (LIPINSKI; BRITZ, 2000, p. 51).

Tratando algumas reflexões éticas sobre a propriedade da informação, Lipinski e Britz (2000, p. 62) alertam para o fato de que “os sistemas econômicos não têm um imperativo moral em sua estrutura”³², possibilitando que injustiças, inclusas as sociais, possam ocorrer nesse processo. A busca por essa justiça também pode ser vista, em uma perspectiva ética, a partir do direito de acesso à informação ou aos dados que podem proporcionar um bem comum.

Isso impacta, obrigatoriamente, na responsabilidade do pesquisador/autor perante a autoria dos dados científicos. Neste sentido, a pergunta mais pertinente seria: qual a responsabilidade do pesquisador? Droit (2012, p. 61) mostra que quando estamos refletindo questões éticas, “há de fato a ideia de que nós criamos nossas próprias ações, de que decidimos fazer isto ou aquilo, e, portanto, de que somos responsáveis. Em toda forma ética há uma concepção de responsabilidade”.

Nisso é relevante perceber que o impacto das tecnologias da informação e da comunicação tem efeito além dos profissionais que lidam com esses aparatos, mas influenciam a sociedade como um todo, seja na relação homem-homem, seja na relação homem-máquinas/instrumentos, seja – por que não? – na relação máquinas/instrumentos-máquinas/instrumentos. “A ciência e a tecnologia tiveram uma parte bem significativa na organização da sociedade contemporânea, a ponto de esta não poder prescindir das primeiras: energia, meios de transporte, comunicações, eletrodomésticos etc.” (FOUREZ, 1995, p. 207).

Se considerarmos que os sistemas eletrônicos mantêm interação com humanos e não humanos – como admite a Teoria Ator-Rede –, estamos diante de novos problemas morais e, por consequência, éticos que estão refletidos em diversos aspectos nos dados científicos primários, inclusa a questão autoral. Com isso, devemos sublinhar a existência de normas universais sem relegar, no entanto, os aspectos históricos e sociais de uma entidade ou de um ambiente social, exigindo

³¹ No original: “Ownership of information refers to any tangible matter (property) over which one has stated rights, namely: to use it, to enjoy it, to alienate it and to exclude others from it”.

³² No original: “economic systems do not have a moral imperative built within its structure”.

pensar a ética de forma a levar em consideração as macro e as micro situações envolvida neste processo, proporcionando uma análise descritiva das ações.

Assim, parece-nos crucial no que concerne à responsabilidade autoral de dados científicos, compreender, de forma geral, sobre a relação de responsabilidade da autoria sobre os dados de investigação. Em sentido mais específico com o objeto da coleta de dados primários, é o entendimento da questão ética da autoria perante os objetos, no caso, dados científicos, que manipula.

4 VESTÍGIOS DE UMA REDE EM CONSTRUÇÃO: AS TEIAS DA TEORIA ATOR-REDE

As ciências e as técnicas não são notáveis por serem verdadeiras ou eficazes – estas propriedades lhes são fornecidas por acréscimo e por razões outras que não a dos epistemólogos –, mas sim porque multiplicam os não-humanos envolvidos na construção dos coletivos e porque tornam mais íntima a comunidade que formamos com estes seres [...] Os saberes e os poderes modernos não são diferentes porque escapam à tirania do social, mas porque acrescentam muito mais híbridos a fim de recompor o laço social e de aumentar ainda mais sua escala. (LATOUR, 2009, p. 106).

Este capítulo diz respeito à introdução na Teoria Ator-Rede³³ (ANT), compreendendo uma tentativa de seguir uma linha de raciocínio em busca dos principais conceitos da Teoria. É interessante notar que à primeira vista as ideias dessa teoria parecem de difícil apreensão. E não deixam de ser. São noções subversivas que pretendem explicar de forma plana o que acontece em uma rede de atores desde sua formação até, caso seja necessário, seu desaparecimento.

Não é de se esperar da Teoria um costumeiro apaziguamento às percepções das redes de atores. Isto ocorre pelo fato dela se contrapor ao que chama de “sociologia crítica” e “sociologia social”. Os estudos ANT buscam desconstruir a percepção de mundo tradicionalmente enraizada em valores políticos e epistemológicos. Isto é, os estudos propostos pela ANT tentam, segundo Latour, reconstruir o que foi perdido na Sociologia ao longo do tempo, ou seja, o poder de realizar e perceber associações.

A Teoria Ator-Rede se opõe aos estudos ditos sociais, os quais buscam a existência de um contexto social que influencia, que atua e que estabelece características explicativas por sua categorização. Esse contexto social é o que muitos pesquisadores utilizam como “uma esfera específica da realidade; pode ser

³³ Embora Latour admita não gostar do nome Teoria Ator-Rede, uma vez que ele é “tão deselegante, tão confuso, tão sem sentido”, ele assume que o acrônimo ANT (Actor-Network Theory) deve ser conservado e evoca a imagem de uma formiga (*ant*, em inglês) para descrever o pesquisador que utiliza a Teoria, alguém míope, farejador de trilhas, viciado em trabalho.

usado como um tipo especial de causalidade para explicar os aspectos residuais que escapam a outros domínios” (LATOURE, 2012, p. 20-21).

Isso significa dizer que as explicações aceitas pela ANT não dizem respeito a existência de “mãos invisíveis” (contextuais) que seriam capazes de controlar (categorizar) e/ou explicar determinada situação. De modo a buscar clareza, pode-se exemplificar com a questão da política (poder). A ideia é que não se use o conceito de política para explicar uma determinada rede de ações coercitivas. Se a causa de determinada ação for uma relação política, então a política das ações executadas e que formam a rede deve ser descrita para então sejam explicadas³⁴.

Latour (2012, p. 51-52) argumenta que “não há grupo relevante ao qual possa ser atribuído o poder de compor agregados sociais, e não há componente estabelecido a ser utilizado como ponto de partida incontroverso”, isto é, não há um processo de relacionamento preconcebido, uma vez que os laços são frágeis e podem apresentar mudanças sem prévio aviso. Isto significa que a não adoção de uma postura delimitadora faz com que os analistas devam seguir seus atores, sem distinção de natureza, tanto na formação quanto no desmanchar de grupos.

O dever [...] não é estabilizar [...] a lista de agrupamentos que compõem o social. [...] O] ponto de partida tem de ser justamente as controvérsias acerca do agrupamento a que alguém pertence, incluindo, é claro, as dos cientistas sociais em torno da composição de mundo[, uma vez que] controvérsias proporcionam ao analista os recursos necessários para rastrear conexões sociais. (LATOURE, 2012, p. 51-52).

Freire (2006, p. 46) parece ir direto ao ponto central da questão tratada pela ANT. Para a autora, a ANT corresponde ao que podemos entender como um método de olhar plano, compreendendo “um esforço de sistematização de princípios e regras metodológicas subjacentes a uma forma de pensar e tratar a realidade que, ao invés de interpretar o mundo a partir de 'grandes divisões', visa descrevê-lo levando em conta sua hibridização”.

³⁴ Deve-se, no entanto, ter cuidado, em relação a análise, no que se refere à política e à sociedade. “Embora exista efetivamente, em cada interação, uma linha pontilhada que leva a alguma entidade virtual, total e sempre preexistente, essa é exatamente a trilha que não deve ser seguida [...]; virtual e escura ela é, e permanecerá virtual e escura” (LATOURE, 2012, p. 241).

Todo o processo de análise se constitui de atores que não são substituídos por entidades que explicariam determinada questão pela existência contextual ou porque a sociedade é dita de certa forma pela imposição de situações invisíveis. A exploração da reunião de coletivos nunca é pré-determinada; como a sociedade, é construída continuamente. Uma vez que não há como atribuir determinada situação ao contexto, “se você mencionar uma ação, terá de apresentar um relato sobre ela” (LATOUR, 2012, p. 84). Como já dito, isto implica que “a presença do social tem de ser repetidamente demonstrada e não simplesmente postulada” (LATOUR, 2012, p. 85).

[...] a incômoda resposta que se obtém em relação aos famosos “contextos” é que existe alguma coisa que possibilita a interação, ao trazer para a cena a maioria dos seus ingredientes necessários, mas que essa “alguma coisa” ao mesmo tempo está presente por trás e é demasiado abstrata para fazer o que quer que seja. A estrutura é muito poderosa, e no entanto demasiado fraca e remota para ter qualquer eficácia. [...] as próprias estruturas permanecem demasiado abstratas porque não foram convocadas, mobilizadas, realizadas ou encaradas em algum tipo de interação local e vivida. (LATOUR, 2012, p. 243-245).

A ANT inicia sua trajetória reivindicando que o que se tem chamado de social pode ser revisto de modo a compreendê-lo como “uma série de associações entre elementos heterogêneos” (LATOUR, 2012, p. 23), isto é, redefine a Sociologia a partir da busca por associações, “não designa[ndo] uma coisa entre outras [...], e sim um tipo de conexão entre coisas que não são, em si mesmas, sociais” (LATOUR, 2012, p. 23). Isso se dá porque o que se pretende estabelecer é que “todos os elementos heterogêneos precisam ser reunidos de novo em uma dada circunstância” (LATOUR, 2012, p. 23), uma vez que não se pode entender as associações como sendo regidas pelo social, mas o social como resultado de associações.

Como se pode perceber há uma redefinição do que conceitualmente se entende como “social”, um colapso que é orientado pela busca por associações mas que pode se tornar ainda mais amplo em virtude da intensidade e diversidade de atores envolvidos. Assim Latour (2012, p. 25) recharacteriza o que se compreende

por social: “não como um domínio especial, uma esfera exclusiva ou um objeto particular, mas apenas como um movimento peculiar de reassociação e reagregação”.

A estrutura não é um corpo acima de nossas cabeças que dispõe de categorias as quais nos encaixamos. A estrutura é fornecedora de elementos e é alimentada ao mesmo tempo, por isso a forma de rede é preciosa. Colocando de forma mais clara: “a ANT é antes de tudo um princípio de projeção abstrato para desenhar qualquer forma, não uma decisão arbitrária sobre qual forma deva estar no mapa” (LATOURE, 2012, p. 257). Estudar como certo ente é mais poderoso ou maior, não deve ser o foco da ANT. A perspectiva é que se estabeleça como tamanho, poder e organização são gerados (LAW, 1992).

Acompanhando isto, Latour (2012) chega a afirmar que nem a sociedade nem o social existem na forma de uma estrutura. A forte afirmação está baseada nas mudanças das ligações entre os atores, ou seja, a sociedade, ou o social, é uma produção constantemente construída, sendo a ANT uma ferramenta que permite analisar esse processo. “A sociedade é a consequência, não a causa das associações. [...] não é o social que explica as associações, mas as associações que explicam o social” (LATOURE, 2012, p. 340).

[...] o social não é um lugar, uma coisa, um domínio ou um tipo de matéria, e sim um movimento provisório de associações novas. [...] A sociedade não é o todo 'onde' todas as coisas estão inseridas, mas aquilo que 'atravessa' tudo, calibrando conexões e dando a cada entidade que encontra uma chance de comensurabilidade. (LATOURE, 2012, p. 341-344).

Isso faz perceber a dificuldade pela qual passa a ANT: por que é tão difícil rastrear o social? Esta é a pergunta feita na introdução à segunda parte do livro *Reagregando o social*, e é, além de uma tentativa de estabelecer clareza ao longo do trecho, uma descrição da tarefa de buscar associações em sua real dificuldade.

Essa dificuldade pode ser entendida a partir do momento em que social é algo que pode designar fenômenos diferenciados: de um lado como uma substância (matéria), de outro lado como movimento. Não há social nestes casos: no primeiro,

um sólido não possui capacidade de associação. No segundo caso é um fluido que não mostra a cola que une o coletivo. Assim, “não há nada mais difícil de apreender do que os vínculos sociais. Ele só é rastreável quando está sendo modificado” (LATOUR, 2012, p. 229). Neste sentido, a quantidade de vínculos que um ator possui está diretamente ligada à sua própria existência.

Esta ideia vem traçando críticas sempre contundentes à sociologia tradicional, a qual compreende a tentativa de unir a questão coletiva e a questão política na forma de uma sociedade. Esta é, para o autor, a responsável por não autorizar a existência do coletivo pleno formado por atividades políticas. Não pode, de forma alguma, extrapolar sua composição de um coletivo, criando um padrão universal que explica associações de forma invisível ou mágica (LATOUR, 2012).

Latour (2012) toma cuidado ao esclarecer que a ANT não atua apenas como um conjunto de entidades já reunidas, como uma sociedade, mas age como um ajuntamento de novas entidades que ainda não estão reunidas, podendo ser chamadas de coletivo. Um coletivo repleto de conexões entre humanos e objetos, humanos e humanos e objetos e objetos, e que não estabelece relação, *a priori*, entre o mundo material e o mundo social.

4.1 Rede de atores e a Teoria Ator-Rede

O que se entende por rede sofreu diversas adaptações ao longo do tempo, principalmente entre os séculos XVIII e XIX, quando o conceito passou a ser compreendido na transição de um estado natural para um estado artificial, ou seja, o conceito de rede começou a ser construído, uma vez que passou a ser pensado a partir de sua relação com o espaço (MUSSO, 2010). Sob o olhar da relação com a sociedade, podemos entender que as redes são como fios que se ligam uns aos outros, sendo apenas possível sua análise a partir de quando se percebe o modo como esses fios se ligam de forma recíproca, sem descuidar das tensões e/ou estrutura interna da própria rede que altera a forma como a individualidade ocorre (ELIAS, 1994).

A rede não é algo rígido, sendo que a ordem a ser estabelecida nasce a partir de diversas unidades, não podendo ser vista de maneira individualizada, uma vez que está “em constante movimento, como um tecer e destecer ininterrupto das ligações” (ELIAS, 1994, p. 35), trazendo uma noção de possível flexibilização do que se pretende analisar, como anota Latour (2009, p. 9): “mais flexível que a noção de sistema, mais histórica que a de estrutura, mais empírica que a de complexidade, a rede é o fio de Ariadne dessas histórias confusas”.

Seguindo neste viés de percurso, compreendemos, à luz de Latour, que a ideia de rede proposta pela ANT é uma nova forma de encarar a problemática da “produção social do conhecimento científico”, porque ela “se conecta ao mesmo tempo à natureza das coisas e ao contexto social, sem contudo reduzir-se nem a uma coisa nem a outra” (LATOURE, 2009, p. 11).

Seguindo por esta linha de pensamento, uma rede é, como aponta Moraes (2004), uma lógica de conexões e não de superfícies, definidas por seus agenciamentos internos e não por seus limitadores externos, não existindo pontos que mereçam ser privilegiados, caracterizando uma totalidade aberta e capaz de se relacionar e crescer em todas as direções.

Em uma das mais claras definições apresentadas por Bruno Latour, o autor diz que uma rede

é uma expressão para avaliar quanta energia, movimento e especificidade nossos relatos conseguem incluir. Rede é conceito, não coisa. É uma ferramenta que nos ajuda a descrever algo, não algo que esteja sendo descrito. [...] a rede não é aquilo que está representado no texto, mas aquilo que prepara o texto para substituir os atores como mediadores. (LATOURE, 2012, p. 192).

Importante notar é que a ideia de rede trazida pela ANT permite uma combinação para explorar os fenômenos que incluem não somente atores humanos, constituindo-se de uma rede heterogênea de atores capazes de agir e fazer agir, sendo estes também atravessados por ações de outros atores.

A ANT, abrindo mão de uma suposta linha divisória entre humanos e não humanos, percebe as entidades como uma grande rede conectada e que se modifica constantemente a partir das relações que se estabelecem em diversos níveis. Como ressaltam Moraes (2004) e Freire (2006), não há previamente uma separação entre o mundo das pessoas e o mundo das coisas. Estes elementos se reúnem e se engendram de forma a não levar em conta princípios essencialistas. Assim sendo, compreender essa configuração permite ir adiante, apoiados em uma perspectiva que não aceita categorizações prévias.

O social que constitui a sociedade representa apenas uma parte dos grupos formadores do coletivo. Para reagrupar o social, será necessário, afora a circulação e a formatação de laços sociais tradicionalmente concebidos, descobrir outras entidades que circulem. (LATOURE, 2012, p. 333).

A conjuntura básica é que devemos permitir e observar a existência de uma simetria entre humanos e não humanos. Isto “significa não impor a priori uma assimetria espúria entre ação humana intencional e mundo material de relações causais” (LATOURE, 2012, p. 114). Com isto, superamos essa divisão dualista em geral imposta pela categorização, privilégio ou até mesmo “esquecimento” de elementos antes da entrada em um domínio, possibilitando que se analise os poderosos da mesma forma que quaisquer outras entidades (LAW, 1992). Nesse aspecto, não há uma separação hierárquica entre os atores.

Se é um grande equívoco metodológico restringir de antemão e no lugar do ator o alcance das entidades que devem povoar o mundo social, seria tanto ou mais patético não reconhecer o trabalho incansável que fazem para limitar o repertório dos actantes e manter as controvérsias a distância. (LATOURE, 2012, p. 325).

Enxergar os objetos como parte de uma rede é tratar sua existência como essencial para a constituição da mesma. Sua incorporação ocupa lugar nas discussões e relatos de pesquisas, não atuando mais como apenas uma peça que tem seu lugar relegado, mas como peça-chave em muitas concepções do que

acontece em um ambiente. Realizar uma ação implica a utilização de outra(s) entidade(s), ou seja, não é possível agir só. A ação é acompanhada por objetos de fundos variados que formam uma rede da ação executada, isto é, forma-se um coletivo.

Esse coletivo é formado pela interação dos seres humanos com outros atores, principalmente não humanos, com outros tipos de materiais. A composição do coletivo, como aponta Law (1992), é uma questão de preferências não somente entre os humanos, mas também entre os não humanos.

O que se deve apreender é a complexidade que envolve o processo. Neste passo não é mais importante privilegiar um humano a um não-humano ou vice-versa. A investigação deve levar em conta a proliferação de actantes que sejam relevantes ao caso. Como mostra Lemos (2010, p. 16),

[...] atores humanos e não humanos agem em rede e constituem-se eles mesmos como rede. Não há claramente sujeito e objeto, nem causa ou efeito predeterminados. Cada nó de uma rede de ação, por exemplo, dirigir um carro, ver TV, usar um celular ou um laptop, convoca uma série de atores humanos e não humanos. As relações sociais, morais ou éticas que daí emergem não podem ser avaliadas a priori ou a partir de um elemento central causal. [...] Considerar as ações na vida social é considerar os diversos fatores como agentes e as ações como parte de uma rede de relações. (LEMOS, 2010, p. 16).

A análise, então, transforma-se em um processo de incertezas, principalmente no que se refere ao conjunto de entidades do mundo. O não delimitar antecipado de entidades, sejam elas de qualquer natureza, faz com que um movimento em torno de fatores complexos seja a florado, um desdobramento que torna a coisa múltipla, seja em suas associações, seja em si própria a partir de observações que permitam esboçar uma rede que, ao ser formada por entidades complexas, adquire também o status de complexa.

Na rede, cada elemento é simultaneamente um ator, cuja atividade consiste em fazer alianças com novos elementos, e uma rede, capaz de redefinir e transformar seus componentes. Os agentes sociais,

portanto, não estão “contidos” unicamente em corpos: um ator é uma rede moldada por relações heterogêneas, ou seja, ele é um efeito de rede que, por sua vez, participa e molda outras redes. Na medida em que a rede se caracteriza por estabelecer um campo de tensões heterogêneas, a síntese não é um resultado necessário. (PEDRO, 2008, p. 4).

O que a ANT propõe é, portanto, uma ruptura radical, como aponta Law (1992), até mesmo com questões éticas, epistemológicas e ontológicas. Sua base parte da negação de que os humanos têm algum tipo de privilégio apenas por serem pessoas, seres especiais. Deve-se utilizar isso como um argumento clarificador da expectativa política da Teoria, uma vez que não diferenciar pessoas de objetos é uma posição analítica e não ética.

Dizer que não há diferença fundamental entre pessoas e objetos é uma atitude analítica, e não uma posição ética. E dizer isso não significa que tenhamos de tratar as pessoas como máquinas. Não temos que lhes negar os direitos, deveres e responsabilidades que usualmente lhes atribuímos. Na verdade, podemos usar essa atitude para aprofundar questões éticas sobre o caráter especial do efeito humano, como, por exemplo, em casos difíceis tais como os de vida mantida artificialmente por conta das tecnologias de tratamento intensivo³⁵. (LAW, 1992, p. 383).

A compreensão dos não humanos – máquinas, textos, dinheiro, conceitos etc. – em conjunto com os humanos, cria uma coletividade. Suas ausências, ou a retirada deliberada desses entes de análises, faz com que a essência da sociedade, em sua acepção coletiva, perca sentido. Como retirar de análise algum tipo de engenhoca que permite o funcionamento de um sistema mais complexo? Como retirar a mesa ou o computador da análise sobre a escritura de um artigo?

Por isso as redes são heterogêneas. Por não se utilizar aqui a concepção de sociedade como uma macroestrutura categorizante, essa heterogeneidade produz o

³⁵ No original: "To say that there is no fundamental difference between people and objects is an analytical stance, not an ethical position. And to say this does not mean that we have to treat the people in our lives as machines. We don't have to deny them the rights, duties, or responsibilities that we usually accord to people. Indeed, we might use it to sharpen ethical questions about the special character of the human effect--as, for instance, in difficult cases such as life maintained by virtue of the technologies of intensive care".

que compreendemos por entidades, e ao mesmo tempo redes sociotécnicas, onde o ator é fundamentalmente conhecido como actante, percebendo que “ao falarmos de ator, deveremos sempre acrescentar a vasta rede de vínculos que o levam a atuar” (LATOURE, 2012, p. 313). Isto é, quando chamamos a baila uma rede sociotécnica, na realidade o que invocamos é uma dinâmica associativa entre actantes, sejam eles humanos ou não.

De forma mais clara, podemos definir actante de acordo com Latour (2001, p. 346):

O segredo é definir o ator com base naquilo que ele faz – seus desempenhos – no quadro dos testes de laboratório. Mais tarde, sua competência é deduzida e integrada a uma instituição. Uma vez que, em inglês, a palavra actor (ator) se limita a humanos, utilizamos muitas vezes ‘actant’ (atuante), termo tomado a semiótica, para incluir não-humanos na definição. (LATOURE, 2001, p. 346).

Os actantes são, desta forma, “agentes, humanos ou não, mobilizados para a composição de redes temporárias engajadas em programas de ação” (HOLANDA; LEMOS, 2013, p. 2). Isto significa dizer que um agente “habita um conjunto de elementos (incluindo obviamente o corpo) que se estende por uma rede de materiais, somáticos e de outros tipos, que circundam cada corpo³⁶” (LAW, 1992, p. 384). Neste sentido, segue Law (1992), o argumento da ANT passa pela extrapolação do ser humano como entidade que carrega diversas atividades – por exemplo pensar, escrever, amar etc. –, fazendo com que outros atores também existam fora do corpo. Ator-Rede: quando um ator também é uma rede.

Um ator, segundo Latour, pode ser definido através do efeito causado pelas suas ações. “Somente podem ser considerados atores aqueles elementos que produzem efeito na rede, que modificam e são modificados por ela e são estes elementos que devem fazer parte de sua descrição” (FREIRE, 2006, p. 55).

Neste âmbito é reafirmada a importância de olhar todas entidades, não caracterizando objetos apenas como ferramental de um ator humano. Os objetos

³⁶ No original: “inhabits a set of elements (including, of course, a body) that stretches out into the network of materials, somatic and otherwise, that surrounds each body”.

têm colocação no espaço e no tempo e são cruciais para a demanda de movimentos de uma rede.

Os objetos técnicos que permeiam nosso cotidiano não são aqui concebidos como meros instrumentos a serviço da sociedade ou como suporte de algo que lhes é externo. São agentes/actantes capazes de engendrar transformações que ultrapassam o âmbito técnico-instrumental, participando da configuração de processos que não mais podemos definir como estritamente sócio-culturais – passamos a nos referir a eles, doravante, como sócio-técnicos. (PEDRO, 2008, p. 3).

Quem são os atores, afinal? Não há teoria que cubra isto, mas apenas que o número de atores de uma rede deve ser aumentado, uma vez que os objetos também agem, pois a vinculação social “é antes o nome de um movimento, um deslocamento, uma transformação, uma translação, um registro. É uma associação entre entidades de modo algum reconhecíveis como sociais no sentido corriqueiro” (LATOUR, 2012, p. 99). Não há forma, apenas movimentos. Movimentos que designam formas novas e diferenciadas. Sendo assim, a multiplicidade de atores é composta, também, pelas “entidades que foram explicitamente banidas da existência coletiva por mais de um século de explicações sociais” (LATOUR, 2012, p. 105).

Isto deve ser explicado a partir de relatos que permitam a clareza necessária e não justificada por um contexto, ou seja, não havendo substituição de uma entidade por outra, de um fenômeno por outro. Devemos investigar as entidades e possibilitar que estas sejam explicadas não como “um feito cognitivo misterioso, mas um empreendimento de construção de mundo muito prático que consiste em ligar entidades a outras entidades, ou seja, em traçar uma rede” (LATOUR, 2012, p. 152).

4.2 Tradução, mediação, caixa-preta

Todo esse processo é conduzido pelo que Latour (2012, p. 158) chama de tradução ou de translação. Mesmo não sabendo que posição cada ator possui no

plano simétrico, é sabido que eles “podem estar associados de tal modo que eles fazem outros fazerem coisas”, entendendo que são as transformações dos eventos que explicam o social. Neste sentido, “[...] a concatenação dos mediadores não traça as mesmas ligações e não requer o mesmo tipo de explicações como um séquito de intermediários transportando uma causa” (LATOURE, 2012, p. 158).

O conceito de tradução é de suma importância para a ANT, tal que a Teoria pode ser chamada de Teoria da Tradução. Seu mérito está na “interpretação dada pelos construtores de fatos aos seus interesses e aos das pessoas que eles alistam” (LATOURE, 2000, p. 178). Isto é, refere-se a “hibridação, mestiçagem, multiplicidade de conexões mais do que à repetição de elementos-chave. A tradução é sustentada por uma ontologia definida por sua hibridação” (MORAES, 2004, p. 326). Essa mistura subjetiva aportada pela tradução é que alimenta a origem do conhecimento, indo além da simples interação linear entre humanos.

Traduzir (ou transladar) significa deslocar objetivos, interesses, dispositivos, seres humanos. Implica desvio de rota, invenção de um elo que antes não existia e que de alguma maneira modifica os elementos imbricados. As cadeias de tradução referem-se ao trabalho pelo qual os atores modificam, deslocam e transladam os seus vários e contraditórios interesses. (FREIRE, 2006, p. 51).

Ou, como explica Latour:

[O social, a associação] não é nem um ator entre muitos nem uma força por trás de todos os atores transportados por meio de um deles, mas uma conexão que transporta, por assim dizer, transformações, usamos a palavra tradução. [...] Assim, a palavra tradução assume [...] uma relação que não transporta causalidade, mas induz dois mediadores à coexistência. [...] não existe sociedade, não existe domínio social nem existem vínculos sociais, mas existem traduções entre mediadores que podem gerar associações rastreáveis. (LATOURE, 2012, p. 159-160).

Em outras palavras, “tradução não significa apenas a mudança de um vocabulário para outro, mas acima de tudo um deslocamento, um desvio de rota,

uma mediação ou invenção de uma relação antes inexistente e que de algum modo modifica os atores nela envolvidos” (MORAES, 2004, p. 326). São negociações que atores impõem sobre outros atores. Assim sendo,

traduzir significa atribuir a um elemento de uma rede um papel a ser representado por ele; significa emprestar-lhe uma identidade, prática que é realizada por todos os elementos de uma rede, em um movimento mútuo e contínuo, a partir dos desejos, expectativas e/ou interesses de cada um dos tradutores. (BENAKOUCHE, 2001, p. 54).

Callon (1986) entende que existem “quatro momentos da translação”. De forma sintética, os momentos são assim elencados e explicados:

- (a) A problematização ou como se tornar indispensável: percorre a identificação de atores que estão envolvidos no processo investigado, construindo e/ou identificando associações que formam uma rede de actantes, isto é, um processo passageiro e que torna o processo dependente das forças de arregimentação existentes no momento, podendo ser alteradas em movimentos de associação e desassociação;
- (b) Os dispositivos de *interessement* ou como os aliados estão encaixados: consequência do primeiro momento, os dispositivos de *interessement* são “o conjunto de ações pelas quais uma entidade [...] tenta impor e estabilizar a identidade dos outros atores que define através de sua problematização. Diferentes dispositivos são utilizados para implementar essas ações”³⁷;
- (c) Como definir e coordenar as funções: recrutamento: este momento pode ser descrito como a estabilização de alianças, um momento de definições de papéis, notando que estes não são, como estabelece a ANT, preconizados de antemão, mas que ocorrem pela imposição descrita pelos dispositivos de *interessement*;
- (d) A mobilização de aliados: verificação da representatividade do mediador, isto é, a consolidação das etapas anteriores, analisando deslocamentos e

³⁷ No original: “Interessement is the group of actions by which an entity [...] attempts to impose and stabilize the identity of the other actors it defines through its problematization. Different devices are used to implement these actions”.

modificações nos actantes e, por conseguinte, nas associações por eles estabelecidas.

A tradução diz respeito a modificações a que os actantes estão propensos durante a vigência do programa, sendo alvos de deslocamentos e negociações, responsáveis pela moldagem entre as associações que formam o social, colocando-os como elementos representativos em uma rede. Isto permite que um discurso emerja e permita clarificar os complexos processos que em que o social está envolto (PEREIRA, 2013).

É possível introduzir agora dois conceitos de essencial importância à ANT: intermediários e mediadores (LATOURE, 2012, p. 65). O primeiro diz respeito à “aquilo que transporta significado ou força sem transformá-los; definir o que entra já define o que sai”. O segundo conceito está posto contrariamente ao primeiro porque o que entra não necessariamente é o que sai. Assim, “os mediadores transformam, traduzem, distorcem e modificam o significado ou os elementos que supostamente veiculam”. Esta diferença conceitual traz um carácter complexo aos atores, fazendo com que os mediadores produzam caminhos que são, como foi dito anteriormente, uma produção contínua de sociedade.

Uma ação, neste sentido, atua como uma ligação a partir dos diversos atores que desempenham seus papéis, sendo “assumida por outros” (LATOURE, 2012, p. 73). Isto serve como elemento diferenciado, já que a ação revela um carácter mediador, transformador, sendo o ator não apenas o emanante de ações, mas, sim, alvo delas, agindo e fazendo agir.

Isso leva a relatar incessantemente como uma ação se desenvolve, uma vez que os mediadores não têm os efeitos de um ato ligado obrigatoriamente a uma causa, mas sim à determinada circunstância que acarreta neste ato. A mediação atua como critério de operacionalização da ação de um actante, isto é, o actante pode ser entendido como um mediador a partir de suas ações.

Mediação é o diálogo ou a ação entre os diversos atores em que não há causalidade facilmente identificável. Ela se dá de acordo com os modos, ou seja, é uma ação a partir da maneira pela qual se dá o

processamento, a troca, o consumo e a produção infocomunicacional local entre os atores. (LEMOS, 2010, p. 11).

Por esse olhar, a mediação é “um elo que coloca os sujeitos em relação, humanos e não humanos. São eles que agenciam e produzem transformações nas redes sociotécnicas” (LEMOS, 2010, p. 17), onde “mediações geram redefinições da realidade, tecendo novas geografias e novos controles. Novas coerções exercem pressão sobre velhas coerções e as deslocam, descentram” (NOBRE; PEDRO, 2010, p. 48).

O interesse que a ANT tem em mediadores decorre da capacidade destes induzirem outros a fazer algo. Neste sentido, “induzir’ não é o mesmo que ‘causar’ ou ‘fazer’: há em seu âmago uma duplicação, um deslocamento, uma translação que modifica simultaneamente todo o argumento” (LATOURE, 2012, p. 312).

Isso possibilita notar que a rede não diz respeito a pontos sendo ligados e que entidades com maior poder são o centro da rede e a gerenciam. Mas há uma difusão do poder existente na rede, podendo estar distribuído e/ou explícito em alguns nós, fazendo com que a ordem não seja fixa, sendo visualizados movimentos que estipulam novos espaços (NOBRE; PEDRO, 2010, p. 52).

É permitida ao pesquisador, deste modo, a compreensão dos fatos científicos a partir da mobilização de agentes diversos, os quais seriam relegados em uma análise final de um produto científico. Cabe então a análise antes das caixas-pretas³⁸ serem fechadas, isto é, o estudo não diz respeito a um fato como algo dado e assumido como tal (caixa-preta), mas sim a todos os acontecimentos que estão referidos além do que entra (*input*) e do que sai (*output*) desta. É de suma relevância estudar a complexidade interna do ator-rede, por mais complicado que isso seja.

Mesmo transiente e precário, o ‘ator-rede’ ganha durabilidade e estabilidade com o fechamento em caixas-pretas. Nelas, os diversos

³⁸ “A expressão caixa-preta é usada em cibernética sempre que uma máquina ou conjunto de comandos se revela complexo demais. Em seu lugar, é desenhada uma caixinha preta, a respeito da qual não é preciso saber nada, senão o que nela entra e o que dela sai. [...] por mais complexo que seja seu funcionamento interno, por maior que seja a rede comercial ou acadêmica para sua implementação, a única coisa que conta é o que se põe nela e o que dela se tira”. (LATOURE, 2000, p. 14).

actantes permanecem engajados em uma única ação, podendo ser então tratados como um único ator. É difícil perceber a complexidade interna das caixas-pretas já que ela é “taken for granted [sic]”, se desloca para um fundo transparente ou estável demais para ser notado. (HOLANDA; LEMOS, 2013, p. 4).

Todo esse processo passa a designar uma rede a partir da descrição, uma vez que as

[...] mediações geram redefinições da realidade, tecendo novas geografias e novos controles. Novas coerções exercem pressão sobre velhas coerções e as deslocam, descentram. A produção em rede aponta para o transbordamento de tais campos que delinea novos ordenamentos na justa proporção de seus movimentos. A produção de um coletivo vai se dando por intermédio desses fluxos que se reapropriam da existência na construção do novo. (NOBRE; PEDRO, 2010, p. 48).

Latour (2012) é contundente no que se refere a descrições, descartando qualquer tipo de tentativa de explicar o mundo a partir de forças que substituam atores, uma vez que um ator que pode ser substituído é um ator que carrega um viés irrelevante em sua condição de actante. Ou seja, “a ANT [...] é uma grade negativa, vazia e relativista, que nos permite não sintetizar os ingredientes do social no lugar do ator” (LATOURE, 2012, p. 317).

A importância de descrever detalhadamente cada aspecto da rede traz consigo uma visão da realidade, aquilo que é anotado, vivido, pensado, visualizado. Assim, uma descrição diz respeito “àquilo que fazem para expandir, relacionar, comparar, e organizar é o que você deve descrever. Não se trata de outra camada a acrescentar à 'mera descrição'. Não tente passar da descrição à explicação: apenas vá em frente com a descrição”. (LATOURE, 2012, p. 217). Contamos com a busca constante para clarear essa face opaca da ciência e suas perspectivas, a partir da ANT, em relação com a tecnologia, incluindo aí os dados de pesquisa, e a autoria.

4.3 Organizando a narração: a Teoria Ator-Rede como método descritivo

A decisão metodológica faz com que escolhas sejam feitas. Por este motivo, a seleção não deve possibilitar a utilização de métodos que não “reconhecem [...] e postulam artefatos como agentes para a ação” (LEMOS, 2010, p. 17). O que se pretende, então, é fazer possível a utilização da Teoria Ator-Rede, uma vez que ela parece pertinente no que diz respeito ao estudo da autoria de dados científicos, possibilitando a agregação e descrição de redes sociotécnicas.

Como a ANT é proposta não exatamente como uma teoria propriamente dita, mas com características de um método, está situada como um instrumento e não como algo produzido a partir de um instrumento, preocupada na análise de eventos, sejam eles de qualquer natureza, bem como nas consequências impostas e/ou sofridas pelos atores envolvidos a partir desses eventos.

Esta teoria/método possibilita estudos a partir do momento em que as entidades são percebidas e analisadas sem que sua natureza seja critério de algum tipo de privilégio, sendo, deste modo, humanos e não-humanos essenciais para explicar as atividades sociais. Isto abre um leque de implicações, como argumenta Latour (2012), uma vez que as definições utilizadas pela Sociologia tradicional restringem o potencial informativo dos objetos não humanos, não os permitindo desempenhar papéis sociais.

Uma leitura do que está “por trás” de um discurso poderia compreender apenas a explicação do mundo a partir da justificativa do social, de um contexto – seja qual for –, isto é, um estágio ainda não emancipado do pensamento. O que queremos demonstrar é, pelo contrário, a possibilidade de “um relato racional daquilo que estão falando” (LATOURE, 2012, p. 79), sem fazer com que entes opacos ou sombrios relatados pelos atores sejam usados como forma de escape de um relatório que não representa estritamente os fatos observados.

É importante fazer referência, em relação a ANT, ao campo empírico no qual a pesquisa pretende ser aplicada. O que acontece, em geral, é uma delimitação do escopo de atores existentes, como um laboratório, por exemplo, onde as forças sociais agem de forma mais ou menos homogênea, caracterizadas por uma “mão invisível” estabilizadora. Em sentido contrário, “o que não se deve estabelecer logo

de início é a escolha de um *locus* privilegiado onde a ação porventura seja mais abundante” (LATOURE, 2012, p. 96), o que permite observar uma proporção de atores não pré-determinada. O aporte da ANT possibilita compreender relações de diversos tipos e que não estão, em geral, explícitas, isto é, não assumem uma linearidade clara e/ou aparente.

Como destaca Latour (2012), a delimitação de entidades envolvidas em um campo é uma perspectiva não adotada pela ANT. Com antecedência nem mesmo são conhecidos os atores envolvidos, nem mesmo quais tipos de associações existem – se é que existem – entre eles. As ações são oriundas de processos heterogêneos e que, por vezes, parecem desconexas, sendo isso aclarado pela narrativa da pesquisa.

A intenção é a aplicação do método que está ligado ao movimento e não à existência ou não de materialidade. Isto parece mais claro quando utilizamos um exemplo: sociedade. Não é sua característica material que a compõe de forma substancial, é seu movimento em busca de constantes associações que a conceituam. Em uma sociedade criada pela internet, por exemplo, seu constante movimento de actantes é que a designam como uma e criam todas suas implicações decorrentes dessa conceituação.

Sendo a ANT uma empreitada que consiste na ligação entre entidades, buscando traçar uma rede, permite combinar um processo de (re)agregação social, fazendo notar que vínculos diversos sejam descritos e, assim, analisados a partir de uma reordenação constante de informações.

Trannin e Pedro (2007) apresentam quatro passos que podem ser cumpridos para seguir atores e trazer à tona a rede:

- (1) Buscar uma porta de entrada: é preciso encontrar uma forma de “entrar na rede”, de acessá-la e, de algum modo, participar de sua dinâmica;
- (2) Identificar os porta-vozes: uma vez que da rede participam múltiplos actantes, humanos e não humanos, é preciso identificar aqueles que “falam pela rede” e que acabam por sintetizar a expressão de outros actantes;

- (3) Acessar os dispositivos de inscrição: ou seja, tudo o que possibilite uma exposição visual, de qualquer tipo, em textos e documentos, e que possibilite “objetivar” a rede sob a forma de indicadores;
- (4) Mapear as ligações da rede: trata-se de delinear as relações que se estabelecem entre os diversos atores e nós que compõem a rede.

Congruente a isso, McBride (1999) anota que uma investigação baseada na Teoria Ator-Rede pode ter como etapas de análise as seguintes peças:

- (a) identificar actantes;
- (b) investigar os atores-rede;
- (c) identificar as interações entre os atores-rede;
- (d) construir um modelo ator-rede, o que pode auxiliar na identificação dos laços entre os indivíduos;
- (e) identificar irreversibilidades, isto é, quando a hierarquia ou poder dos atores se apresentam de forma diferenciada, como pessoas, processos e/ou estrutura organizacional, por exemplo, advindas de questões políticas, econômicas etc.;
- (f) identificar inibidores e promotores, ou seja, aquilo que diz respeito a qualquer estrutura, indivíduo, dispositivo etc. que possa ser considerado algo que atrapalhe ou estimule a formação da rede;
- (g) identificar ações, possibilitando haver um alinhamento entre os atores-rede, isto é, uma percepção de seus atos.

Ainda que caiba ressaltar que os passos anotados por Trannin e Pedro (2007) e McBride (1999) não contemplam todas as situações de mapeamento e de estabelecimento de ações que a ANT propicia, visto a possibilidade de fenômenos que existem em uma infindável quantidade de objetos de estudo, são utilizados como um norte a ser compreendido nas questões que se impõem a partir da ANT. Assim, para este trabalho, esses passos serviram de guia em sua construção, com as especificidades que a ele são concernentes.

Nunca se deve esquecer que, no entanto, pensar as inscrições, seus dispositivos, seus atores-rede é, antes de qualquer coisa, pensar sobre suas relações e suas construções socialmente tensas devido às suas origens políticas.

Estes fatores requerem um processo de registro fiel ao que se propõe, não havendo possibilidade de estabilizações categóricas.

Neste sentido, é importante, como relato, fazer ecoar as palavras de Latour (2012, p. 88): “registrar e não filtrar, descrever e não disciplinar”. Optando então pela descrição, através de uma “lencitologia” – um processo lento, demorado e, ainda por cima, difícil –, buscamos traçar um relato da autoria no compartilhamento de dados de pesquisa, tomando por base uma teoria/método de seguir associações.

Estas observações ficam mais claras a partir do momento em que o texto desenvolvido relata a rede, ou seja, produz “um bom relato como aquele que tece uma rede” (LATOURE, 2012, p. 189), onde atores não são tratados apenas como intermediários, mas sim como mediadores com potencial transformador. Latour discorre longamente – como de praxe – sobre o que entende ser um texto ANT:

Definirei um bom relato ANT como aquele que tece uma rede. Refiro-me com isso a uma série de ações em que cada participante é tratado como um mediador completo. Em palavras mais simples: um bom relato ANT é uma narrativa, uma descrição ou uma proposição na qual todos os atores fazem alguma coisa e não ficam apenas observando. Em vez de simplesmente transportar efeitos sem transformá-los, cada um dos pontos no texto pode ser tornar uma encruzilhada, um evento ou a origem de uma nova translação. Tão logo sejam tratados, não como intermediários, mas como mediadores, os atores tornam visível ao leitor o movimento do social. Assim, graças a inúmeras invenções textuais, o social pode ser tornar de novo uma entidade circulante não mais composta dos velhos elementos que antes eram vistos como parte da sociedade. O texto, em nossa definição de ciência social, versa portanto sobre quantos atores o escritor consegue encarar como mediadores e sobre até que ponto logra realizar o social. [...] O bom texto tece redes de atores quando permite ao escritor estabelecer uma série de relações definidas como outras tantas translações. (LATOURE, 2012, p. 189).

Pelo que foi até aqui exposto, o processo de descrição é intensamente defendido por Latour. A descrição – refletida nos objetivos deste trabalho – atua como eixo norteador da visualização da rede, de sua formação e de seu desmanche. Pensar a rede de atores é um processo somente realizado com a exaustiva descrição das incertezas. Nas palavras de Latour (2012, p. 169): “a solução [...] é

aprender a alimentar-se de incertezas, em vez de decidir de antemão como deve ser a aparência do conjunto de equipamentos do mundo”. A descrição, desta forma, permite que seja possível remover a opacidade da rede.

O bom texto nunca é um retrato não mediatizado daquilo que descreve – nem sequer é um retrato. Sempre faz parte de um experimento artificial para reproduzir e sublinhar os traços gerados por testes em que os atores se tornam mediadores ou os mediadores são transformados em fiéis intermediários. (LATOURE, 2012, p. 199).

Assim sendo, “para os sociólogos das associações, a regra é a *performance*³⁹ é aquilo que tem de ser explicado, a exceção perturbadora, é qualquer tipo de estabilidade a longo prazo e em larga escala” (LATOURE, 2012, p. 60).

O importante, neste sentido, é que os conceitos propostos, aliados ao que é trabalhado em relação à autoria, permitem “brechas para pensar a importância da produção de saberes [...], ao fazer ou deixar falar aquilo que normalmente é silenciado no território da Ciência”. Assim sendo, “o que se pode encontrar [...] são pistas a um caminho possível de problematização e questionamento daquilo que está sendo feito e da organização de certos modos de fazer” (HÜNING; GUARESCHI, 2011, p. 68).

O compreender o processo, seguindo a perspectiva da Teoria Ator-Rede, é descrito pelas autoras:

O olhar para o processo não se centra exclusivamente na apresentação de resultados inequívocos ou conclusivos, de hipóteses confirmadas ou refutadas, ou de uma grande 'descoberta' científica [...]; mas na visibilização e reflexão sobre a própria maquinaria de produção da ciência, no circuito mobilizado nessa produção, da movimentação e articulações que, tendo acompanhado o fato científico, artefato ou tese científica, são usual e deliberadamente omitidos nas formas finais de comunicação da Ciência. (HÜNING; GUARESCHI, 2011, p. 69).

³⁹ Performance aqui está relacionada ao desempenho, ao alcançar algo, isto é, por um ato pronunciado a partir do relato.

Por fim, devemos compreender aqui que a ANT está ligada, neste trabalho, a explicar elementos que, de outra forma que não a partir da descrição, caracterizar-se-iam a partir especulações ou suposições para a construção conceitual do que vem a ser autoria. Sua ideia é, em síntese, servir como referencial conceitual-metodológico para considerar o papel dos atores-rede, permitindo visualizar a rede de mediadores e seus modos de atuação, partindo para uma análise da rede sociotécnica. Neste trabalho, a descrição do *locus* e da pesquisa em si são apresentados nos capítulos 5 e 6.

4.4 E depois do passeio: anseios entre a ANT e a autoria

A grande reivindicação da ANT é, como percebemos, opor o conceito usual de sociedade, colocando a perspectiva em deslocamento para a ideia de coletivo, sendo esta uma expansão daquela, sofrendo uma redefinição, o que se faz através de uma crítica constante à estrutura, ao contexto, ao social. O social, para a ANT, é visto como uma construção contínua, isto é, algo incessantemente produzido, coletado, sendo considerado uma empreitada que deve ser cumprida.

Já não existe sociedade, natureza, estrutura, contexto, nem mesmo atores e/ou redes preconcebidas. O preestabelecido caiu, ruiu. Aliás, nem foi cogitado. O que existe é fluxo, movimento. Este faz a sociedade, faz a natureza, faz a cultura, faz os atores, faz os vínculos. O que é social não é contexto posto de antemão e suas preconcebidas estruturas políticas, econômicas, jurídicas etc. O que é social é sua possibilidade de movimento, de fluxo. Em outras palavras, as possibilidades de associações.

Para a ANT, a relação inseparável entre ator e rede emerge a partir das associações mantidas naquele momento, das transformações criadas a partir da afetação entre os atores. [...] esses movimentos de tradução não obedecem a leis causais anteriores, nem se mantêm imutáveis no tempo. (PRIMO, 2012, p. 630).

A inexistência de estabilidade social ao se deparar com o estado de coisas inicial é perceptível. Instável no sentido de produzir constante movimentação, dinamismo nas relações, nos interesses, nas interações. A ordem só é atingida a partir de redes heterogêneas, fazendo com que actantes tomem parte de suas posições. Não há uma ou mais explicações generalistas – nível macro – que vislumbrem explicar o que acontece microssociologicamente, defende a ANT. O relato que tece a rede é composto por certa quantidade de mediadores e sua trama de elementos descritos performativamente com suas traduções.

Então buscamos as associações. Mas como? Um norte é a busca pela movimentação necessária destas que servem como uma destituição da entidade superior a que os sociólogos do social estão presos. Assim, as associações tomam a forma de redes, isto é, conexões contínuas que não são o resultado, mas uma perspectiva de investigação. E dentro dessa constituição o importante é que ressalte “muito mais vividamente [...] as conexões, fios, meios de transporte e veículos que ligam lugares” (LATOIR, 2012, p. 255).

Não deveremos dizer: 'Frente a um objeto, ignorem seu conteúdo e investiguem apenas os aspectos sociais que o cercam'. Digamos antes: 'Frente a um objeto, atendem primeiramente para as associações de que ele é feito e só depois examinem como ele renovou o repertório de laços sociais'. (LATOIR, 2012, p. 333-334).

É importante observar o lema dos estudos ANT: “siga os atores”, ou seja, atentar aos movimentos que permitem seguir as entidades que estão em circulação os fazendo atuar. A existência do homem está em constante relação com os instrumentos que ele mesmo constrói.

Law (1992, p. 390) sintetiza algumas questões que podem ser respondidas pela ANT:

Quais são os tipos de elementos heterogêneos criados ou mobilizados e justapostos para gerar os efeitos organizacionais? Como eles são justapostos? Como são superadas as resistências? Como é, se for o caso, que a durabilidade material e a

transportabilidade necessárias ao ordenamento organizacional das relações sociais são obtidos? Quais são as estratégias sendo performadas através das redes do social como uma parte do próprio? Até onde vão essas redes? Quão amplamente elas são performadas? Como elas interagem? Como é, se for o caso, que o cálculo organizacional é tentado? Como, se for o caso, são os resultados dos cálculos traduzidos em ação? Como, se for o caso, que os elementos heterogêneos que compõem a organização geram um relacionamento assimétrico entre centro e periferia? Como é, em outras palavras, que um centro pode vir a falar em nome dos esforços do que se tornou uma periferia e lucrar com esses esforços? Como é que um gerente gerencia? Vista desta forma, organização é uma conquista, um processo, uma consequência, um conjunto de resistências superadas, um efeito precário. Seus componentes – hierarquias, arranjos organizacionais, relações de poder, e fluxos de informação – são as consequências incertas da ordenação dos materiais heterogêneos⁴⁰. (LAW, 1992, p. 390).

Law (1992, p. 381) mostra, com clareza, como a ANT enxerga a ciência: “um processo de 'engenharia heterogênea' no qual elementos do social, do técnico, do conceitual, e do textual são justapostos e então convertidos (ou 'traduzidos') para um conjunto de produtos científicos, igualmente heterogêneos⁴¹”. Isto implica dizer que “pensar as redes implica também em *pensar com a rede*” (PEDRO, 2008, p. 3) como proposta de método, de atitude, de apreender a Teoria. Compreender a coisa como ela realmente é, é capturar sua abstração a partir de uma reconceitualização do que se entende como social.

Com uma visão na relação entre sociedade e tecnologia, McBride (2000) diz que a “ANT é uma abordagem para estruturar e explicar as relações entre sociedade e tecnologia. Ele oferece explicações de como a tecnologia se torna aceitável e é

⁴⁰ No original: “What are the kinds of heterogeneous bits and pieces created or mobilized and juxtaposed to generate organizational effects? How are they juxtaposed? How are resistances overcome? How it is (if at all) that the material durability and transportability necessary to the organizational patterning of social relations are achieved? What are the strategies being performed throughout the networks of the social as a part of this? How far do they spread? How widely are they performed? How do they interact? How it is (if at all) that organizational calculation is attempted? How (if at all) are the results of that calculation translated into action? How is it (if at all) that the heterogeneous bits and pieces that make up organization generate an asymmetrical relationship between periphery and center? How is it, in other words, that a center may come to speak for and profit from, the efforts of what has been turned into a periphery? How is it that a manager manages? Looked at in this way, organization is an achievement, a process, a consequence, a set of resistances overcome, a precarious effect. Its components-- the hierarchies, organizational arrangements, power relations, and flows of information are the uncertain consequences of the ordering of heterogeneous materials”.

⁴¹ No original: “a process of "heterogeneous engineering" in which bits and pieces from the social, the technical, the conceptual, and the textual are fitted together, and so converted (or "translated") into a set of equally heterogeneous scientific products”.

ocupada por grupos da sociedade. Ela sugere como a tecnologia é socialmente construída⁴².

A própria tecnologia utilizada desde o processo de compartilhamento até o processo de reuso de dados científicos, por exemplo, é um tipo de híbrido que não pode ser entendida como um intermediário. Pelo contrário, é um meio que, agindo como mediador, altera as formas, inclusive, de fazer ciência, atuando como reconfigurador de um processo informacional desde a produção até o reuso de dados científicos.

O aproveitamento das tecnologias da informação possibilita “rastrear associações de um modo antes impensável. [...] tornam visível o que antes só existia virtualmente, [...] permitindo] acompanhar o trabalho de construção do ator” (LATOUR, 2012, p. 299-300). Neste sentido fica claro que “seguir os atores” é “seguir os actantes”, seguir todos aqueles que interferem – medeiam – na rede.

No entanto, seguir, acompanhar passos, como propõe Latour, não significa uma relação de interação e investigação cotidiana com o laboratório, mas estar em posição de perceber como algo acontece no local de investigação. Ou seja, apreender as relações que são construídas e destruídas a partir dos actantes, humanos ou não, como artefatos tecnológicos ou artefatos conceituais.

Podemos tomar como arranjo de exemplo a própria questão autoral que neste trabalho atua como uma composição híbrida. Sua estrutura (o ator-rede autoria) vai além de uma forma e não possui nenhum tipo de desenho normatizado. Esse híbrido é composto, entre outros elementos, por sujeitos, dados, repositório de dados, redes eletrônicas, instituições etc. – isto é, humanos e não humanos. Com isto, a importância da ANT como teoria/método apropriada para estabelecer relações entre humanos e não humanos (ideias, coisas, inscrições etc.) e como essas entidades em que pesquisadores estão envolvidos – incluindo os próprios pesquisadores – fazem parte de grandes redes.

⁴² No original: “ANT is an approach to structuring and explaining the links between society and technology. It offers explanations of how technology becomes acceptable and is taken up by groups in society. It suggests how technology is socially constructed”.

O olhar proporcionado pela ANT extrapola a tecnologia como artefato, possibilitando que sejam compreendidas como tecnologias da inteligência, uma vez que

[...] as tecnologias de informação e comunicação podem ser entendidas como tecnologias da inteligência menos por projetar ou exteriorizar a riqueza e complexidade dos processos cognitivos do que por revelar o quanto sua complexidade deriva não apenas da riqueza de nossos sentidos e faculdades, mas também dos objetos, suportes, dispositivos e tecnologias que nos circundam e compõem uma rede sociotécnica de grande complexidade. (PARENTE, 2007, p. 103).

Podemos perambular pela ANT visando recolher informações sobre actantes realizando, basicamente, duas perguntas: (1) Se um elemento X não existisse, ele faria falta? (2) Se um elemento X não existisse, o que seria modificado? As respostas a estas duas questões estão intimamente ligadas pela ação de seguir as entidades que fazem parte do coletivo de associações, uma vez que revelam os elementos mais sensíveis de determinado campo.

Conhecer a estrutura que está posta a partir de um trabalho que agregue entidades heterogêneas é também um trabalho de superação de poderes, de resistências. A possibilidade de conhecer esses elementos e seus vínculos permite que consigamos revelar e ordenar sua coletividade em constante construção e desconstrução.

5 METODOLOGIA: TRAÇOS E LINHAS DE UM MÉTODO DESCRITIVO

“Nenhum pesquisador deve achar humilhante a tarefa de descrever – que, muito pelo contrário, é a realização máxima e mais rara”. (LATOURE, 2012, p. 199).

De caráter qualitativo, este estudo atua na abordagem do processo de forma mais aguda do que o resultado/produto, já que, como ressalta Latour, a rede formada pelos actantes está em constante movimento e o que se tem é um retrato momentâneo dos aspectos que podem ser coletados. Isto é, necessitamos de uma abordagem que compreenda o dinamismo da rede, uma vez que o mundo a ser observado e analisado é o movimento desta rede, possibilitando, como mostrado no objetivo geral proposto por este trabalho, analisar a autoria científica a partir do compartilhamento de dados de pesquisa.

Ressaltamos alguns princípios sobre o método qualitativo – e que são utilizados como norteadores da pesquisa – elencados por Appolinário (2006):

- (a) A análise dos dados pode ocorrer desde o momento da coleta, uma vez que o pesquisador observa, descreve e analisa impressões dos eventos. Appolinário (2006, p. 159) vai nesta direção, dizendo que “[...] a abordagem qualitativa apresenta certos elementos – como a recursividade, por exemplo – que implicam o fato de a análise poder se iniciar até mesmo ao longo da fase de coleta de dados”;
- (b) a análise é sistemática e compreensiva, ou seja, está relacionada ao rigor científico e não a uma rigidez capaz de delimitar de forma negativa o objeto investigado, permitindo que, dessa maneira, observações aprofundadas sejam feitas;
- (c) ao final é possível analisar os pormenores observados e sistematizados a partir do repertório teórico-conceitual-metodológico que se estabeleceu para tal.

Podemos também caracterizar esta pesquisa como descritiva, já que tem manifesta predileção pela descrição dos elementos que compõem certo fenômeno,

fazendo com que, à luz da ANT, possibilite compreender os actantes em seus percursos como atores-rede, isto é, suas ligações e suas movimentações.

5.1 O *locus* de pesquisa

O *locus* desta pesquisa é o Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio). Este foi criado pelo Ministério da Ciência, Tecnologia e Inovação (MCTI)⁴³ em 2004 e está ligado à Convenção sobre Diversidade Biológica (CDB) da Organização das Nações Unidas e às Diretrizes da Política Nacional de Biodiversidade.

O Programa de Pesquisa em Biodiversidade (PPBio) foi criado objetivando

articular a competência regional e nacional para que o conhecimento da biodiversidade brasileira seja ampliado e disseminado de forma planejada e coordenada. Os objetivos específicos do Programa são: (a) apoio à implantação e manutenção de redes de inventário da Biota; (b) apoio à manutenção, ampliação e informatização de acervos biológicos do País (coleções *ex situ*); (c) apoio à pesquisa e desenvolvimento em áreas temáticas da biodiversidade; (d) desenvolvimento de ações estratégicas para políticas de pesquisa em biodiversidade. (PPBIO, 2005, p. 18-19).

A estrutura básica do PPBio está disposta em três componentes: (a) Coleções biológicas; (b) Inventários Biológicos; e (c) Projetos Temáticos. O primeiro componente é responsável pelo suporte e manutenção de herbários, coleções vivas etc., atuando como fonte de informação sobre a composição, distribuição e conteúdo da biodiversidade de um ambiente. O segundo componente possibilita, através de levantamentos da biodiversidade, sítios de coleta, dados e metadados, disseminar informações a partir de levantamentos biológicos na Amazônia. Já o terceiro componente visa desenvolver métodos para o manejo sustentável da biodiversidade e bioprospecção. (PPBIO, 2014b).

⁴³ Na época Ministério de Ciência e Tecnologia (MCT).

O Programa também conta com Núcleos Regionais, os quais possibilitam

realizar de forma descentralizada e autônoma as ações definidas em conjunto com os Núcleos Executores; identificar áreas prioritárias para projetos de inventário e implantação de Sítios de Coleta Padronizada; elaborar o diagnóstico local das necessidades na região com relação aos estudos e aproveitamento da biodiversidade, de forma a promover a eficácia na utilização dos recursos humanos regionais. (PPBIO, 2014a).

Em relação ao Repositório, trata-se de um sistema via web que visa prover acesso a um banco de dados que contenha dados científicos sobre a biodiversidade Amazônica, e que parte do princípio que a coleta de dados sobre os indivíduos que compõem esse ecossistema deve estar associada a uma documentação (metadados) e disponibilizada em acesso livre para pesquisadores interessados no tema (PPBIO, 2015).

A escolha por este Repositório se deve a sua estrutura no que tange, principalmente, a sua explícita política de dados – como parte do PPBio –, publicada na Portaria 693 no Diário Oficial da União de 20 de agosto de 2009. A existência de uma política de dados oficialmente instituída por Portaria do Ministro do Estado de Ciência e Tecnologia, consonante com os objetivos do Programa de Pesquisa em Biodiversidade, ocorre devido à

[...] necessidade de um arcabouço de princípios, regras e orientações para todos os participantes do Programa de Pesquisa em Biodiversidade e usuários das bases de dados geradas no âmbito do Programa, no que diz respeito à abrangência, à coleta, ao armazenamento, à propriedade, à autoria, ao compartilhamento, à citação, ao acesso e uso de dados e das bases de dados. (BRASIL, 2009, p. 8).

Isso fortalece os parâmetros que devem ser seguidos para a coleta e a disponibilização de dados científicos, permitindo, segundo o a política publicada no Diário Oficial da União (BRASIL, 2009, p. 8), a “geração e disseminação de

informações e conhecimento sobre a biodiversidade brasileira”, afetando questões “econômicas e ambientais”.

5.2 Observação e entrevista

No que se refere à observação como método de pesquisa, Richardson (2010, p. 259) relata que ela, “sob algum aspecto, é imprescindível em qualquer processo de pesquisa científica, pois ela tanto pode conjugar-se a outras técnicas de coleta de dados como pode ser empregada de forma independente e/ou exclusiva”, ou seja, podemos considerar que a observação “é a base de toda investigação no campo social[, permitindo ...] o exame minucioso ou a mirada atenta sobre um fenômeno no seu todo ou em alguma de suas partes; é a captação precisa do objeto examinado” (RICHARDSON, 2010, p. 259).

Para fins deste estudo, a utilizamos na forma não-participante, já que a função investigativa, neste caso, não pretende “toma[r] parte nos conhecimentos objeto de estudo como se fosse membro do grupo observado, mas apenas atua[r] como espectador atento” (RICHARDSON, 2010, p. 260). A observação vai além, permitindo o exame de “documentos (livros, jornais, impressos diversos) na medida em que estes não somente nos informam dos resultados das observações e pesquisas feitas por outros mas traduzem também a reação dos seus atores” (RUDIO, 1978, p. 39).

Uma vez que o trabalho propõe sua investigação em um domínio de conhecimento especializado – compartilhamento de dados digitais de pesquisa do Programa de Pesquisa em Biodiversidade –, além da pesquisa observacional, foram realizadas entrevistas com gestores dos dados no portal PPBio que, além de se caracterizarem como curadores, são especialistas no domínio e autores de dados científicos e documentos bibliográficos.

A entrevista pode ser considerada uma forma flexível de coleta de dados, caracterizada pela forma de interação social entre o entrevistador e o entrevistado, com a finalidade de coletar dados a partir dessa fonte de informação (HAGUETTE, 1987). Aliado a isso, a entrevista é um procedimento metódico, podendo ser

considerado um instrumento de investigação social por excelência (LAKATOS; MARCONI, 2007), permitindo, enquanto instrumento científico, “a tentativa de captação do real, sem contaminações indesejáveis nem da parte do pesquisador nem de fatores externos que possam modificar aquele real original” (HAGUETTE, 1987, p. 81).

Para esta pesquisa foi utilizada a técnica de entrevista como método capaz de permitir flexibilidade no instrumento de coleta de dados, bem como possibilitar ao entrevistador a condução da situação pelo caminho que melhor se adequar ao objeto de estudo. Assim sendo, foi buscada a modalidade focalizada de entrevista (FLICK, 2009) que, embora não siga uma estrutura formal, permite a utilização de roteiros expostos em ordem não pré-estabelecida, buscando esclarecimentos sobre a temática envolvida (Apêndice B).

Para a realização das entrevistas foi apresentado aos participantes um termo de consentimento, o qual tem a função de exprimir “a anuência livre e esclarecida de participação de alguém como sujeito da pesquisa” (PESCUMA; CASTILHO, 2005, p. 75). Esse termo foi assinado pelos participantes que aceitaram realizar a atividade (Apêndice A).

A seguir é descrito como foi realizado o processo de coleta e de análise dos dados.

5.3 Sobre a coleta e a análise de dados

É necessário que façamos, nesta seção, o relato sobre a coleta e a análise de dados realizada. Neste sentido, estando nosso estudo ligado ao conceito de autoria de dados científicos digitais a partir de seu compartilhamento, fica latente que sua efetivação como objeto está constituída no repositório de dados científicos.

Iniciamos nossa incursão de busca por um repositório desse tipo e vislumbramos algumas possibilidades, mas foram estabelecidos dois critérios básicos que deveriam ser obedecidos para que um fosse escolhido para a investigação: (a) deveria ser um repositório em âmbito nacional. Esta especificação

permitiria, além da investigação proposta, obter uma visão do cenário nacional em relação ao tratamento de dados de pesquisa; (b) deveria dispor de uma explícita política de dados, mesmo que ampla, relacionada aos ativos que ali se encontram. Isso possibilita que critérios mínimos sejam estabelecidos durante todo o ciclo de vida dos (conjuntos de) dados, demonstrando a preocupação do sistema desde a coleta até a disponibilização.

Ao buscarmos por um repositório de dados que pudesse servir como *locus* de pesquisa, uma seleção foi realizada e elegemos o Repositório de Dados Ecológicos do Programa de Pesquisa em Biodiversidade, como já mencionado. A composição do Repositório possibilita que, uma vez publicamente disponíveis, os dados “possam ser usados por agências governamentais, ONGs e setor privado para gerenciamento de decisões”⁴⁴ (PEZZINI et al, 2012, p. 266).

Para que a pesquisa fosse realizada, dois passos foram considerados essenciais, a saber: (1) o conhecimento sobre o sistema que disponibiliza os dados, isto é, o repositório em si. Nesse passo entendemos que isso engloba desde a utilização dos aplicativos disponibilizados para envio de dados e metadados – Morpho – bem como o sistema que mantém o Repositório ativo para acesso aos dados de investigação – Metacat⁴⁵ –, sendo essencial que sejam conhecidas suas funcionalidades e interfaces; (2) entrevistas realizadas com os gestores/curadores do repositório, com a finalidade de compreender como estão dispostas as relações entre dados-repositório e a autoria em si, uma vez que os gestores/curados também são, em algum momento, participantes de pesquisas e autores de dados científicos.

A etapa seguinte foi o contato com os responsáveis pelo Programa e pelo Repositório. O primeiro contato com os pesquisadores do Programa de Pesquisa em Biodiversidade (PPBio) foi realizado em 12 de março de 2014 via e-mail. Foi realizada a busca pelos coordenadores do Programa, visando apresentação do pesquisador e, em linhas gerais, como o projeto se constituiria, bem como para saber a viabilidade para realização da pesquisa junto ao Repositório de Dados de Estudos Ecológicos do PPBio e os indivíduos responsáveis pelo seu gerenciamento.

⁴⁴ No original: “can be used by government agencies, NGOs and the private sector for management decisions”.

⁴⁵ Morpho é o software utilizado para envio de dados ao repositório, gerenciado pela aplicação Metacat. Ambos são descritos posteriormente.

Feito isso, e com a autorização confirmada, foi repassada, pelo investigador principal do PPBio, uma listagem com os nomes e endereços eletrônicos dos responsáveis pelo gerenciamento do Repositório, permitindo que fosse possível contatá-los para entrevistas. As entrevistas serviram para compreensão sobre o funcionamento do Repositório no aspecto operacional, tanto tecnológico quanto político.

Essa etapa se constituiu em trajeto importante para a pesquisa, uma vez que o sistema de armazenamento e recuperação como um todo influi no processo de análise, já que é composto pelo Repositório propriamente dito e pela sua interface que disponibiliza as informações básicas sobre os dados a partir dos metadados, além de possibilitar a coleta de informações sobre as ações do Repositório no que tange à sua funcionalidade. Em conjunto, existem aspectos que vão além da questão técnica e tecnológica, como a política de dados que é estabelecida pelo Programa, além de outros elementos que puderam ser coletados ao longo do trabalho⁴⁶.

Para a coleta de dados foram realizadas entrevista com três gestores do repositório de dados PPBio entre setembro e novembro de 2014. Essas pessoas foram destacadas e se encaixam no perfil de porta-vozes, já que são ou foram responsáveis pela curadoria/gestão do repositório em épocas diferentes, desde sua implementação até os dias atuais, bem como atuam como coletores de dados e produtores de material bibliográfico a partir da utilização dos dados científicos coletados.

Com a finalidade de obter dados relevantes, buscamos concentrar a pesquisa nesses indivíduos, convidando-os a participar de entrevistas online, separadas, via software de voz sobre IP⁴⁷ (VoIP) Skype⁴⁸ com gravações de áudio com o software MP3 Skype Recorder⁴⁹. As entrevistas foram gravadas para que pudéssemos, além de anotar determinadas informações relevantes durante a fala dos entrevistados, possibilitar a (re)investigação contínua das informações ali contidas, contando também com transcrições realizadas. Essas sessões de entrevistas foram possíveis

⁴⁶ Os elementos, considerados actantes, serão trabalhados nas seções seguintes.

⁴⁷ Tecnologia que permite comunicação multimídia – principalmente áudio e vídeo – através de redes de computadores, como a internet

⁴⁸ <http://www.skype.com>

⁴⁹ <http://voipcallrecording.com>

graças à disponibilidade dos entrevistados e sua compreensão sobre a importância da questão que foi proposta.

As questões propostas aos entrevistados (Apêndice B), como é possível notar ao analisá-las, não são fixas e/ou fechadas em si mesmas, já que têm a intenção de criar um arcabouço em forma de rede e seu emaranhado é construído em conjunto com outros blocos ou temas, sendo seu estabelecimento buscador de clareza aos relatos. Como de fato ocorre em uma entrevista semiestruturada, as falas dos declarantes não são lineares e assuntos que seriam abordados em momentos posteriores podem ser discutidos de forma antecipada e/ou retomados após certo período de tempo.

Outro ponto relevante são os documentos relativos ao Repositório e a sua sistemática, como a Política de dados do Programa de Pesquisa em Biodiversidade e publicações em periódicos e apresentações em congressos realizadas por membros do PPBio. Primeiramente, a recolha desse tipo de material foi feita antes das entrevistas, já que foi necessário conhecer de forma aprofundada o próprio Programa e o Repositório, além de como está expressa a Política que lida com os dados científicos ali depositados. Em momento posterior, foram realizadas coletas a partir de informações obtidas junto aos gestores do Repositório, bem como a partir de indicações dos mesmos.

Assim, as entrevistas também possibilitaram conhecer documentos sobre a forma de operar, o histórico do Repositório e sua missão. Alguns artigos foram enviados por e-mail para que pudessem ser utilizados. Outros trabalhos tiveram seu link fornecido para acesso.

Na questão narrativa, optamos por trazer o mínimo possível da fala direta dos entrevistados. Esta medida visa, além de promover maior fluidez ao texto, incentivar a plena descrição em busca de pontos que possam ser relevantes, sem nos deixarmos levar pelo discurso em si. Tal ponto nos parece convergente às ideias de Latour, pensando a Teoria Ator-Rede como uma forma de construção contínua. Para as falas dos entrevistados, adotamos as reduções E1, E2 e E3 para, respectivamente, “Entrevistado 1”, “Entrevistado 2” e “Entrevistado 3”, com o intuito de preservar o anonimato dos declarantes.

Ao imaginarmos um mapa recortado, buscamos uma descrição de uma fotografia, propondo-nos a relatar o que pode ser compreendido como a autoria de dados científicos a partir de seu compartilhamento, sendo isso consoante aos anseios deste trabalho, expressos nos objetivos geral e específicos.

Há, neste sentido, um caminho para o afastamento do intervencionismo, uma vez que entendemos que qualquer observação de um processo, seja ela participante ou não, acarreta na modificação do estado de coisas. Deste modo, o distanciamento do objeto, embora tenhamos consciência da dificuldade e da provável impossibilidade de neutralidade, foi uma busca constante que pretendemos demonstrar nas seções a seguir.

6 PESQUISA EMPÍRICA: A DESCRIÇÃO DO OBJETO

6.1 O Repositório de Dados de Estudos Ecológicos do Programa de Pesquisa em Biodiversidade (PPBio): a interface de acesso aos dados e aos metadados

O Repositório de Dados de Estudos Ecológicos⁵⁰ é gerido pelo Programa de Pesquisa em Biodiversidade⁵¹ (PPBio), sendo este um programa de pesquisa do Instituto Nacional de Pesquisas da Amazônia⁵² (INPA). O programa foi criado pelo Ministério da Ciência, Tecnologia e Inovação⁵³ (MCTI) em 2004 e está ligado à Convenção sobre Diversidade Biológica (CDB) da Organização das Nações Unidas e às Diretrizes da Política Nacional de Biodiversidade.

O Programa é composto por diversas ações relacionadas aos estudos ecológicos, compreendendo desde a coleta e armazenamento de coleções biológicas, treino de pessoal especializado em criação de taxonomias e produção de guias para pesquisa de campo e bioprospecção, isto é, organismos que possuam valor econômico e possam ser utilizados comercialmente. Percebendo essas questões, um aspecto que permeia todos componentes é a publicação de dados coletados em sítios delimitados pelo Programa (PEZZINI et al., 2012).

Após a coleta os dados devem se tornar disponíveis para o público em um breve período de tempo, sendo os metadados disponibilizados após 30 dias da coleta e os dados propriamente ditos em até 12 meses da sua recolha. Além do mais, pesquisadores com interesse em utilizar os dados são incentivados a contatar o(s) responsável(is) para discutir possíveis questões de autoria, como a utilização e direitos de uso, além da colaboração na manipulação dos dados (BRASIL, 2009; PEZZINI et al., 2012).

Assumindo que nossa pesquisa deve permitir, nesta etapa, a narrativa sobre a base de conjuntos de dados e metadados do PPBio, no momento⁵⁴ da escrita final deste trabalho são registrados 467 conjuntos de dados⁵⁵ dos mais variados tipos

⁵⁰ Disponível em: <http://ppbio.inpa.gov.br/knb/style/skins/ppbio/>. Acesso em: 13 out 2014.

⁵¹ Disponível em: <http://ppbio.inpa.gov.br/>. Acesso em: 13 out. 2014.

⁵² Disponível em: <https://www.inpa.gov.br/>. Acesso em: 13. out. 2014.

⁵³ Na época Ministério de Ciência e Tecnologia (MCT).

⁵⁴ Em 07 abr. 2015.

⁵⁵ Embora, como vimos, dados possuam definição complexa, podemos dizer que um conjunto de dados é um agregado, em geral distribuído a partir de um arquivo digital compactado, que contém um ou mais dados e/ou tabelas sobre a coleta realizada.

relativos a assuntos da biodiversidade. Estes conjuntos de dados são acessíveis através de uma interface web que apresenta seus metadados para descrição e consulta.

A busca por conjuntos de dados pode ser realizada a partir da documentação existente através de metadados, listando todos os conjuntos existentes no repositório ou, por padrão, realizar pesquisas em todos os campos disponíveis na documentação dos conjuntos. Também é possível restringir a consulta aos campos “título”, “resumo”, “palavras-chave” ou “pessoas envolvidas”. Ainda é possível utilizar o coringa “%” quando há interesse em encontrar conjuntos que revelem alguma similaridade linguística, como, por exemplo, a expressão “igarap%” possibilita resultados como “igarape”, “igarapé”, “igarapes”, “igarapés”. Isso pode ser verificado na Figura 4.

A partir da página de resultados é possível visualizar o rol de conjuntos de dados recuperados com base na expressão de busca e uma listagem a partir dos seus respectivos metadados “título”, “contato”, “organização” e “palavras-chave”, como consta na Figura 5, além da opção “Abrir”, a qual leva até a tela que mostra os metadados referentes a um conjunto de dados em sua totalidade (Figura 6).



Repositório de Dados do PPBio

Bem vindo ao Repositório de Dados do Programa de Pesquisa em Biodiversidade - PPBio. Esse repositório contém dados de levantamentos realizados no âmbito do Programa de Pesquisa em Biodiversidade e projetos parceiros. Os dados armazenados nesse repositório em breve estarão conectados à rede KNB ("Knowledge Network for Biocomplexity"), um repositório internacional de dados. Algumas das séries de dados contidas aqui foram geradas por pesquisadores em esforços individuais, enquanto outras são resultado de esforços conjuntos de grupos de contribuintes. As descrições de cada conjunto de dados contêm mais informações sobre pessoas e instituições envolvidas. Os dados seguem a política de dados do PPBio, disponível em <http://ppbio.inpa.gov.br/politicadados>.

Para perguntas, comentários e sugestões, por favor contate: ppbio@inpa.gov.br.

Busca por Dados

Buscar somente nos campos "Título", "Resumo", "Palavras-Chave", "Pessoas Envolvidas"

Buscar todos os campos

Esta ferramenta permite a busca por conjuntos de dados de interesse. Ao inserir um texto no quadrado e clicar no botão "Buscar", a busca será conduzida apenas nos campos "Título", "Resumo", "Palavras-Chave" e "Pessoas Envolvidas". Ao optar pela opção "Buscar todos os campos", a busca ocorrerá em todos os campos (isso fará com que a busca leve mais tempo).

O caractere "%" pode ser usado como um "coringa" (ou "wildcard") nas buscas (por exemplo, "%biodiversidade%" localizará qualquer frase que contenha a palavra biodiversidade).

Buscar todos os metadados cadastrados no Repositório do PPBio

Registrar Informação sobre um novo conjunto de dados do PPBio



Use o software **Morpho** para registrar, documentar e inserir dados.

O Morpho é uma ferramenta de gerenciamento para ecólogos. Essa ferramenta foi criada para prover um modo fácil e multi-plataforma de acesso e manipulação dos metadados (ex. documentação) e dados (ambos localizados na rede). O Morpho permite que ecólogos criem metadados, ou seja, descreveram seus metadados em um formato padrão, e criem um catálogo de dados e metadados de maneira que seja possível a busca, edição e visualização das coleções de dados. Além disso, o Morpho prover modos de acesso ao servidor da rede, para a busca, visualização e recuperação de dados ecológicos públicos!

Esse repositório é uma iniciativa do PPBio, baseado em um software desenvolvido pelo "Knowledge Network for Biocomplexity" (KNB) e acomoda metadados segundo o padrão EML (Ecological Metadata Language)

Figura 4: Tela de busca do repositório

Fonte: Repositório de Dados Ecológicos do PPBio



Página do PPBio Repositório de Dados

70 data packages found

Title	Contacts	Organization	Keywords	Actions
» Abundância e Distribuição de Ervas Terrestres em Parcelas Ripárias na Reserva Ducke: Variação Lateral ID: drucker.3.10	Drucker	Instituto Nacional de Pesquisas da Amazônia – INPA	Ervas Parcelas Ripárias Reserva Ducke Floresta de Terra Firme PELD-PPBio	Abrir
» Abundância e Distribuição de Ervas Terrestres em Parcelas Ripárias na Reserva Ducke: Variação Longitudinal ID: fecosta.42.9	Drucker	Instituto Nacional de Pesquisas da Amazônia – INPA	Ervas Reserva Ducke Parcelas Ripárias Floresta de Terra Firme PELD-PPBio	Abrir
» Altitude, distância de Igarapé e número de árvores do módulo de Cotríguaçu, Sítio Sinop, Mato Grosso, Brasil. ID: naman.183.3	de Jesus Rodrigues	Universidade Federal do Mato Grosso – UFMT	PPBio Cotríguaçu Sítio Sinop Número de árvores Distância do Igarapé Altitude Variáveis Ambientais	Abrir
» Características Estruturais de Igarapés de Terra Firme da Reserva Ducke ID: menger.39.4	Pereira de Mendonça	Instituto Nacional de Pesquisas da Amazônia – INPA	Igarapés de Terra-Firme Características Estruturais Tipos de Substratos Parcelas Permanentes Aquáticas Reserva Florestal Adolpho Ducke Amazônia Central	Abrir

Figura 5: Tela com os resultados da busca

Fonte: Repositório de Dados Ecológicos do PPBio

Data Set Citation:	
When using this data, please cite the data package: Pereira de Mendonça F. Características Estruturais de Igarapés de Terra Firme da Reserva Ducke menger.39.4 (http://ppbio.inpa.gov.br/knb/metacat/menger.39.4/ppbio)	
General Information:	
Title:	Características Estruturais de Igarapés de Terra Firme da Reserva Ducke
Identifier:	menger.39.4
Abstract:	Os dados aqui disponibilizados são parte das informações disponíveis para as 38 Parcelas Permanentes Aquáticas distribuídas sobre os igarapés da Reserva Florestal Adolpho Ducke, Manaus, Amazonas, Brasil.
Keywords:	<ul style="list-style-type: none"> • Igarapés de Terra-Firme • Características Estruturais • Tipos de Substratos • Parcelas Permanentes Aquáticas • Reserva Florestal Adolpho Ducke • Amazônia Central • PELD-PPBio • PELD-PPBio
Data Table, Image, and Other Data Details:	
Metadata download:	Ecological Metadata Language (EML) File
Data Table:	Caract_estruturais.txt (View Metadata Download File 

Involved Parties

Data Set Owners:	
Individual:	Fernando Pereira de Mendonça
Organization:	Instituto Nacional de Pesquisas da Amazônia – INPA Coordenação de Pesquisas em Biologia Aquática – CPBA,
Address:	Avenida André Araújo, 2936. Aleixo, CEP 69011-970, Manaus, Amazonas Brasil
Phone:	+55 92 3643 3253 (voice)
Email Address:	mendonca@inpa.gov.br
Web Address:	http://lattes.cnpq.br/3772778427025258

Figura 6: Excerto da tela que contém o registro completo de conjunto de dados

Fonte: Repositório de Dados Ecológicos do PPBio

Já na tela referente ao conjunto de dados (Figura 6), são encontradas *tags* (etiquetas) que descrevem esses dados. Essas *tags* são baseadas em EML (Ecological Metadata Language) e têm a função de descrever os dados e suas metainformações, sendo compostos por diversos itens, a saber:

- (a) Data set citation: instrução para a utilização dos conjuntos de dados, da forma como deve ser realizada a citação de modo que segue a orientação, inclusive com indicação de link direto para o conjunto.
- (b) General information: consta título do conjunto de dados; identificador único: permite acesso direto; abstract: pode compreender objetivos da pesquisa que deu origem aos dados, bem como a razão do seu uso inicial ou a descrição

- geral dos próprios dados; palavras-chave: expressões que descrevem/representam o conjunto de dados e servem para auxiliar a sua recuperação.
- (c) Data table, Image, and other data details: neste item é possível visualizar e/ou fazer o download dos metadados em formato EML enviados através de software específico para esse fim; fazer download do(s) conjunto(s) de dados; visualizar os atributos que descrevem os elementos coletados. Este último item é variável, uma vez que os metadados refletem a descrição daquilo que foi coletado, fazendo com que cada coleta possa responder por dados diferentes dos já existentes.
- (d) Involved parties: composto pelos itens: Data set owners: pessoa ou organização a quem é creditada a criação dos dados, como pesquisadores, bolsistas, estudantes etc.; Data set contacts: pessoa ou instituição principal para contato em caso de utilização e/ou interpretação dos dados; Associated parties: pessoas ou organizações que tenham relação com a coleta de dados, como, por exemplo, criador de metadados (Custodian/Steward). Inclui as seguintes informações:
- i. Individual: nome
 - ii. Organization: instituição a que está vinculado
 - iii. Address: endereço pessoal
 - iv. Phone: telefone para contato
 - v. Email address: endereço de e-mail
 - vi. Web Address: links de acesso ao website pessoal, institucional, do currículo etc.
- (e) Data set characteristics: como o próprio nome denota, diz respeito aos atributos do conjunto de dados, como: Geographic Region, composta por Geographic Description: local onde os dados foram coletados; Bounding Coordinates: coordenadas precisas do local de coleta; Time Period: delimitação das datas de início e fim da coleta.
- (f) Sampling, Processing and Quality Control Methods: pode conter métodos da coleta de dados e instrumentos utilizados para tal.
- (g) Funding Project: informações sobre o projeto de pesquisa (coordenador e descrição) a que a coleta está vinculada e possíveis financiadores.

(h) Data Set Usage Rights: deve conter informações sobre acesso e uso dos dados. Deve estar consonante com a política de dados existente no PPBio.

Após essa descrição sobre o *locus* de pesquisa, interessa-nos relatar o conteúdo dos dados coletados tanto na observação, através das entrevistas e de documentos. As informações obtidas através de entrevistas com os gestores envolvidos com o repositório de dados PPBio foram analisadas e comparadas, visando aglutinar implicações observadas para o Repositório e para a questão autoral. Ressaltamos que, a partir da determinação imposta pela ANT, não se trata de uma comparação e/ou substituição de uma descrição do coletivo em favor de uma análise discursiva.

6.2 O Repositório e suas ferramentas

Antes de entrar na questão específica dos dados científicos, é importante relatar sobre um aspecto fundamental: a importância de existir um local onde os dados estejam depositados. Embora muitas coletas e pesquisas sobre biodiversidade sejam realizadas em ambientes razoavelmente conhecidos e relatados, como em um sítio ecológico, há necessidade de saber o que está sendo feito ou coletado pelos investigadores. Como encontrar e/ou utilizar dados de outros pesquisadores? Como manter isso a longo prazo? Como reunir a dispersão de dados coletados por grupos próximos? Estas questões foram o mote central da necessidade⁵⁶ de um repositório que se adequasse às exigências daqueles que buscam dados ecológicos.

Pensando nisso, trabalhar com um repositório composto por, basicamente, um banco de dados estruturado onde são apresentadas tabelas formalmente definidas difere da realidade das pesquisas realizadas em Ecologia. Ou seja, a prévia estruturação de um banco de dados pode não atender as necessidades dos

⁵⁶ É imperativo que se frise o termo “necessário”. Uma vez que há detectado uma necessidade (um local onde dados possam ser armazenados e recuperados) há a busca por uma solução (a criação de um repositório). No entanto, isso não significa que todas as carências detectadas devam ser sanadas logo de início. Pelo contrário, a possibilidade de iniciar um projeto faz com que algumas soluções sejam implementadas já com o andamento do plano inicial.

pesquisadores que coletam dados, caso haja incompatibilidade de campos, usualmente dispostos em colunas, a serem preenchidos. De certo modo, isso é simples de ser visualizado: ao se trabalhar com aves, as questões podem dar-se a partir do comprimento das asas, mas o pesquisador pode necessitar produzir uma nova variável, como o comprimento das patas, isto é, uma variável que não é comum para o estudo de aves. Não há como prever se os dados que são coletados para atender às características do animal estarão cobertos pelo banco de dados já constituído (E1; E2; E3).

A utilização de ferramentas que propiciem o depósito dos conjuntos de dados em associação com os metadados é, nesse sentido, fundamental. O software Metacat é utilizado pelo PPBio como forma de agregar essas questões – conjuntos de dados e metadados – atuando no gerenciamento de dados científicos, principalmente sobre ecologia e ciências ambientais. Faz isso a partir da sintaxe XML (eXtensible Markup Language), o qual, a partir das recomendações da W3C (World Wide Web Consortium), define um conjunto de regras para a codificação de documentos em formato legível por máquinas e humanos.

O servidor Metacat é um repositório de metadados e dados que faz com que os conjuntos de dados possam ser descobertos através de metadados consistentemente descritos. Este sistema permite o armazenamento de qualquer tipo de dado como um arquivo anexado nos metadados, enquanto armazenar os dados em si em um sistema de banco de dados como MySQL ou PostgreSQL requerem atributos definidos e fixos. Tais sistemas de banco de dados podem ser desenvolvidas para fins específicos a partir do repositório de dados (por exemplo, DB para medições de biomassa de árvore)⁵⁷. (PEZZINI et al., 2012, p. 270).

Esse empreendimento é importante porque possibilita, além do armazenamento e gerenciamento, um tipo de padronização sobre os campos dos metadados com a possibilidade de busca, comparação, aglomeração e, por consequência, reúso do conjunto de dados, possibilitando que novas descobertas

⁵⁷ No original: “The META-CAT server is a repository for metadata and data that makes datasets discoverable through consistently described metadata. This system allows the storage of any kind of data as an attached file in the metadata, whereas storing the data itself in a database system such as MySQL or PostgreSQL requires defined and fixed attributes. Such database systems can be developed for specific purposes from the data repository (e.g. DB for tree biomass measurements).”

sejam realizadas com base nos dados coletados por outros investigadores. Ou seja, permite que o depósito seja um processo menos custoso para o repositório no que se refere à documentação dos dados, além de fornecer ferramentas para busca e recuperação.

Dessa forma, o propósito do Metacat, tendo como base sua função principal de servir como repositório, possibilita a recuperação de dados distribuídos em diversos servidores ao redor do mundo, e não somente a partir de buscas locais. O problema enfrentado nesse caso diz respeito à impossibilidade de consulta entre mais de uma tabela, o que pode ser um impeditivo para a comparação e/ou combinação das mesmas (DRUCKER, 2011).

Visando buscar solução quanto à questão documental, no Metacat é utilizado o EML (Ecological Metadata Language), que está baseado no XML. Esse padrão foi pensado a partir do aumento “da profundidade, da amplitude e da complexidade que a ecologia tem sido acompanhada através de mudanças significativas no tipo, magnitude e complexidade dos dados que são adquiridos e analisados pelos ecologistas”⁵⁸ (MICHENER, 2006, p. 3).

Com isso se pretende solucionar dois pontos:

Em primeiro e mais importante lugar, definir uma estrutura comum que todos os ecologistas possam usar para documentar dados ecológicos para que outros ecologistas possam interpretar corretamente os dados. A segunda finalidade para o EML é proporcionar uma estrutura para que as aplicações de software possam ser desenvolvidas. Os tipos de aplicações previstas variam de uma ferramenta básica que permite ao usuário realizar pesquisas muito específicas orientadas para os conjuntos de dados para aplicações avançadas que poderiam executar a integração automática de conjuntos de dados⁵⁹. (BLANKMAN; MCGANN, 2003).

⁵⁸ No original: “the depth, breadth and complexity of ecology has been accompanied by significant changes in the types, magnitude and complexity of data that are acquired and analyzed by ecologists”.

⁵⁹ No original: “The first and most important purpose is to define a common structure that all ecologists can use to document ecological data so that other ecologists can correctly interpret the data. The second purpose for EML is to provide a structure so that software applications can be developed. The kinds of applications anticipated range from a basic tool that allows a user to perform very specific targeted searches for datasets to advanced applications that could perform automatic integration of datasets”.

Dessa maneira, o que é provido a partir do EML é uma descrição dos dados – a partir dos metadados – que podem incluir unidades de medida e precisão, entre outras características que possam ser relevantes. Isso permite que os softwares disponíveis, como o Metacat, possam lidar algoritmicamente com a informação ali escrita e produzir resultados consistentes para a busca, mesmo em ambientes com grandes quantidades de conjuntos de dados.

O EML apresenta, para tal, uma estruturação em conjunto de módulos que permite as funções relatadas acima, podendo ser assim apresentadas (DRUCKER, 2011; MICHENER, 2006):

- (a) EML-resource: Informação geral sobre os metadados, como quem os criou, título e palavras-chave;
- (b) EML-party: Informação detalhada sobre pessoas e organizações envolvidas
- EML-coverage: Abrangência temporal, geográfica e temporal do conjunto de dados
- EML-project: Contexto da pesquisa que originou o conjunto de dados, por exemplo, objetivos e financiamento
- (c) EML-methods: Procedimentos de campo e de laboratório usados para criar o conjunto de dados e procedimentos de controle de qualidade dos dados
- (d) EML-attribute: Nome, definição, unidade e domínio das variáveis
- (e) EML-access: Níveis de acesso permitidos para usuários ou grupos de usuários.

Com relação a edição de metadados, o repositório utiliza a ferramenta Morpho, software que fornece uma interface para criação de arquivos com metadados já codificados em EML, além da busca, visualização e download de pacotes de dados já existentes. Essa ferramenta também possibilita a busca de metadados nas redes interna (local) e externa (rede de servidores de dados interligados) a que o sistema está instalado, possibilitando que uma recuperação com maior especificidade possa ser realizada⁶⁰. Ademais, permite também que sejam ligadas duas tabelas, criando sistemas referenciais. As telas iniciais do Morpho compõem as Figuras 7 e 8.

⁶⁰ O manual de utilização do software Morpho – Armazenamento e disponibilização de dados em ecologia utilizando o software Morpho – produzido pelo PPBio e que conta com um pass-a-passo sobre a utilização da aplicação está disponível em: http://ppbio.inpa.gov.br/sites/default/files/Manual_Morpho_PPBio_2012.pdf. Acesso em: 13 out. 2014.

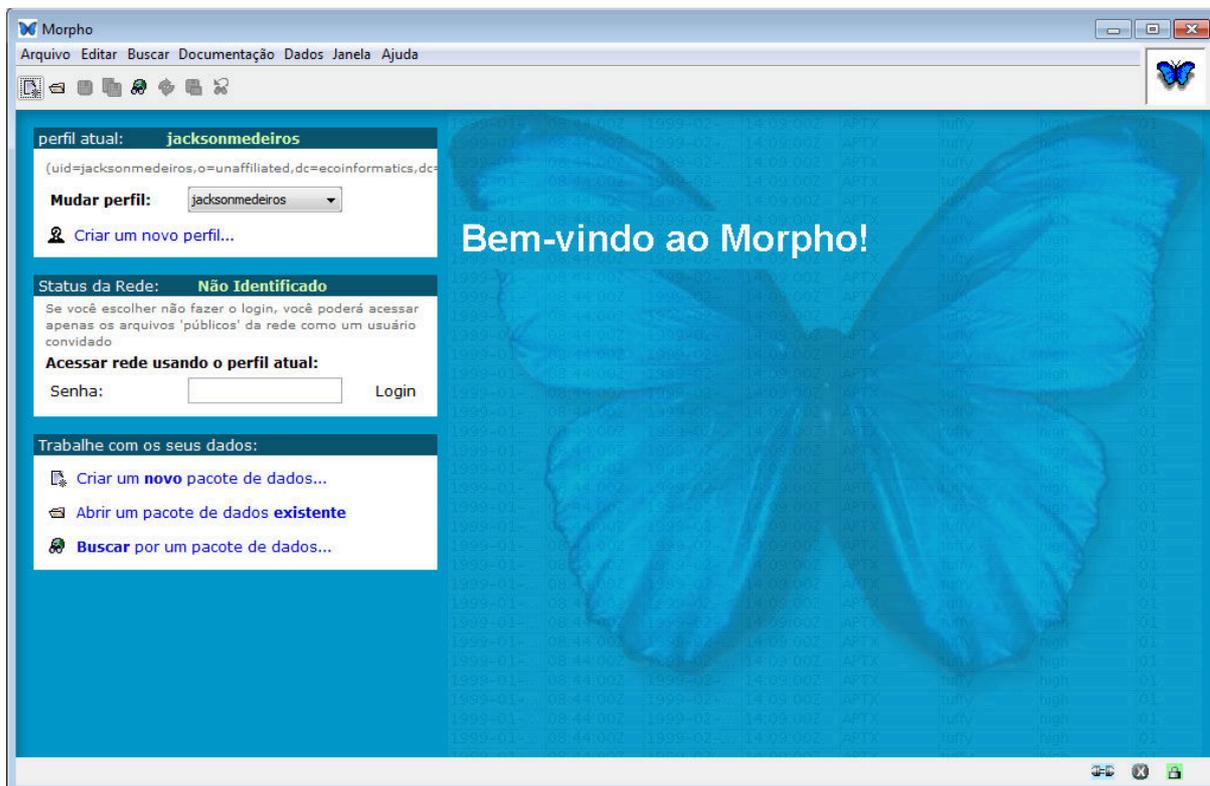


Figura 7: Tela inicial do Morpho

Fonte: Morpho

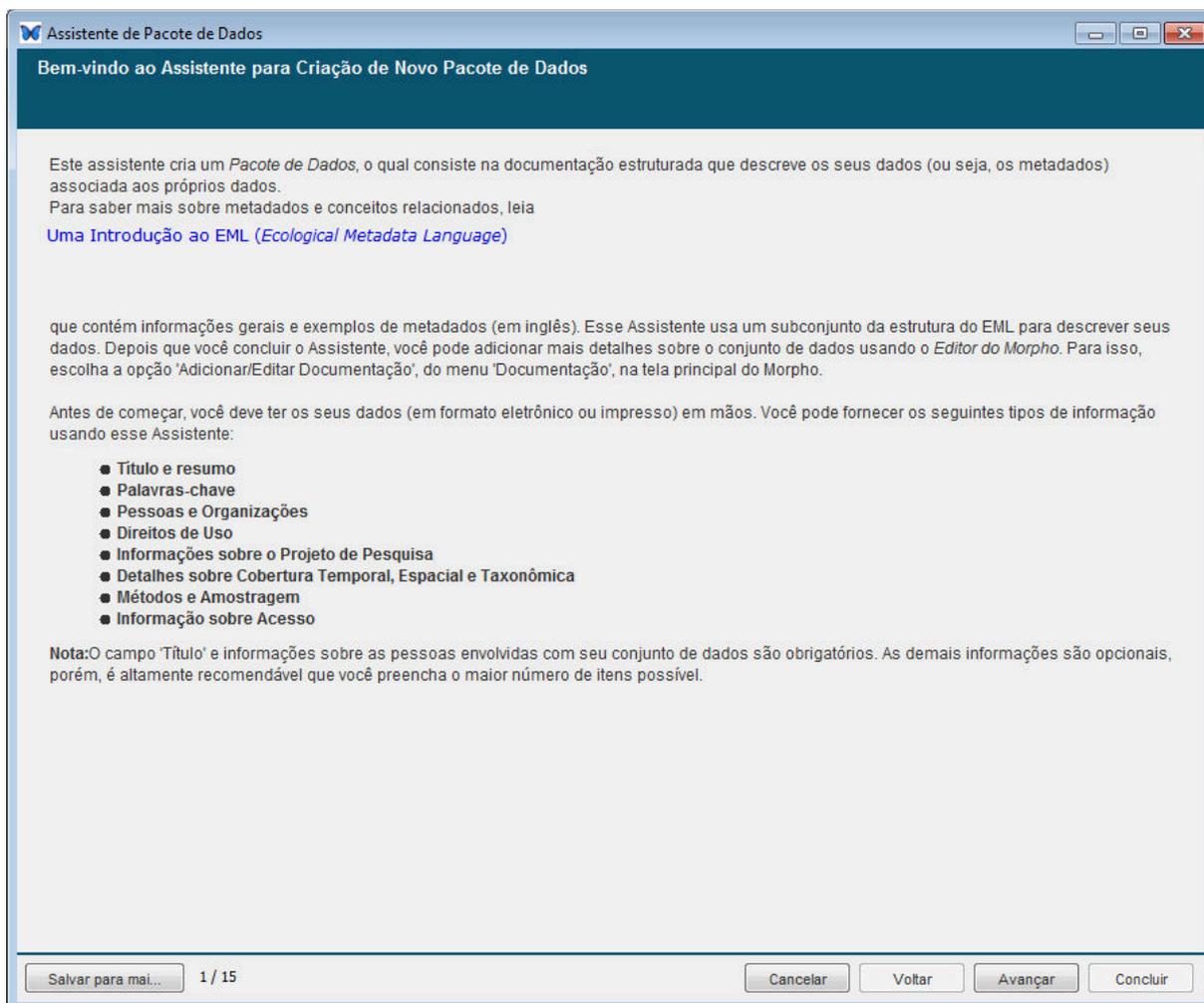


Figura 8: Tela de apresentação do Morpho – Assistente para criação de novo pacote de dados

Fonte: Morpho

A partir da descrição do repositório e sua ferramenta para envio de dados e metadados, passamos a relatar agora a questão do conteúdo do repositório, isto é, os dados científicos primários e como o repositório compreende e lida com esse objeto.

6.3 O entendimento sobre os (conjuntos de) dados científicos

Algo que deve ser levado em conta quando estamos analisando dados de pesquisa diz respeito à forma como esses se apresentam. Dados, como visto no capítulo 2, são sempre complexos, inclusive na sua definição. Este problema advém dos campos de pesquisa e das próprias pesquisas realizadas. Isso porque é constatado que a produção de dados de pesquisa, além de absorver processos oriundos do próprio campo, depende de metodologias e abordagens que são administradas pelo pesquisador, incluso aqui o campo da ecologia. Drucker (2011, p. 115) faz menção a isso:

Essas diferentes abordagens em ecologia, aliadas às tradições de pesquisa distintas tanto em suas subdisciplinas como em áreas afins, levam à produção de dados altamente heterogêneos. Tais dados podem ser representados por contagens de indivíduos, medidas de variáveis ambientais (por exemplo temperatura ou precipitação) ou representações de processos como competição ou herbivoria, ou de seus efeitos (por exemplo grau de abertura de dossel ou quantidade de dano em folhas). As terminologias utilizadas também variam de acordo com a linha de pesquisa, bem como a forma de estruturar os dados digitalmente. (DRUCKER, 2011, p. 115).

A questão definitiva do que vem a ser um dado, ou um conjunto deles, é um problema inicial que não passa pelo Repositório como um ponto formalmente colocado. Isto decorre da própria forma como o Programa busca gerir seus dados, já que se constitui em um sistema mais amplo do que o próprio Repositório, tendo em vista que mantém, fora do Repositório, coleções de taxonomias recebendo outro tipo de tratamento. Isso é consonante ao âmbito científico internacional na área de biodiversidade (E3).

Isso, na realidade, vai ao encontro dos problemas que o Repositório busca tomar conhecimento e tratar como sistema que agrupa conjunto de dados ecológicos que podem, ainda, ser arranjados com outros tipos de dados e/ou atributos, como medidas de espécies, tamanho de bico, asa, diâmetros de plantas, medidas ambientais, como abertura solar, fertilidade de solo etc. Isso faz com que se fuja de um padrão pré-estabelecido de base de dados, e se tenha que lidar com uma

permissividade muito grande de dados, como questões taxonômicas que classificam a biodiversidade.

Ou seja, o Repositório de dados do PPBio é um sistema de recebe dados coletados a partir de levantamentos da biodiversidade, bióticos ou abióticos, servindo, de modo geral, como premissa para aceitação de dados que compõem o agregado de (conjuntos de) dados. Em função da essência pela qual o Repositório foi criado, sua missão é o armazenamento, preservação e divulgação de dados coletados a partir de projetos cobertos pelo PPBio.

Durante a coleta de dados brutos praticamente não há normas ou padrões que sejam impostos. Dizemos “praticamente” uma vez que algumas chaves-primárias ou chaves-básicas são obrigatórias, satisfazendo um processo mínimo de manutenção colocado pelo Repositório. Outras colunas (atributos/características) a serem preenchidas são estabelecidas pela área de coleta – a área de plantas não pode ter o mesmo padrão para a coleta de dados sobre aves – ou pelo próprio pesquisador, já que, por exemplo, durante a coleta um animal ou planta pode ser “mapeado” e não há nenhuma coluna de parâmetro que atue nas novas características atribuídas a ele. Mesmo com a pouca quantidade de padrões a serem seguidos no momento da coleta, esse se torna um processo oneroso, fazendo com que possa existir uma possível lacuna entre o produzido e o disponibilizado.

Esse processo pode ser facilmente reproduzido: costumeiramente o investigador, ao coletar os dados para sua pesquisa, o faz com um tipo de tabela, como uma planilha, por exemplo. No entanto, sua criação é frequentemente acompanhada de uma série de anotações, como um bloco de notas. Assim, é comum que sejam estabelecidos procedimentos particulares para marcações e/ou valorações de determinados dados, fazendo com que a compreensão de um analista ou do próprio banco de dados/repositório seja prejudicada a partir da existência de uma célula em branco – que pode ser compreendida como valor nulo ou não coletado. Isso pode levar a erros não intencionais e faz com que haja necessidade de limpeza dos dados coletados com a criação de uma nova tabela, um processo trabalhoso e redundante (DRUCKER, 2011).

A Figura 9 e a Figura 10 mostram, respectivamente, um exemplo de uma planilha de levantamento ecológico sem anotações que permitam atribuir significado aos dados e um exemplo da planilha da Figura 9 preparada para importação em banco de dados ou pacote estatístico. Esse processo consiste em uma tabulação consistente dos dados após o processo de limpeza dos mesmos, buscando sua clareza para inserção no Repositório de dados.

Parcela	Espécie	Qde	Altitude
LO2 3500	<i>Rapatea paludosa</i>	133	55
	<i>Monotagma spicatum</i>	60	
	<i>Pepinia sprucei</i>	473	
	<i>Ischnosiphon arouma</i>	3	
LO4 0000	<i>Rapatea paludosa</i>	0	45
	<i>Monotagma spicatum</i>	19	
	<i>Pepinia sprucei</i>	120	
	<i>Ischnosiphon arouma</i>		

Figura 9: Exemplo de uma planilha de levantamento ecológico sem anotações que permitam atribuir significado aos dados

Fonte: DRUCKER (2011, p. 116)

Data	Parcela	Espécie	Contagem	Altitude (m)
26/04/2004	LO2 3500	<i>Rapatea paludosa</i>	133	55
27/04/2004	LO2 3500	<i>Monotagma spicatum</i>	60	55
28/04/2004	LO2 3500	<i>Pepinia sprucei</i>	473	55
29/04/2004	LO2 3500	<i>Ischnosiphon arouma</i>	3	55
30/04/2004	LO4 0000	<i>Rapatea paludosa</i>	0	45
01/05/2004	LO4 0000	<i>Monotagma spicatum</i>	19	45
02/05/2004	LO4 0000	<i>Pepinia sprucei</i>	120	45
03/05/2004	LO4 0000	<i>Ischnosiphon arouma</i>	*	45

Figura 10: Exemplo da planilha da Figura 9 preparada para importação em banco de dados ou pacote estatístico

Fonte: DRUCKER (2011, p. 116)

Os dados são vistos com a finalidade de reproduzir algo que já foi feito, visando sua reutilização. Isso significa a possibilidade de qualquer investigador

refazer uma pesquisa que tenha uma questão abrangida pelos dados disponibilizados, o que se baseia em uma lógica central da ciência, a possibilidade de reproduzir, questionar, refutar e avançar o conhecimento. No entanto, há de levarmos em conta que apenas os dados por si mesmo não permitem reprodução total das pesquisas realizadas, uma vez que a falta de elementos importantes da pesquisa podem dificultar esse processo. Para que efetivamente exista reúso, outros elementos devem ser disponibilizados pelos pesquisadores.

Reúso é a utilização para outros propósitos, como ampliar escalas de análise. Além disso, é possível que análises conjuntas sejam realizadas com base em estudos separados, o que possibilita ampliar o que já havia sido encontrado, especialmente em sistemas naturais, onde mesmo em locais muito próximos, a extrapolação de análises pode não refletir a realidade (E1).

Essas análises em conjunto possibilitam, entre outras funções, que sejam apresentadas e estabelecidas políticas públicas ou privadas para conservação, código florestal etc., uma vez que trazem dados concretos e correspondem àquilo que acontece em determinado local. No caso dos dados ecológicos, mostram, por exemplo, a real importância da vegetação das margens de rios, distância de igarapés, lençol freático, o que isso afeta a distribuição de espécies e as consequências da sua conversão em área rural, entre outros.

No contexto do PPBio, essa é a lógica que guia o sistema de gerenciamento de dados proposto: ampliação de análises, levando a resultados que possam agregar a compreensão de um sistema amazônico gigantesco e que auxilie no direcionamento de políticas no sentido de maior conhecimento do ambiente e, em consequência, na preservação ambiental.

Como uma das preocupações do Programa é possibilitar o acesso da sociedade aos dados coletados, “a possibilidade de analisar os dados integrados através de escalas temporais e regionais levou os pesquisadores a aumentar a gama de questões respondíveis e incentivar mais pesquisadores para participar⁶¹” (PEZZINI et al., 2012, p. 268). Assim, todo esforço colocado em um posicionamento

⁶¹ No original: “The possibility to analyze data integrated across temporal and regional scales has led researchers to increase the range of answerable questions and encourage more researchers to participate”.

de livre acesso aos dados é também uma atitude no sentido de provocar maiores níveis de colaboração para reutilização.

A disponibilização dos dados é um momento crucial para a efetivação científica, mas, em alguns casos, ainda pode sofrer certa resistência por parte dos pesquisadores, até mesmo por temor do destino dos dados. Mesmo entendendo que

a partir do momento que você se dispõe a colocar uma coisa na internet, infelizmente está sujeito a todo tipo de coisa. Mas a partir do momento em que a pessoa também está colocando ali, está pensando que quer ser útil, quer que aquilo seja utilizado por outras pessoas. (E2).

Há, nesse sentido, uma quebra do possível receio de disponibilização dos dados coletados e utilizados em trabalhos como teses e dissertações, passando a ser visto como um processo de incentivo à reutilização, não havendo retrabalho de campo e, por consequência, economia de dinheiro, em geral, público, bem promotor de avanços científicos significantes.

Com relação à reutilização dos dados, também existe junto ao Repositório um versionamento – não explícito – que é realizado. “Não explícito” porque pode ser entendido como a possibilidade de recuperação dos dados antigos, mas essa não é uma opção visível para o usuário, já que não existem links disponíveis para acesso via interface. O conjunto de dados está armazenado no servidor, possibilitando que, ao tomar conhecimento do nome do arquivo que armazena o conjunto de dados antigo, o usuário tenha acesso a ele.

O versionamento, ou melhor, a criação de novas versões dos conjuntos de dados a partir de modificações que o pesquisador que os coletou encaminha para o Repositório, pode ser compreendida pela ótica de quem coleta e pela ótica tecnológica. Pela visão do primeiro, o versionamento resulta de um entendimento tácito do que é uma versão de conjunto de dados, não existindo diretrizes declaradas nesse sentido, sendo que as novas versões podem ser compostas, por exemplo, por conjuntos de dados que não produzem grandes alterações no que foi previamente armazenado no Repositório. Isso, inclusive, faz parte de indicações

realizadas pelos gestores àqueles que querem disponibilizar seus dados no repositório, atuando como um tipo de acordo entre as partes (E3).

Nesse sentido, o versionamento decorre de uma função própria da ciência, a de responder uma pergunta previamente estabelecida a partir de uma metodologia instituída para esse fim. Ou seja, parte da ideia de que a coleta de dados visa responder uma questão que está colocada como problema. Caso essa questão permaneça a mesma, bem como a metodologia utilizada, há versionamento dos dados. Caso contrário, um novo conjunto de dados pode ser criado.

De razão tecnológica, os dados recebem uma nova versão sempre que são reenviados para o curador e, posteriormente, para o Repositório propriamente dito, não havendo influência humana no processo de atribuição das versões. O salvamento de novas versões é automático a partir do software utilizado pelo Repositório. Caso o investigador requeira alguma alteração, ao salvar os dados, o sistema automaticamente altera o versionamento, alterando, por exemplo, de versão x0.1 para x0.2 e assim por diante em quantas alterações forem necessárias ou requeridas. É importante frisar que, embora armazene todas as versões dos dados, o sistema do repositório disponibiliza, através de sua interface, apenas a última versão.

Compreender como os dados científicos primários são vistos pelo Programa e seu Repositório, posiciona os conjuntos de dados como actantes, atores-rede que mobilizam diversos outros atores-rede, configurando-se na produção mais crua do processo de autoria. No entanto, os dados devem passar por um processo de gestão/curadoria bem como haver preocupação com sua preservação.

6.4 Dos (conjuntos de) dados aos metadados: curadoria e preservação

Os dados podem ser encontrados em circunstâncias cotidianas. No entanto, o conjunto de dados que se enquadra em nosso plano de estudo é oriundo de regiões de análise ecológica. Neste contexto, podem ser encontrados dados novos, como

novas espécies ou outra forma de analisar as já existentes, além do confronto com dados que já estão ou que estão sendo coletados.

O depósito de dados científicos ocorre da seguinte maneira: o investigador que coletou os dados entra em contato com o responsável pela curadoria/gestão do Repositório, enviando, a partir de um modelo pré-estabelecido, os metadados sobre aquilo que foi coletado. Há, então, o processo de conferência de aspectos mínimos para uma documentação e, neste estágio, especial atenção é dada à metodologia descrita nos metadados, já que ali consta, em detalhes, como os dados foram coletados.

Nesse ponto existe um processo de negociação entre curador – indivíduo responsável por critérios de qualidade e clareza dos dados, bem como o recebimento de dados e de metadados e exigência do cumprimento do tempo limite para depósito – (repositório) e investigador (depositante). Estando a documentação descritiva (metadados) de acordo com o que é estabelecido pelo Repositório, ocorre o imediato depósito. Caso contrário, iniciam-se as conversas para que ajustes possam ser realizados visando adequação da tabulação feita pelo pesquisador de campo. Esse processo é realizado até ser possível a formação consistente e clara de metadados de acordo com as exigências do Repositório.

É importante ressaltar dois eventos que ocorrem costumeiramente: o primeiro se refere ao fato de que o investigador que coleta e tabula dados não os envia diretamente à base, não inserindo qualquer tipo de dado ou metadado no sistema. Todo o processo é acompanhado por um gestor/curador dedicado à tarefa de conferência qualitativa, no sentido de verificação de parâmetros já estabelecidos, visando redução de possíveis problemas em relação aos dados e sua documentação, resultantes, também, da complexidade que age sobre os estudos ecológicos e da abertura que o Repositório permite para depósito de qualquer tipo de dado.

O segundo evento se refere à diferença entre coleta e envio de dados. Muitas vezes, pesquisadores que coletam dados estão envolvidos, por exemplo, na confecção de dissertações e teses, o que faz com que eles possam compreender que os dados úteis da pesquisa são apenas aqueles que foram utilizados na fundamentação do trabalho. Caso isso ocorra, há um movimento por parte do

curador em informar que todos os dados, mesmo aqueles não utilizados para a concretização do trabalho, também são considerados dados de pesquisa e, mesmo não sendo úteis para o investigador, podem ser constituir de material valioso para outros.

Mesmo informações que parecem muito bobas para a pessoa, e as vezes a pessoa nem anota, podem ser importantes para outras coisas. [...] Em que hora você começou a coletar? Qual é a hora que você terminou? É muito importante [...] e pode refletir em outra informação. (E3).

Tratar a documentação que acompanha os dados é outro aspecto de imprescindível importância. Essa etapa deve responder a algumas questões, a saber: como, quando, onde e por quem os dados foram coletados? Com quais objetivos? Estas indagações são essenciais para que possa ser realizado o reuso dos dados, ou seja, para que possam ser utilizados em novas análises (DRUCKER, 2011).

Esse procedimento de documentação amplia a utilidade e colabora para a longevidade do conjunto de dados, pois os próprios pesquisadores que levantaram os dados tendem, com o passar do tempo, a esquecer detalhes dos procedimentos de coleta relevantes para sua interpretação, ou mesmo as abreviações usadas ao anotá-los. (DRUCKER, 2011, p. 117).

É nesse sentido que é essencial que se descreva a tabela de atributos, introduzindo o nome da variável, sua unidade, como foi obtida etc. A documentação é importante para compreender o estudo como um todo. Sem a descrição, provavelmente, não há como lidar com o estudo já realizado. Caso não houvesse uma iniciativa organizada a partir do Repositório, seguramente o processo se tornaria inviável pela falta de critérios mínimos de tabulação (E1; E3).

Como relatado anteriormente, o mesmo é observado em relação aos dados propriamente ditos. Neste caso, há exigência de contemplação do que se convencionou chamar de chaves-primárias, isto é, um conjunto de elementos que

são considerados essenciais para a coerência dos dados – atributos relativos à coleta de dados: sítio, parcela⁶², trilha etc. Além do mais, há supervisão para conferência de possíveis erros de digitação, formato correto para o Repositório – já que apenas são aceitos dados brutos e não tabelas para análise.

Essa figura central no processo de checagem do conjunto de dados é a do curador. Sua função é, basicamente, conferir a consistência da descrição dos dados, os metadados, verificando se há atendimento satisfatório ao preenchimento de campos minimamente necessários. Esta preocupação ocorre por dois motivos fundamentais: em primeiro lugar, pela necessidade de credibilidade que deve haver na descrição dos dados, permitindo ao Repositório manter uma lógica de construção de suas tabelas que promovam a difusão do conhecimento ali expresso em forma de dados. Em segundo lugar, porque o investigador de campo, dentro de suas atribuições, não necessariamente conhece ou está em sintonia com o processo final, podendo, mesmo ao saber que é necessária uma boa descrição para uso futuro, incorrer em problemas de clareza e consistência da descrição (E2; E3).

Outro ponto notável sobre a curadoria é a capacitação, em forma de cursos, sobre a documentação de dados, fazendo com que sejam apresentados modos consistentes de preenchimento de campos de metadados, entre outras questões que visam o reuso dos dados. Estabelecer o reuso depende, quase exclusivamente, dos metadados que acompanham os dados científicos. Assim sendo, é possível dizer que caso os metadados não acompanhem os dados científicos, os dados isolados não podem fazer parte do Repositório, já que não apresentam valor de documentação/descrição.

Ou seja, a existência de tabelas de dados por si mesmas não caracteriza um possível conjunto de dados a ser disponibilizado pelo Repositório. Exemplificando: uma tabela, ainda que bem organizada, deve demonstrar as unidades e/ou códigos que cada coluna representa, resumo, cobertura temporal, geográfica e/ou taxonômica, palavras-chave etc., servindo como um caminho a ser seguido para documentar, isto é, descrever de forma detalhada cada elemento componente dos dados enviados. Inconsistências podem ocorrer, mas existem alguns itens que

⁶² Menor unidade a ser coletada, como um terreno delimitado, demarcado, por exemplo.

podem colaborar com a clareza: NULL para vazio, NC para não coletado e NA para não amostrado (E1).

Sem a documentação dos dados, os dados se tornam inúteis. [...] Numa tabela de dados, geralmente tem assim, exemplo: 'tamanho'. [...] tamanho do quê? Em que unidade? [...] a documentação dos dados vai estar vindo para facilitar a interpretar aquela tabela de dados, saber o que significa aquela coisa, em que unidade está cada coisa, assim como os dados foram coletados. Então é uma maneira de você garantir que outras pessoas também possam realizar essa mesma coleta em outros locais. Ou seja, tem que ser replicado. Também para poder saber detalhes de como onde foi coletado, como foi coletado, quando foi coletado. [...] outras questões como palavras-chave e título é para facilitar a busca. (E2).

Os metadados também agem como um importante componente também no que diz respeito à longevidade dos dados. Neste sentido, há duas perspectivas de trabalho com preservação dos dados: primeiro é a preservação para recuperação após certo período de tempo através da descrição. Em segundo lugar, a preservação física dos dados, já que existem medidas comprometidas com a guarda segura através de redundância e replicação, respectivamente, processo de manter um sistema sempre disponível através de repetição de dados em bases alternativas à principal e backup de dados entre servidores em locais distintos.

No INPA existem, atualmente, três servidores capazes de executar de forma redundante o Repositório de dados do PPBio. Também há replicação dos dados através de parceria com o LNCC⁶³ (Laboratório Nacional de Computação Científica), órgão vinculado ao MCTI. É utilizada também a rede que mantém o Metacat – KNB⁶⁴ (Knowledge Network for Biocomplexity) e NCEAS⁶⁵ (National Center for Ecological Analysis and Synthesis) – com futura entrada no DataONE⁶⁶ (Data Observation Network for Earth), atuando com replicação automática uma vez que os dados sejam colocados no repositório do PPBio (E3).

⁶³ <http://www.lncc.br>

⁶⁴ <https://knb.ecoinformatics.org/>

⁶⁵ <https://www.nceas.ucsb.edu>

⁶⁶ <https://www.dataone.org>

Ao perceber os (conjuntos de) dados e seus metadados como forma de elementos interligados, há necessidade de depreender que o processo de curadoria e gestão dos dados científicos demonstra duas preocupações latentes: (a) manter o serviço a longo prazo, sustentar o Repositório de forma a possibilitar a recuperação dos dados depositados para avanço científico e (b) utilizar serviços de replicação em redes nacionais e internacionais, o que, além de visar segurança na manutenção de dados, fomenta as relações para parcerias científicas e visibilidade do Repositório e do Programa.

6.4 O gerenciamento político: a política de dados

O Repositório de dados do PPBio nasce a partir de ideias para divulgação, de forma aberta, dos dados coletados por pesquisadores que se utilizam da infraestrutura de coleta de dados, como sítios e parcelas permanentes. Ou seja, a questão inicial, no início da implementação do Repositório, era um local de depósito de dados que pudesse suprir uma necessidade imediata. Partiu-se da ideia de que havia um problema a ser resolvido, então uma solução foi buscada. Aprimoramentos tecnológicos e políticos são ajustes posteriores.

Apesar disso, como relatam os entrevistados, foi perceptível a necessidade de haver uma política de dados explicitamente estabelecida. O início do Programa, em 2005, deu origem à Política de dados do Programa de Pesquisa em Biodiversidade em 2009 (BRASIL, 2009), como consta no Anexo A⁶⁷. Esse processo, como toda política a ser definida em uma instituição, dependeu de uma série de fatores, uma vez que implica no entendimento da missão, neste caso do Repositório e do próprio PPBio, bem como aspectos que concernem ao MCTI, INPA e outras entidades. Além do mais, o empreendimento de construção de uma determinada política deve contar com a colaboração de investigadores de campo através de suas opiniões.

E assim foi feito. Pesquisadores de campo e/ou interessados na investigação ecológica com coleta de dados poderiam apresentar seus interesses e/ou anseios a uma política de armazenamento e disponibilização de dados. Solicitações anotadas,

⁶⁷ Também disponível em: http://ppbio.inpa.gov.br/sites/default/files/politica_dou.pdf

o resultado foi um cartapácio que detalhava de forma exaustiva como os dados deveriam ser dispostos no repositório, com, por exemplo, identificadores únicos⁶⁸, possibilitando identificar autoria, sítio de coleta etc. Após determinada redução no calhamaço gerado inicialmente, e entre ideias e possibilidades, consolidou-se a existência de uma política de dados como é apresentada no Diário Oficial da União (DOU), capaz de subsidiar pontos básicos para disponibilização de dados (E1).

Assim sendo, a política de dados do PPBio pode ser encarada como uma consolidação de um (longo) processo – tanto de entendimento de anseios quanto de obrigações –, possibilitando que novas parcerias fossem criadas e/ou facilitadas. Além do mais, faz compreender que a disponibilização de dados em um repositório não se constitui como uma ameaça para o investigador que coleta dados ecológicos, mas, sim, de uma vantagem que é capaz de atrair benefícios para o pesquisador, para a própria pesquisa e também para o bem-estar social, já que novos conhecimentos podem ser extraídos daquilo que está devidamente armazenado, formando-se um processo cultural baseado na política de dados estabelecida.

A publicação de uma política de dados desenvolvida no âmbito do PPBio tem por propósito, o qual está ligado aos objetivos do Programa, o “fomento à geração e disseminação de informações e conhecimento sobre a biodiversidade brasileira para diferentes segmentos da sociedade” (BRASIL, 2009, p. 8). Visa promover o gerenciamento de diretrizes para os participantes do Programa, no que tange aspectos ligados à geração e uso dos dados, as regras para os participantes do Programa para coleta, armazenamento, propriedade, autoria, compartilhamento, citação, acesso e uso dos dados e das bases de dados, buscando evitar problemas de acesso, uso e disseminação (BRASIL, 2009, p. 8).

O anexo da política de dados do PPBio é, por assim dizer, a política propriamente dita, contando com as seguintes seções:

- (a) Definições gerais, onde são explicadas as competências de termos utilizados no Repositório e na política de dados, como metadados, dados, dados preliminares, sensíveis, consolidados, ostensivos, núcleos de biogeoinformática, comitê gestor da informação, participantes do Programa,

⁶⁸ O identificador permanece, sendo apresentado sob a metatag Identifier. Isso caracteriza, como o próprio nome leva a crer, a identificação única daquele conjunto de dados.

- núcleos executores e regionais, indicador de confiabilidade do dado, divulgação ampla e restrita, além do período de embargo⁶⁹;
- (b) Das obrigações, atribuições e competências, dizendo respeito aos encargos dos núcleos de biogeoinformática, do comitê gestor da informação, do conselho diretor, do comitê científico e dos participantes;
 - (c) Das bases de dados – usos e acessos, respondendo sobre as condições de uso e acesso das bases de dados;
 - (d) Da propriedade intelectual, que trata sobre proteção e propriedade intelectual;
 - (e) Dos prazos, referente ao período para envio de dados e metadados (BRASIL, 2009).

Na seção “Sobre as obrigações, atribuições e competências” que cabem aos Núcleos de Biogeoinformática, consta: “Manter registro das publicações que utilizaram dados do PPBio” (BRASIL, 2009, p. 8). Cumpre notar que isso se constitui em um trabalho inglório para o Repositório, sendo, por política definida, aquele que deve buscar qualquer registro de publicações realizadas com base nos dados científicos disponibilizados pelo próprio Repositório, bem como a manutenção de ferramentas computacionais que forneçam infraestrutura para o pleno funcionamento. Este trabalho é uma atribuição do curador e/ou gestor.

Essa convenção é retomada a partir do termo de compromisso (Anexo B⁷⁰) assinado pelos investigadores que coletam dados e o Programa, tendo sua validade a partir do que foi coletado dentro dos sítios delimitados. No entanto, está também ligada a um processo de convencimento e articulação entre os indivíduos para que seja cumprido, uma vez que se dispõe de tempo limite para que isso ocorra.

No que se refere à seção “Da gestão e autoria de dados”, poucas informações são explicitadas, mas permitem compreender que as coletas realizadas com recursos do PPBio “são de interesse público para o desenvolvimento científico – tecnológico” (BRASIL, 2009, p. 9), sendo a participação do PPBio restrita à função de resguardar a autoria dos dados nas bases de dados e nas publicações resultantes (BRASIL, 2009).

⁶⁹ Parte desses termos estão descritos durante o capítulo. As definições a partir do Diário Oficial da União estão no Anexo A.

⁷⁰ Também disponível em:

<http://ppbio.inpa.gov.br/sites/default/files/TERMO%20DE%20COMPROMISSO%20e%20Normas%20de%20conduta.pdf>

Essa função de resguardo está vinculada à ideia de que os dados após coletados e disponibilizados no Repositório devem conter como autor uma entidade humana ou institucional. Assim, para a publicação de documentos – em geral artigos científicos – que são resultado de análise dos dados, o analisador (como indivíduo e/ou entidade) é o autor/proprietário do documento textual criado, mas o crédito referente aos dados científicos utilizados dizem respeito ao indivíduo e/ou a entidade que mantêm uma situação de autor/proprietário sobre o que foi coletado.

O processo de coautoria em trabalhos resultantes de análise dos dados segue um regime mais propositivo ao mencionar um sistema de diálogo com convite e permissão para uso dos dados:

[...] a co-autoria ou outras formas de citação da participação na geração, análise e publicação dos dados deverão ser definidas pelas partes envolvidas, refletindo a participação intelectual, de acordo com o código de ética da ciência (BRASIL, 2009, p. 9).

Esse segmento, embora abstrato, refere-se às atitudes relacionadas às questões de produção intelectual que devem ser tomadas “de acordo com o código de ética da ciência”. Embora tácito, o referido código estabelece de forma implícita aos iniciados na Ciência, um processo de dar créditos àqueles que os merecem. Ou seja, são os novos analisadores/re(utilizadores) dos dados os responsáveis por creditar o que é devido à quem os coleta e os disponibiliza, em geral, através de citações e referências.

Na seção Data Set Usage Rights nos metadados da interface de acesso e na seção 4 (Sobre as condições de uso e acesso das bases de dados) item 4.2 do DOU, encontramos os seguintes parágrafos sobre a disponibilização de dados e possível interesse em reutilização:

Todos os dados do PPBio serão publicados no máximo 1 ano após sua coleta, desde que creditados os responsáveis pela mesma e disponibilização dos dados em qualquer publicação que os utilizem. Recomendamos aos interessados em utilizar os dados entrar em

contato com os responsáveis e assim discutir o interesse de uso e possibilidade de participação no corpo de autores. (PPBIO, 2015)

[...] O acesso e uso com intenção comercial ou de forma que possa resultar na geração de produtos ou processos passíveis de exploração econômica, deverá ocorrer mediante a celebração de contrato entre as partes interessadas. (BRASIL, 2009, p. 9).

É possível perceber que a política de disponibilização de dados pelo Programa está ligada, no que se refere à autoria, apenas aos trabalhos oriundos da análise de dados, como artigos e livros. A questão autoral dos dados propriamente ditos é apenas explicitada no termo de compromisso para coleta de dados em sítios do PPBio (Anexo B).

(6) A política de dados não estabelece nenhuma regra sobre autoria dos trabalhos científicos resultantes dos dados coletados, ficando a definição de autores e co-autores a critério do pesquisador. No entanto, as publicações científicas e qualquer outro meio de divulgação de trabalho de pesquisa, apoiados pelo Edital _____, deverão citar, obrigatoriamente, o apoio prestado pelos financiadores _____, utilizando suas respectivas logomarcas, de acordo com as normas de Uso da Marca da cada instituição. (PPBIO, [2011]).

Neste sentido, a Política de dados atua como algo

[...] bem simples: a gente só pede pra que a pessoa que queira utilizar os dados que não se aproprie dos dados de forma indevida. Que cite a fonte [...] a gente sugere que entre em contato com a pessoa que seria o proprietário dos dados pra [...] conversar de fazer [...] coautoria, se a pessoa quer participar, se tem problema, o que acha de fazer tal trabalho... Porque a grande ideia do repositório, fora essa questão de disponibilizar o acesso aos dados, também é você promover a integração entre as diferentes instituições de pesquisa e áreas de pesquisa. (E2).

Toda a disponibilização de dados coletados explorada pela Política de dados está colocada de forma a colaborar com as agências de fomento na possibilidade de acesso aos dados primários de pesquisas realizadas com verbas públicas. Ou seja, os dados são considerados públicos, já que o financiamento também é público. Isso é visto, além de uma comprovação da utilização do financiamento empregado, como uma prestação de contas do que foi feito a partir do que foi investido com capital oriundo da sociedade (E1; E2). Mas a quem deve ser creditado? O que o Repositório entende como um proprietário/autor?

6.6 A autoria

Ao iniciar a compreensão do quesito autoria, é bom ter em mente que as pessoas envolvidas no gerenciamento do Repositório têm visão da importância do processo de atribuição de autoria e seu valor no processo científico: “uma coisa importante para incentivar as pessoas a disponibilizar é dar crédito a elas” (E3). Também é relevante destacar que, como aparece anteriormente nas imagens relativas ao software utilizado para envio e gerenciamento de dados ao repositório – Morpho e Metacat –, a autoria é atribuída através da expressão “Proprietário”.

A expressão “Proprietário”, no entanto, procede da tradução feita a partir dos softwares utilizados na administração do envio e disponibilização de dados e metadados, o que causa, em algumas circunstâncias, certo desconforto na utilização do termo. Isso ocorre uma vez que ele atribui um sentido de posse, ser dono de algo, tornando-se sobrecarregado se levarmos em consideração a semântica do termo. Sendo assim, devemos tratar essa questão com certo cuidado com a finalidade de não atribuir sentido de restrição a partir do poder existente por uma entidade, seja ela individual ou institucional (E1; E2; E3).

Durante o processo de inserção de dados no Repositório, o pesquisador envia os dados e os metadados coletados para um curador de dados. No momento do envio, existem metadados responsáveis pela designação de proprietário e partes associadas. Para o Repositório é isso que define a autoria do conjunto de dados e o curador e/ou o Repositório não tem nenhuma participação na atribuição de autoria. De outro modo: “o proprietário é como se fosse o autor principal”, onde

“normalmente a gente põe ele em ordem crescente de importância de quem coletou o dado. Então o primeiro proprietário fica como se fosse o primeiro autor, o segundo proprietário como segundo autor e assim por diante” (E2).

No que se refere as partes associadas, estão dispostas “pessoas que de alguma forma colaboraram com o trabalho”, podendo compreender “um identificador de espécies, um ajudante de campo, alguém que, de repente, digitou seu material”, isto é, “qualquer pessoa que de alguma forma tenha contribuído e que você ache que seja importante citar você coloca como partes associadas”⁷¹ (E2). Isso pode ser visto nas Figuras 11 e 12.

⁷¹ Vale ressaltar que o processo de autoria está construído como as ações do pesquisador para a construção desse ente (conjunto de dados). Cumpre lembrar que o PPBio, como repositório, apresenta cursos de capacitação para os investigadores interessados, assumindo um tipo de responsabilidade para mostrar o melhor caminho para se atingir consistência na tabulação, envio e manutenção dos dados.

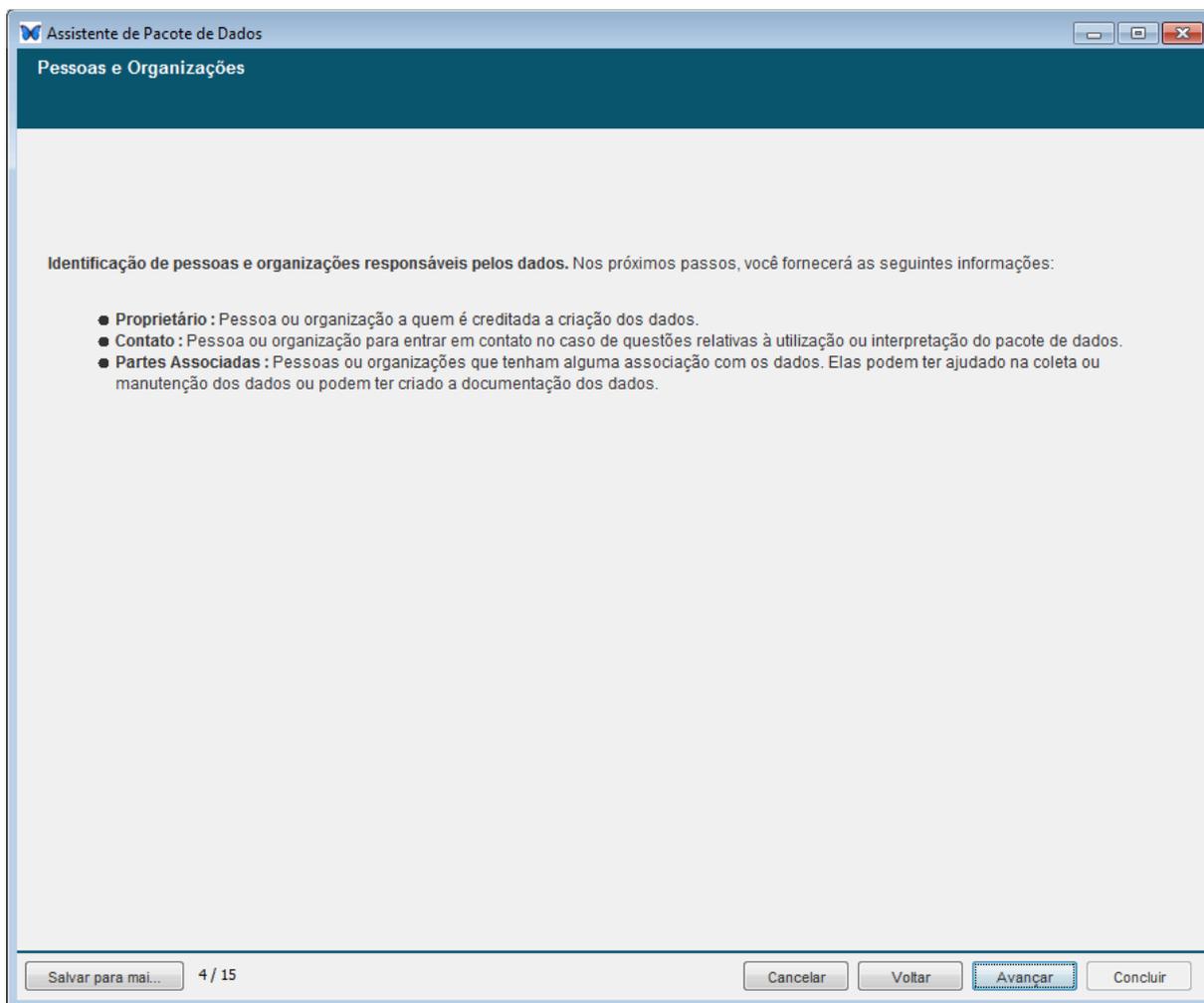


Figura 11 – Pessoas e organizações

Fonte: Morpho

Proprietário Detalhes

Você pode escolher uma das entradas adicionadas anteriormente. - ...

Saudação:

Nome: Traduções

Sobrenome: Traduções

Organização: Traduções

Nome da Posição:

Endereço 1:

Endereço 2:

Cidade: Estado:

Caixa Postal: País:

Telefone: Fax:

E-mail: URL:

Um dos três exigidos

OK Cancelar

Figura 12 – Proprietário detalhes

Fonte: Morpho

No entanto, em caso de citação do conjunto de dados analisado ou em análise, como em outras pesquisas ou revalidações, as partes que devem ser citadas dizem respeito apenas aos chamados proprietários, excluindo-se os indivíduos colocados como partes associadas. Dessa forma, o processo pelo qual está se planejando a inserção de dados no Repositório passa, obrigatoriamente, pela reflexão do pesquisador daquele(s) que está(ão) sendo atribuído(s) como proprietário(s) principal(is) do conjunto de dados, ainda que não haja um processo de licenciamento formal dos dados.

Isso reflete na compreensão do sistema de autoria e licenciamento. Havendo um sistema de reconhecimento do pesquisador e, embora a propriedade dos

conjuntos de dados seja vista como um bem público a que todos aqueles que têm a iniciativa de produzir novos conhecimentos possam buscá-los e utilizá-los, é necessário que se caracterize um investigador como o indivíduo responsável pela coleta, análise e disponibilização de dados, atuando como autor nesse processo, sendo afetado pelo bônus e pelo ônus da autoria

Esse processo de reconhecimento, oriundo da própria ciência, é de grande importância para quem está envolvido na atividade de coleta – e por isso deve ter seu trabalho certificado e condecorado – mas também para a continuação dos estudos impostos pelos dados. Fazer do investigador um partícipe reconhecido faz com que, além do incentivo em colaborar cada vez mais efetivamente na coleta e disponibilização dos conjuntos de dados, crie um senso de responsabilidade com os próprios dados, sua clareza, consistência etc.

Isso pode ser ilustrado da seguinte forma: um investigador que vai até determinado sítio ecológico e coleta os dados de uma espécie desconhecida de ofídio, o faz a partir de suas anotações que descrevem aquele réptil. Ao enviar o conjunto de dados e seus metadados para o Repositório, a ideia de quem é o autor, para o pesquisador, está encerrada, uma vez que ele ou seu grupo são os responsáveis pelo colhimento daqueles dados, sendo os observadores e “criadores”.

Se seu destacado esforço for retribuído com o reconhecimento de quem utilizar os dados no futuro, comumente através de citações, seu processo de investigação será recompensado, fomentando o compartilhamento de novos dados ou aprimorando os que já foram enviados e criando novas versões, bem como sabendo que os atributos que existiram na coleta, como clareza na descrição, consistência, entre outros, devem ser mantidos para que haja novos reúsos e novos reconhecimentos.

Fato que deve ser considerado primordial, e é constantemente lembrado por todos entrevistados, é um tipo de diretriz que começa a se popularizar a partir das agências de fomento no que se refere à obrigatoriedade de disponibilização dos dados de pesquisa, onde os dados disponibilizados são considerados produtos da pesquisa realizada com verbas da financiadora. Neste sentido, o ente público é o financiador de uma pesquisa na qual se espera algum tipo de retorno com expectativa de que haja uma retroalimentação em benefícios sociais daquele que

pagou para que conhecimentos – em forma de dados – fossem gerados e compartilhados.

7 REUNINDO A RAZÃO: ELEMENTOS EM AÇÃO

“O mundo comum ainda precisa ser coletado e composto”.
(LATOURE, 2012, p. 173).

Do mesmo modo que Latour provoca em *A vida de laboratório: a produção dos fatos científicos* – “o TRF⁷² é *totalmente* uma construção social” (1997, p. 160) –, caracterizamos a autoria, parafraseando o autor, como uma construção social, sendo razoável lembrar aqui suas palavras: “seria supérfluo opor os aspectos conceituais aos aspectos materiais da atividade de um laboratório” (LATOURE, 1997, p. 63). As associações foram explicitadas, esperamos agora ligar pontos que enxergamos no movimento da rede.

Ao reconhecermos a complexidade do objeto de estudo, buscamos nesta seção uma espécie de reunião dos elementos que formam a razão da autoria de dados científicos, os quais compreendem aspectos simétricos dos atores, humanos e não humanos, possibilitando a compreensão das associações entre elementos heterogêneos. Também gostaríamos de deixar claro que este capítulo não trata de uma interpretação do que foi narrado com ou sem a inclusão de “categorias mágicas” que, supostamente, explicam o que foi descrito, uma vez que isso seria um contrassenso em relação à Teoria Ator-Rede.

Assim, a seção busca coligar “um tipo de conexão entre coisas que não são, em si mesmas, sociais” (LATOURE, 2012, p. 23), onde a reunião desses elementos visa agrupar as matérias, aplicações, ferramentas, conceitos, práticas etc. heterogêneas, já que o social é “um movimento peculiar de reassociação e reagregação” (LATOURE, 2012, p. 25) constante, permitindo usufruir do que há de social em cada ator-rede.

Reiteramos que a caixa-preta da autoria é, na atualidade, vagamente compreendida. A elucidação dos movimentos do ator-rede autoria pode ser vista como gestos que abrem a caixa que estava fechada, desconstruindo algo antes

⁷² Thyrotropin-Releasing Factor, em português Hormônio liberador de tireotrofina, é um hormônio que estimula a hipófise, sendo um dos alvos do estudo realizado por Bruno Latour no Laboratório Salk, nos Estados Unidos.

vedado ou finalizado. Abrir a caixa-preta da autoria, por essência um conceito operacional, é descerrar uma série de outros conceitos, concepções, ideias, noções que atuam de forma contínua e em rede, como atores-rede, podendo ser considerados operadores para a ação. Trazer à tona esses conceitos já parece um trabalho e tanto – como narrado no capítulo 6 –, mas o que devemos pensar é na sua correlação e nos modos como eles integram um dispositivo altamente mutável e naturalmente instável.

Desse modo, o que temos é algo heterogêneo, nos moldes do que ensina a ANT, uma vez que os atores se articulam a permitir o engendramento de conjuntos híbridos, sendo eles atores e rede ao mesmo tempo, atuando de forma estratégica na associação de outros agentes. Essa associação heterogênea compõe relações complexas entre o que é humano e o que é tecnológico, já que do mesmo modo que o humano cria e adapta a tecnologia como ferramenta de fazer, a própria criação tecnológica afeta o indivíduo em todas as esferas de atuação do ser.

Parece-nos que, guiados por Latour, acertamos o caminho e sua forma de enfrentar os acontecimentos sociais/coletivos, já que a explicação dos conjuntos não está no conjunto em si, mas nas associações e nos movimentos produzidos, isto é, no fazer e no desfazer das ligações e que pode, aí sim, alterar substancialmente a estrutura do coletivo. Com isso, suscitamos que as concepções sobre autoria sofrem alterações a partir da forma com que a tecnologia avança, principalmente em relação aos dados digitais de pesquisa.

Os objetos conceituais e pragmáticos aqui trazidos não podem, de forma alguma, ser vistos de maneira isolada, mas são atores-rede afetados diretamente pela tecnologia. Devem, então, ser considerados atuantes importantes no processo de produção, compartilhamento e uso de dados digitais de pesquisa. Pensar a autoria de dados científicos é pensar atos, ações sempre complexas e que envolvem diversos agentes de uma sociedade – de um coletivo. É fazer agir uma série de outros agentes e dispositivos que tem sua posição constantemente modificada a partir de uma série de modos de existência.

Compreendemos, então, que a construção de um fato científico e social, como a autoria, é coordenada por uma série de elementos que foram levantados com o auxílio da ANT e que podem ser mediadores. Elementos como a coleta de

dados, a curadoria, os próprios dados científicos, seus metadados, as fontes de fomento, a política de dados, o repositório e seus softwares de gerenciamento e envio de dados e metadados, o licenciamento, a propriedade e a responsabilidade.

O primeiro ponto que gostaríamos de trazer diz respeito à **coleta técnica**. Quando um pesquisador se dirige a um sítio, parcela ou qualquer outra área de coleta de dados, está imbuído de um senso técnico para trazer à tona detalhes de sua observação do mundo que o rodeia, no caso aspectos da biodiversidade da Amazônia.

No entanto, minúcias relacionadas à tabulação de dados podem ser facilmente deixadas de lado devido a não ser dada a merecida atenção a aspectos de forma que devem ser contemplados para que os dados possam ser inseridos e disponibilizados no Repositório. Cabe ao pesquisador a consciência e a instrução sobre os procedimentos a serem adotados, como, por exemplo, a capacitação para coleta e tabulação de dados e o suporte técnico que abrangem a **curadoria**.

Como visto, a capacitação ocorre em formato de cursos, treinando os pesquisadores na coleta de dados para que depois possam ser úteis – no que se refere aos dados e metadados – para fins de novas pesquisas. Isso diz respeito à uma tabulação consistente, em formato que o banco de dados aceite a partir de características próprias.

No que se refere ao suporte técnico, esse é o âmbito que vai desde o assessoramento na curadoria de dados, ou seja, a análise dos dados pós coleta para que eles possam ser adicionados ao banco até a preservação e a administração do Repositório. É nesse espaço que o curador tem trabalho essencial, pois é ele quem coordena as atividades mencionadas anteriormente, ou seja, a formatação de dados e metadados capaz de servir ao seu propósito, isto é, estar disponível em um repositório.

Somado a isso está a preocupação em preservar os dados e, por consequência, o conhecimento contido neles a partir tanto do cuidado já descrito, isto é, uma tabulação e uma documentação condizente com os dados e sua possível (re)utilização, bem como todo aparato tecnológico utilizado para possibilitar sua

recuperação e assim possibilitar novos estudos e/ou aplicações em serviços e/ou bens para a sociedade.

Isso é consonante com a estratégia que o compartilhamento de dados científicos e a e-Science visam promover, algo próximo do que Gray (2009) chama de um novo modelo de ciência, com dados gerando novos conteúdos, a partir do tratamento, da produção, do gerenciamento, do acesso e do reuso dos dados em grandes centros de dados ou em computadores pessoais, poupando, assim, esforços humanos e financeiros. Promover o compartilhamento dos dados é promover um maior alcance e maior eficácia da ciência, fazendo com que ela atue como um bem e um serviço social.

Outro atuante que devemos levar em consideração são as **agências de fomentos**. Ao compreender que essas existem no sentido de amparar pesquisas para algum propósito, seja ele social, econômico, entre outros, deve haver uma separação entre os fundos privados e públicos. Quando se trata de um fundo privado, o interesse do investidor é, a partir de certa aplicação – geralmente financeira – na área, obter retorno no que diz respeito ao desenvolvimento de melhores produtos e, por consequência, competir em determinada fatia de mercado e obter lucros. Ao se trabalhar com fundos públicos, o que se pretende é estimular a produção de um local (país, região, estado etc.), aplicando recursos em áreas, geralmente, estratégicas e multidisciplinares, com vistas a alavancar a produção do local e melhorar a qualidade de vida de seus cidadãos.

Assim, o grande mote desses dados é proporcionar, através de uma **política de dados**, a possibilidade de que as agências financiadoras possam ter acesso e (re)utilizar os dados, buscando benefícios para os cidadãos. Isso parece ficar claro na Política de dados do Programa, já que ela está preocupada, entre outras coisas, com a difusão conhecimento, bem como o desenvolvimento de sistemas eletrônicos capazes de tratar o que pode se tornar informação sobre a biodiversidade brasileira e suas consequências em âmbitos econômico e social.

Nesse sentido, a popularização de normas pelas agências de fomento, como vistas a exigir a liberação de dados coletados em pesquisa com financiamento, proporciona, além da geração de novos conhecimentos, o estabelecimento de uma cultura de participação, de envolvimento, de engajamento nas questões sociais e na

preocupação com o bem-estar dos cidadãos. Prestar contas para a sociedade é um começo para pensar um fluxo contínuo de reúso e preocupação/consciência social.

Pensando, então, no caráter público dos recursos que subsidiam pesquisas, com especial atenção para a biodiversidade, foco no nosso trabalho, é correto pensar que o ente público atua como um tipo de detentor desse conhecimento, sendo necessário que haja um local para depósito e disponibilização desses dados. Em nosso caso, o **Repositório** de Dados Ecológicos do PPBio.

A importância dos repositórios está na possibilidade de que eles proporcionam acesso universal à grande quantidade de dados/informação científica, possibilitando que a visibilidade de autores e instituições, bem como dos próprios dados que agregam, seja potencializada. Tomando consciência disso, parece lógico chegar à compreensão da função do Repositório como importante dispositivo de compartilhamento de dados – também informação e conhecimento –, promovendo, para além da disponibilização dos dados, a integração entre pesquisadores e instituições.

Isso parece mais claro quando aceitamos que os **dados científicos de pesquisa** são, além de representativos de uma realidade, de uma materialidade, de determinados aspectos da natureza, construtos sociais. Ou seja, advêm da natureza e da operação do homem com ela e atuam em uma constante mutação do que vivemos hoje em dia, sendo modificados não apenas a partir de desafios tecnológicos, mas sociais e políticos. Isso também está ligado aos **metadados**, produtos da documentação dos dados científicos. Vemos também que definir dados de pesquisa ainda é algo complexo para pesquisadores que trabalham com coleta de dados.

Não existe, como podemos aferir, uma separação entre dados e metadados, uma vez que há necessidade de coexistência de ambos elementos. Dados sem metadados têm capacidade minimizada de comprimir sua função minimizada, quiçá inutiliza-los.

O Repositório, e sua atuação como meio de disseminação de dados, parte dos dados e metadados, suportado pelas **aplicações de gerenciamento do próprio repositório e do envio de dados e metadados**, respectivamente Metacat

e Morpho, agindo como centro útil para depósito, promovendo também o reconhecimento científico daqueles que coletam e sistematizam um conjunto de dados, isto é, os autores. Isso é possível a partir da interface de acesso aos dados, onde é permitido realizar pesquisas a partir da documentação do conjunto de dados e ter acesso à listagem de conjuntos recuperados com base na expressão de interesse utilizada no momento da busca.

Ao reconhecimento científico está ligada diretamente outra questão, o **licenciamento** de dados. Esse processo pouco explorado no contexto do PPBio pode afetar de maneira substancial a forma como a autoria está posta. Como a intenção é promover algo que possa ser de utilidade para o contexto de pesquisa do Programa, as pendências que seriam geradas com a discussão de direitos autorais são, em geral, deixadas de lado, simplificando – ou nos arriscando a dizer, mostrando falta de debate sobre esse conteúdo – a forma de disponibilizar acesso aos dados. Também é importante frisar que essa é uma preocupação do Programa, mas não a mais importante, sendo, em primeiro lugar, dada importância à disponibilização para reuso. O que se vê, desse modo, é que direito autoral está ligado ao aspecto moral da relação estabelecida entre o autor e sua obra que também é, ao mesmo tempo, patrimonial – em outras palavras, um bem econômico.

Chegando à **propriedade**, com base no que foi narrado, podemos dizer que não há preocupação profunda com essa questão. É claro que existe a ideia de garantir a autoria a quem deve ser creditada, mas isso passa por uma questão moral por parte do reutilizador. Ou seja, com base na Política de dados, há indicação para que o autor dos dados científicos seja contatado para saber da sua possibilidade de participar de novos estudos com base em reuso. Isso promove, para além do sentido de propriedade, um sentido de que os dados possam ser úteis para um indivíduo que desenvolve seus estudos ou para a sociedade como um todo. Essa promoção da utilidade dos dados para fins científicos e sociais provoca interações entre os pesquisadores e suas instituições

A disponibilização dos dados é um momento crucial para a efetivação científica, mas, em alguns casos, ainda pode sofrer certa resistência por parte dos pesquisadores, até mesmo por temor pelo destino dos dados. Mesmo entendendo que no momento em que se decide publicar algo em um ambiente aberto como a

internet há uma dualidade de sentimentos, é possível compreender que ao mesmo tempo em que se está suscetível a uma série de possibilidades negativas, há também entusiasmo por colaborar com outros indivíduos.

Há, nesse sentido, uma quebra do processo de receio de disponibilização, incentivando a reutilização, não havendo retrabalho de campo e, em consequência, economia de dinheiro, em geral, público. Agregado a isso estão os estímulos do pesquisador para colaborar e disponibilizar os dados coletados, como:

- (a) sentimento de ser útil, fazendo com que conhecimento seja compartilhado e sirva como uma espécie de interação e de troca com outros, passando a ser um processo de obtenção de novos conhecimentos: se hoje eu compartilho, amanhã alguém irá compartilhar e novos passos poderão ser dados;
- (b) senso de reconhecimento, fazendo com que seja possível aumentar seu capital social e científico através do resultado de novas pesquisas que utilizam os dados compartilhados;
- (c) obrigatoriedade de compartilhar o que foi coletado, já que ao realizar uma recolha de determinado sítio, o pesquisador de campo assina um termo no qual assume a responsabilidade por enviar ao Repositório, em um prazo estipulado, tanto os dados quanto os metadados.

Neste caso, a utilização de licenciamentos, assegurando os direitos autorais, parece primordial para, além de conferir responsabilidade, possibilitar que os indivíduos se sintam confortáveis a colaborar de forma mais segura, tendo certeza de que os dados depositados no Repositório terão o crédito reconhecido, percebendo que novos projetos podem vingar a partir da interação e da colaboração com outros pesquisadores.

O último item que elencamos em nossa reunião de elementos é a **responsabilidade**. Este elemento conceitual traz uma série de implicações que permeiam o processo de autoria de início a fim. Tem sua formação na coleta, no compartilhamento e também no reuso dos dados, deixando traços, inclusive, na percepção sobre que tipo de dados devem ser coletados e liberados para disseminação, o que estaria próximo da “possibilidade de prever os efeitos do próprio comportamento e de corrigi-lo com base em tal previsão” (ABBAGNANO, 2007, p. 855).

Pensar a responsabilidade como conceito ator-rede é primordial em diversos sentidos, mas a literatura dispõe de pouco conteúdo sobre a relação entre ética e autoria e, em consequência, sobre responsabilidade autoral, mas esse ponto é de extrema relevância como a construção do conceito de autoria, uma vez que a responsabilização pode trazer um caráter de reconhecimento ou de punição, caso, por exemplo, dados inconsistentes causem algum tipo de dano ao sistema ecológico.

Ao analisarmos o Repositório e suas políticas vemos a autoria a partir da responsabilidade na sua atuação como agente de credibilidade dos conjuntos de dados dispostos no Repositório, uma vez que não existe um tipo de verificação sobre a qualidade dos dados, mas a credibilidade vem a partir da premissa de que “todos os dados que entram no nosso Repositório são coletados por instituições de pesquisa [...]” (E2). Ou seja, a própria credibilidade da instituição a que estão vinculados os pesquisadores é uma aferição da possibilidade de publicar/disponibilizar, funcionando como certo tipo de “confiança científica” em parceiros. Podemos ilustrar esse caso a partir do termo de compromisso (Anexo B) do PPBio, que rege a disponibilização de dados vinculados ao uso da infraestrutura e às normas de conduta dos pesquisadores enquanto estiverem trabalhando em sítios permanentes de pesquisa.

Fica clara uma coordenação de esforços para a responsabilização sobre a autoria, já que o indivíduo tem a incumbência de depósito e boa atitude junto ao sítio pesquisado, oportunizando, uma vez obedecidas as normas acordadas, desenvolver novas pesquisas nos locais utilizados bem como receber o devido crédito pela publicação de dados.

A importância dessas características sobre a responsabilização pode ser observada como Latour (2000) pensa isso a partir da persuasão, refletida nas citações que um autor recebe, sendo isso um indicativo de força, fazendo com que a arregimentação de outros atores – citantes e citados – permitam com que os argumentos sejam considerados válidos. Isso pode ser estendido para o compartilhamento dos dados de pesquisa, cumprindo a atividade científica plenamente e não como algo isolado, promovendo (re)utilização daquilo que é um bem coletado, o próprio conjunto de dados de pesquisa.

A autoria busca superar uma visão antagônica entre o ruim e o bom, entre o certo e o errado, aceitando que essas dicotomias são cooriginadas e acarretam na tensão constante e perpetuada durante a “vida” do coletivo, ou seja, desde a formação do coletivo até sua dissolução. No entanto, como mencionado a partir de Droit (2012), a execução de nossas ações, a decisão de fazer algo a partir de uma consciência do correto ou do errado, tornando-nos responsáveis pelas consequências de nossas escolhas.

Na responsabilidade há uma pretensa capacidade de prestar atenção à inquietação que existe entre a possibilidade de dizer sim ou não diante de determinadas situações, de ter o poder e a capacidade de discernir entre os aspectos mais implícitos de uma situação, aspectos de possibilidade e de consequência. A forma como isso se coloca varia conforme os casos a serem analisados, mas não há como fugir de um processo ético/moral capaz de realizar julgamentos baseados tanto na possibilidade como nas consequências.

As possibilidades de tratar e/ou disponibilizar dados e informações de forma eletrônica permite com que os objetos digitais possam ser, como destaca Chartier (2002), deslocados, recortados, estendidos e recompostos. Isto faz com que a autoria seja posta em crise, com dificuldades, inclusive, de delimitações do que define autoria a partir de suas responsabilidades – para que possa ser entidade de um processo de atribuições de sanções de incentivo e/ou punição, conforme orienta Foucault.

A autoria concerne um ambiente propício para a base científica como a coleta, armazenamento, disponibilização e responsabilização, atuando com base em aspectos tecnológicos e que agem, diferentemente da autoria de textos, na coleta e composição da existência, da natureza, do que é material, seja qualquer instância coletada.

De forma ampla, os dados – como criação intelectual – derivam do esforço de algum indivíduo, de algum autor. Havendo autoria constituída, há uma autoria que é instituída a partir do produto de seu fazer. Como aponta Chartier (1998, p. 28), o resultado do fazer vai além de “uma propriedade [que] não apenas [é] pensada em termos econômicos e financeiros, mas em termos de controle e exatidão”.

Se a obra – os dados científicos digitais – é um produto intelectual ou um produto do esforço de coleta, da reunião de dados que certas entidades mantêm controle, o que resta é a propriedade sobre o conteúdo deste conjunto, e neste sentido, a propriedade é um conjunto de etapas que torna o dado científico algo público, mas também um conjunto com vistas ao reconhecimento científico do pesquisador.

A autoria age como um tipo de instrumentalização de um arranjo de fenômenos, isto é, uma ferramenta que, atuante como conceito e/ou como processo, permite que seja consolidado um sistema de comportamento, instituído a partir de regras explícitas – como a política de dados – bem como regras implícitas, originárias de um longo período de embates, acordos e símbolos científicos, sendo possível depreender que a autoria mantém vínculo com normas, com regras e com diretrizes. Ao compreendermos que as próprias normas são advindas dos sujeitos e suas associações, e por isso seus interesses, percepções, posições políticas etc., os indivíduos já posicionam o coletivo como a entidade a ser analisada. Este coletivo é, por tudo que defendemos neste trabalho, um processo de resgate de posições, de compreender atores que são redes e não podem ser vistos separadamente, sendo, sem dúvida, coletivamente existentes.

Isso está ligado ao que entendemos por dado, informação e conhecimento, em suas acepções originais. Notamos isso ao entender que dados, inclusos os científicos, são um átomo ou um fato (SEMIDÃO, 2014), ou seja, são objetivos, sendo fragmentos antecessores, na maioria dos casos, ou sucessores, em alguns casos, de uma entidade funcional, como a informação. Esta que, por sua vez, é componente de significação dos dados a partir de seu ordenamento e estruturação, permite sua utilização a partir das inscrições em determinado suporte (SETZER, 1999), possibilitando conhecimento, ou seja, a rede estabelecida pelos conceitos que possibilita, a partir de determinada visão de mundo, a elaboração de novos conhecimentos e, por consequência, novos conceitos.

O que estabelecemos aqui é a compreensão de que os procedimentos ligados à autoria de dados científicos são, em essência, conceituais em forma, sem descartar o ferramental pragmático utilizado. Não estamos ligados, nesse sentido, a um processo exclusivamente maquinário e/ou tecnológico, já que isso poderia tornar

a análise um apelo reservado ao imperativo tecnológico, mas, sim, aos atores-rede conceituais que formam a ideia de autoria e que fecham a caixa-preta que buscamos abrir ou, pelo menos, estabelecer menor opacidade.

Mesmo em forma de objetos conceituais, apreender esses atores-rede é compreender a sua igualdade, simetria, fazendo com que as relações entre humanos e não humanos não sejam denominadas por um nome, mas por ações que modificam um cenário. Assim, vistos como funções conceituais, seus “usos” são incertos por outros atores, mas compõem uma gama de associações que constroem o conceito de autoria.

Podemos concluir que não há autoria fora da ação. Embora isso seja algo que parece óbvio, – e o é se pensarmos que nenhum tipo de coleta/posição/discurso ocorre fora de um contexto de ação – remete-nos a tessitura do fato, à criação de um objeto complexo, como a autoria, a partir de um emaranhado coletivo, colaborativo em essência, cercado por uma contextualização de tempo e espaço. Em consequência, há de se atribuir a esse processo a responsabilidade por seus atos, arcando com o bônus e o ônus de pensar e de agir, de refletir e de promover modos de anotações, de escritas, de dados. Isso nos parece ser consoante ao que Foucault descreve como função-autor.

Esses processos podem ser vistos como convênios, ou seja, um tipo de contrato social e estão colocados entre o(s) indivíduo(s) autor(es) e o(s) ato(s) de autoria, as instituições que são utilizadas como depósito e disseminadoras de dados, bem como as questões políticas e éticas envolvidas no processo. Esse grande espectro leva à responsabilidade da autoria, isto é, a autoria é um ato que congrega autonomia e responsabilidade visando busca por reconhecimento e persuasão de determinado público. A autonomia está colocada na criação de determinado conjunto de dados e a responsabilidade está pautada na questão ética, ambas já apresentadas. Ou seja, não há autoria apenas com o autor, mas também a partir de uma série de mediadores, humanos e não humanos, instrumentais e conceituais.

8 CONCLUSÕES

Trabalhar com questões do âmbito social faz com que sempre relativizemos algo. Isto ocorre devido à impossibilidade de se estabelecer regras universais para compreensão de tais fenômenos, uma vez que esses estão sempre condicionados a fatores culturais – em amplo senso – que alteram a forma como a análise é realizada dependendo da comunidade escolhida.

A possibilidade de trabalhar com a autoria de dados de pesquisa permitiu que compreendêssemos esse aspecto como uma questão de amplo escopo, de forma a entender que determinada comunidade compartilha determinados valores, determinadas atitudes/práticas que impactam o coletivo (micro e macro). Assim sendo, o que se pretendeu defender neste trabalho – e acreditamos que a missão foi cumprida – é a compreensão da autoria como algo constituído através de práticas, comportamentos, instrumentos e conceitos organizados a partir e sobre um (meio) ambiente, tomando em consideração seus atos como uma forma de alterar a natureza pré-estabelecida, isto é, categórica. Isto leva a crer que a autoria é composta pelos elementos descritos e pode ser compreendida na forma da ação de seus atores-rede, como visto no capítulo anterior.

Diferentemente de um trabalho de cunho filosófico, o qual busca explicações e regras universais capazes de atuar como fio condutor de aspectos da vida social, nossa empreitada se pautou na investigação de fatores constituintes de um aspecto, isto é, a autoria de dados científicos. Neste sentido, não falamos aqui em uma teoria estabilizada – como preconiza a Teoria Ator-Rede utilizada em nosso estudo – mas em um processo associativo capaz de revelar estruturas que, coletivamente, compõem, ao longo do tempo ou em parte dele, um todo.

De certo modo, foi preciso que abrissemos mão de uma série de modelos que poderiam comprometer a análise baseada na Teoria Ator-Rede. E podemos dizer que este processo não foi fácil, uma vez que o estabelecimento de relações aparece, em nossa ideação, a partir de nossos modelos mentais já estabilizados a partir da agregação de valores constantemente reforçados. Ao olharmos por outro viés, percebemos que a proposta a ANT, mesmo arriscada, foi produtiva ao possibilitar a compreensão do fluxo de informações existente na rede; informação que liga os atores-rede; atores-rede que ligam a informação.

Isto é, o retrato espaço-temporal de uma organização que se modifica de forma constante, que emana e é emanada por n ações. Retomando Latour (2012, p. 131), entendemos após este trabalho que “a rede é uma expressão para avaliar quanta energia, movimento e especificidade nossos próprios relatos conseguem incluir. Rede é conceito, não coisa”. Isto descreve fielmente a constituição de redes, já que elas não são estruturas que permitem o encaixe de atores em determinadas posições, mas são conceitualmente construídas a partir do enlace de diversos desses actantes, capazes de captar a energia necessária para composição da rede, de forma a entender a perspectiva de um coletivo.

Assim sendo, a Teoria Ator-Rede mostrou-se um método revelador em dois sentidos: em primeiro lugar como uma teoria/método capaz de mobilizar um coletivo emergente a partir de, como sua proposta original, uma investigação nos atores humanos e não humanos que permeiam a rede investigada. O segundo sentido revelador, o qual existe em decorrência do primeiro, diz respeito ao processo que é estabelecido para que se possa trazer à baila os elementos que formam a rede, isto é, os próprios atores-rede. Estas possibilidades de pensar um ambiente, alicerçada em seus atores-rede, faz pensar a própria ciência como um cenário propício para a fuga de categorias pré-estabelecidas.

No entanto, não podemos nos furtar de dizer que esta forma de investigação é mais penosa do que o estabelecimento prévio de categorias e um “encaixe de atores” que ali residem. Compreender que, mais do que atores, os atores são atores-rede, fez com que os alicerces naturais fossem desestabilizados, fazendo emergir uma cultura que até então poderia ser considerada como já estabelecida, já fundada a partir de preceitos postos, mas isto refletiria um ponto estagnado sobre o entendimento da ciência, principalmente no que se refere à autoria de dados de pesquisa.

Passamos por um trajeto penoso, porém valioso. Valioso, porém penoso. Conforme prega a ANT, podemos conferir que a autoria é, ao mesmo tempo, um processo e um produto coletivamente construído, sendo, neste sentido, negociada por uma série de actantes que se envolvem em todas as etapas do compartilhamento de dados científicos.

A identificação de uma porta de entrada na rede possibilitou dela participar. Neste estudo, a porta de entrada, a controvérsia identificada de forma a fazer o *script* abrir o acesso à rede é colocado a partir da parca compreensão que a literatura fornece sobre a autoria e o próprio autor. O que é? Do que é composta? Como se comporta? Qual o impacto conceitual à sua existência ou não? É atividade/sujeito uno ou plural? Com que(m) se relaciona? Afeta o social/coletivo? Afasta-se de uma ideia já estabelecida a partir da autoria de textos? Esse assunto nos parecia nebuloso, acordado tacitamente de uma maneira para ser deixado de lado. Cremos ter feito um trabalho de não deixar isso de lado, confirmando nosso pressuposto de que a autoria tem sua essência alterada a partir do paradigma da e-Science e do compartilhamento de dados científicos, uma vez que compreende não necessariamente a criação intelectual, mas uma forma de observação e descrição da natureza, do material, do que é coletado como análise de um fenômeno.

Procuramos também tratar e discutir dados científicos e seu compartilhamento como um movimento que ganha cada vez mais corpo e que está promovendo certa algazarra que os pesquisadores mais tradicionais talvez não consigam acompanhar. Os dados jogam fora uma série de preceitos já estabilizados fechados em suas caixas escuras. No entanto, dados por dados são apenas ruídos.

Assim, as descrições realizadas apontam para o que entendemos como um dos fatores que constroem e destroem sociedade/coletivos, isto é, momentos de decisão, onde a fragilidade e a firmeza de atores-rede constitui a rede a partir de suas mediações, entendendo que essas desorganizações estabelecem a criação e a dissolução de coletivos, como na autoria. Acreditamos que isso vai ao encontro do que Bruno Latour fala sobre a rede estar em constante movimento, podendo atribuir a ela um fator fundamental no movimento, um caráter desestabilizador e, posteriormente, agregador.

Através de nosso trabalho, chegamos a considerar que a autoria é um tipo híbrido de convênio conceitual, ou seja, é um acordo sobre determinada atividade e/ou assunto que se estabiliza a partir de tratados previamente estabelecidos. Ainda é relevante notar que esses tratados podem ou não ser percebidos como tal, havendo possíveis conjecturas não entendidas *per se* ou explicitadas. No entanto, no fulcro de qualquer questionamento está a compreensão do que é autoria, uma

vez que esse conceito liga o insumo ao conhecimento gerado, passando pelo compartilhamento e pela colaboração.

Por fim, buscando situar o compartilhamento de dados científicos e a colaboração científica, vemos que não são elementos a se pensar para o futuro. São ações no momento; que já acontecem e permeiam os dados científicos desde sua coleta até seu reúso. Perspectivas de estudos, fundamentações e aplicações não parecem faltar. Talvez devêssemos pensar em uma ciência para além da informação, para além do documento. São as inscrições nos documentos que trazem as informações necessárias com base em dados. Compreender a essência da moral/ética envolvida no processo de autoria talvez seja o bojo de ação. Acreditamos nisso.

Para trabalhos futuros prospectamos o desenrolar de estudos que envolvam a fundo a relação entre ética e autoria, uma vez que, além da pouca relação que existe na literatura, a compreensão desse fenômeno visa permitir maior entendimento tanto da relação do homem com a natureza fornecedora de dados e informações quanto da sua relação com outros homens, possibilitando que a responsabilidade emerja como conceito a ser trabalhado e discutido em âmbito acadêmico, reverberando na sociedade.

REFERÊNCIAS

ABBAGNANO, Nicola. **Dicionário de Filosofia**. 3. ed. São Paulo: Martins Fontes, 2007. 1014 p.

ANTONIO, Irati. Autoria e cultura na pós-modernidade. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 189-192, maio/ago. 1998. Disponível em: <<http://www.scielo.br/pdf/ci/v27n2/irati.pdf>>. Acesso em: 7 maio 2014.

APPELBE, Bill; BANNON, David. eResearch - Paradigm Shift or Propaganda? **Journal of Research and Practice in Information Technology**, v. 39, n. 2, p. 83-90, maio 2007. Disponível em: <<http://www.jrpit.acs.org.au/jrpit/JRPITVolumes/JRPIT39/JRPIT39.2.83.pdf>>. Acesso em: 30 ago. 2012.

APPOLINÁRIO, Fábio. **Metodologia da ciência**: filosofia e prática de pesquisa. São Paulo: Thomson, 2006. 209 p.

ARRUDA, Susana Margaret de; CHAGAS, Joseane. **Glossário de Biblioteconomia e ciências afins**: português – inglês. Florianópolis: Cidade Futura, 2002. 229 p.

BARBOSA, Antonio Luiz Figueira. Proteção do trabalho intelectual. In: _____. **Sobre propriedade do trabalho intelectual**: uma perspectiva crítica. Rio de Janeiro: UFRJ, 1999. p. 21-87.

BAUMGARTEN, Maíra. Comunidades ou coletividades? O fazer científico na era da informação. **Política & Sociedade**, v. 3, n. 4, abr. 2004. Disponível em: <<https://periodicos.ufsc.br/index.php/politica/article/view/2003/1750>>. Acesso em: 20 set. 2013.

BEAVER, Donald. Reflections on scientific collaboration (and its study): past, present, and future. In: COLLABORATION IN SCIENCE AND IN TECHNOLOGY: PROCEEDINGS OF THE SECOND BERLIN WORKSHOP ON SCIENTOMETRICS AND INFORMETRICS, 2., 2000, Berlin. **Proceedings...** Berlin: Free University

Berlin, 2000. Disponível em: <<http://www.collnet.de/workshop/Berlin2000.pdf>>. Acesso em: 13 dez. 2014.

BENAKOUCHE, Tamara. Duas culturas, três culturas...ou redes? Dilemas da análise social da técnica. In: BAUMGARTEN, Maíra. (Org.). **A era do conhecimento: Matrix ou Ágora?** Porto Alegre: UFRGS; Brasília: UnB, 2001. p. 45-60.

BLANKMAN, David; MCGANN, Jeanine. **Ecological Metadata Language: Practical Application for Scientists.** Albuquerque, NM: LTER Network Office, 2003.

BORGMAN, C. L. The Conundrum of Sharing Research Data. **Journal of the American Society for Information Science and Technology**, v. 63, n. 6, p. 1059-1078, jun. 2012. Disponível em: <<http://onlinelibrary.wiley.com.ez45.periodicos.capes.gov.br/doi/10.1002/asi.22634/pdf>>. Acesso em: 17 jun. 2012.

BORGMAN, Christine L. **Scholarship in the digital age: information, infrastructure, and the Internet.** Cambridge, Massachusetts; London, England: MIT Press, 2010. 336 p.

BRANCO, Sérgio. **O domínio público no direito autoral brasileiro: uma obra em domínio público.** Rio de Janeiro: Lumen juris, 2011. 294 p. Disponível em: <<http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/9137/Sergio+Branco+-+O+Dominio+Publico+no+Direito+Autoral+Brasileiro.pdf?sequence=1>>. Acesso em: 18 mar. 2015.

BRANCO JÚNIOR, Sérgio Vieira. Fundamentos para o domínio público no direito autoral brasileiro, **Liinc em revista**, v. 7, n. 2, p. 437-465, set. 2011. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/view/439/328>>. Acesso em: 15 mar. 2015.

BRASIL. Portaria nº 693 de, 20 de agosto de 2009. Institui, no âmbito do Programa de Pesquisa em Biodiversidade - PPBio, a Política de Dados. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 21 ago. 2009, Seção 1, p. 8-9.

BRASIL. Lei 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 19 fev. 1998. Disponível em: <<http://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=20/02/1998&jornal=1&pagina=11&totalArquivos=80>>. Acesso em: 12 out. 2014.

BRASIL. Lei 12.853, de 14 de agosto de 2013. Altera os arts. 5º, 68, 97, 98, 99 e 100, acrescenta arts. 98-A, 98-B, 98-C, 99-A, 99-B, 100-A, 100-B e 109-A e revoga o art. 94 da Lei nº 9.610, de 19 de fevereiro de 1998, para dispor sobre a gestão coletiva de direitos autorais, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 14 ago. 2013. Disponível em: <<http://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=1&data=15/08/2013>>. Acesso em: 12 out. 2014.

CALLON, Michel. Some elements of a sociology of translation: domestication of the scallops and the fishermen of Saint Brieuc Bay. In: LAW, John. (Ed.). **Power, action and belief: a new Sociology of Knowledge?** Sociological Review Monograph, p. 196-233. London: Routledge and Kegan Paul, 1986.

CAMPOS, Arnaldo. O direito de autor. In: _____. **Breve história do livro**. Porto Alegre: Mercado Aberto, 1994. (Série Revisão; 40).

CARBONI, Guilherme C. O direito de autor e seus desafios: os conflitos com a liberdade de expressão, o direito de acesso ao conhecimento, à informação e à cultura e o direito ao desenvolvimento tecnológico. In: REIS, Renata et al. (Org.). **Propriedade intelectual: interfaces e desafios**. Rio de Janeiro: ABIA, 2007. p. 59-123.

CHARTIER, Roger. **A aventura do livro: do leitor ao navegador: conversações com Jean Lebrun**. São Paulo: UNESP; Imprensa Oficial do Estado, 1998. 159 p.

CHARTIER, Roger. História intelectual do autor e da autoria. In: FAULHABER, Priscila; LOPES, José Sérgio Leite. **Autoria e história cultural da ciência**. Rio de Janeiro: Beco do Azogue, 2012a. p. 37-64.

CHARTIER, Roger. **O que é um autor?** Revisão de uma genealogia. São Carlos: EDUFSCAR, 2012b. 90 p.

CHARTIER, Roger. Os **desafios da escrita**. São Paulo: UNESP, 2002. 144 p.

CIANCONI, Regina de Barros. **Gestão do conhecimento**: visão dos indivíduos e organizações no Brasil. 2003. 287f. Tese (Doutorado em Ciência da Informação)– Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2003.

CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Brique de Lemos, 2008. 451 p.

CURIA, Luiz Roberto. Apresentação. In: SANTOS, Manuella. **Direito autoral na era digital**: impactos, controvérsias e possíveis soluções. 2009. São Paulo: Saraiva. 2009. p. xv-xvii.

DROIT, Roger-Pol. **Ética**: uma primeira conversa. São Paulo: Martins Fontes, 2012. 95 p.

DRUCKER, Debora Pignatari. Avanços na Integração e Gerenciamento de Dados Ecológicos. **Natureza & Conservação**, v. 9, n. 1, p. 115-120, jul. 2011. Disponível em: <<http://peld.inpa.gov.br/sites/default/files/Drucker%202011.PDF>>. Acesso em: 20 nov. 2014.

ELIAS, Norbert. **A sociedade dos indivíduos**. Rio de Janeiro: Jorge Zahar, 1994. 201 p.

FLICK, Uwe. **Introdução à pesquisa qualitativa**. 3. ed. Porto Alegre: Artmed, 2009. 405 p.

FOUCAULT, Michel. **Microfísica do poder**. 25. ed. São Paulo: Graal, 2012.

FOUCAULT, Michel. **O que é um autor?** 3. ed. Lisboa: Vega, 1997. 160 p.

FOUREZ, Gérard. **A construção das ciências**: introdução à filosofia e à ética das ciências. São Paulo: UNESP, 1995. 319 p.

FRAGOSO, João Henrique da Rocha. **Direito autoral**: da antiguidade à internet. São Paulo: Quartier Latin, 2009. 406 p.

FREIRE, Leticia de Luna. Seguindo Bruno Latour: notas para uma antropologia simétrica. **Comum**, Rio de Janeiro, v. 11, n. 26, p. 46-65, jan./jun. 2006. Disponível em:
<<http://www.ifcs.ufrj.br/~lemetro/pesquisadores/Leticia%20de%20Luna%20Freire/latour.pdf>>. Acesso em: 14 maio 2013.

GANDELMAN, Henrique. **De Gutenberg à internet**. 5. ed. rev. e ampl. Rio de Janeiro: Record, 2007. 295 p.

GRAY, Jim. Jim Gray on science: a transformed scientific method. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. (Ed.). **The fourth paradigm**: data-intensive scientific discovery. Washington: Microsoft Research, 2009. Disponível em:
<http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_jim_gray_transcript.pdf>. Acesso em: 20 maio 2013.

HAGUETTE, Teresa Maria Frota. A entrevista. In: _____. **Metodologias qualitativas na Sociologia**. 14. ed. Petrópolis: Vozes, 1987. p. 81-86.

HEY, Tony; HEY, Jessie. E-science and its implications for the library community. **Library Hi Tech**, v. 24, n. 4, p. 515-528, 2006.

HÜNING, Simone Maria; GUARESCHI, Neuza Maria de Fátima. Michel Foucault, Bruno Latour e algumas linhas de fuga na produção de conhecimentos. **Ciências Sociais Unisinos**, São Leopoldo, v. 47, n. 1, p. 64-71, jan./abr. 2011. Disponível em: <<http://www.redalyc.org/articulo.oa?id=93820778007>>. Acesso em: 14 maio 2014.

HOLANDA, André; LEMOS, André. Do paradigma ao cosmograma: sete contribuições da Teoria Ator-Rede para a pesquisa em Comunicação. In: ENCONTRO ANUAL DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO, 22., 2013, Salvador. **Anais...** Salvador: COMPOS, 2013. Disponível em: <http://compos.org.br/data/biblioteca_2050.pdf>. Acesso em: 10 out. 2013.

HURD, Julie M. The Transformation of Scientific Communication: A Model for 2020. **Journal of the American Society for Information Science**, v. 51, n. 14, p. 1279-1283, 2000.

KATZ, J. Sylvan; MARTIN, Ben R. What is research collaboration? **Research Policy**, v. 26, n. 1, p. 1-18, 1997.

KUHN, Thomas Samuel. **A estrutura das revoluções científicas**. 3. ed. São Paulo; Perspectiva, 1990. 257 p. (Debates; 115).

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos da metodologia científica**. São Paulo: Atlas, 2007. 315 p.

LANNOM, Laurence. Research Data. **D-Lib Magazine**, v. 17, n. 1/2, jan./fev. 2011. Disponível em: <<http://www.dlib.org/dlib/january11/01editorial.html>>. Acesso em: 7 abr. 2013.

LATOUR, Bruno. **A esperança de Pandora: ensaios sobre a realidade dos estudos científicos**. Bauru: EDUSC, 2001.

LATOUR, Bruno. **Ciência em ação**: como seguir cientistas e engenheiros sociedade afora. São Paulo: UNESP, 2000. 438 p.

LATOUR, Bruno. **Jamais fomos modernos**: ensaio de antropologia simétrica. 2. ed. São Paulo: Ed. 34, 2009. 150 p.

LATOUR, Bruno. **Reagregando o social**: uma introdução à Teoria Ator-Rede. Salvador; Bauru: EDUFBA; EDUSC, 2012. 399 p.

LATOUR, Bruno. Redes, sociedades, esferas: reflexões de um teórico ator-rede. **Informática na educação**: teoria e prática, Porto Alegre, v. 16, n. 1, jan./jul. 2013. Disponível em: <<http://seer.ufrgs.br/index.php/InfEducTeoriaPratica/article/view/36933/25939>>. Acesso em: 4 nov. 2013.

LATOUR, Bruno. WOOLGAR, Steve. **A vida de laboratório**: a produção dos fatos científicos. Rio de Janeiro: Relume Dumará, 1997. 310 p.

LAW, John. Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity. **Systems Practice**, v. 5, n. 4, p. 379-393, 1992.

LE COADIC, Yves-François. **A Ciência da Informação**. 2. ed. Brasília: Briquet de Lemos, 2004. 124 p.

LEMOS, André. Você está aqui! Mídia locativa e teorias “materialidades da comunicação” e “ator-rede”. **Comunicação & Sociedade**, ano 32, n. 54, p. 5-29, jul./dez. 2010.

LIPINSKI, Tomas A.; BRITZ, Johannes J. Rethinking the ownership of information in the 21st century: ethical implications. **Ethics and Information Technology**, v. 2, p. 49-71, 2000. Disponível em: <<http://link.springer.com/article/10.1023%2FA%3A1010064313976>>. Acesso em: 10 ago. 2014.

MANSO, Eduardo J. Vieira. **O que é direito autoral**. São Paulo: Brasiliense, 1987. 92 p.

MCBRIDE, Neil. **Using Actor-Network Theory to predict the organisational success of a communications network**. 1999. Disponível em: <<http://www.cse.dmu.ac.uk/~nkm/PAPERS/WTCPAP.pdf>>.

MCGARRY, Kevin. **O contexto dinâmico da informação**. Brasília: Briquet de Lemos, 1999. 206 p.

MEDEIROS, Jackson da Silva. CAREGNATO, Sônia Elisa. Compartilhamento de dados e e-Science: explorando um novo conceito para a comunicação científica. **Liinc em Revista**, v. 8, n. 2, p. 311-322, 2012. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/488/373>>. Acesso em: 10 maio 2013.

MERTON, Robert King. **Sociologia: teoria e estrutura**. São Paulo: Mestre Jou, 1970. 758 p.

MICHENER, William K. Meta-information concepts for ecological data management. **Ecological Informatics**, v. 1, p. 3-7, 2006. Disponível em: <http://www.researchgate.net/profile/William_Michener/publication/222535033_Meta-information_concepts_for_ecological_data_management/links/02e7e51e459d7a4380000000.pdf>. Acesso em: 21 nov. 2014.

MIRANDA, Antonio; SIMEÃO, Elmira; MUELLER, Suzana. Autoria coletiva, autoria ontológica e intertextualidade: aspectos conceituais e tecnológicos. **Ciência da Informação**, v. 36, n. 2, p. 35-45, maio/ago. 2007.

MORAES, Marcia. A ciência como rede de atores: ressonâncias filosóficas. *História, Ciências, Saúde — Manguinhos*, Rio de Janeiro, v. 11, n. 2, p. 321-333, 2004.

MORIN, Edgar. O conhecimento do conhecimento científico. In: _____. **Ciência com consciência**. 14. ed. Rio de Janeiro: Bertrand Brasil, 2010. p. 37-94.

MUSSO, Pierre. A filosofia da rede. In: PARENTE, André. (Org.). **Tramas da rede: novas dimensões filosóficas, estéticas e políticas da comunicação**. 2. ed. Porto Alegre: Sulina, 2010. p.17-38.

NATIONAL SCIENCE FOUNDATION. **Cyberinfrastructure Vision for 21st Century Discovery**. 2007. Disponível em: <<http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf>>. Acesso em: 13 jun. 2012.

NOBRE, Júlio Cesar de Almeida; PEDRO, Rosa Maria Leite Ribeiro. Reflexões sobre as possibilidades metodológicas da Teoria Ator-Rede. **Cadernos UniFOA**, n. 14, dez. 2010. Disponível em: <<http://www.foa.org.br/cadernos/edicao/14/47.pdf>>. Acesso em: 22 maio 2013.

OECD (ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT). **OECD principles and guidelines for access to research data from public funding**. Paris: OECD, 2007. Disponível em: <<http://www.oecd.org/sti/sci-tech/38500813.pdf>>. Acesso em: 14 jan. 2015.

PARENTE, André. Rede e subjetividade na filosofia francesa contemporânea. **RECIIS – R. Eletr. de Com. Inf. Inov. Saúde**, Rio de Janeiro, v. 1, n. 1, p. 101-105, jan./jun. 2007. Disponível em: <<http://www.reciis.cict.fiocruz.br/index.php/reciis/article/download/35/67>>. Acesso em: 20 jul. 2013.

PEDRO, Rosa Maria Leite Ribeiro. Redes e controvérsias: ferramentas para uma cartografia da dinâmica psicossocial. In: ESOCITE - JORNADAS LATINO-AMERICANAS DE ESTUDOS SOCIAIS DAS CIÊNCIAS E DAS TECNOLOGIAS, 7., 2008. Rio de Janeiro. **Anais...** 2008. v. 1. p. 1-18.

PEREIRA, Ana Lúcia Lage. **A rede como espaço multirreferencial de aprendizagem: construção do conhecimento na produção de inovação em TIC em um Instituto de Ciência e Tecnologia brasileiro**. 2013. 267 f. Tese (Doutorado em

Difusão do Conhecimento)–Faculdade de Educação, Universidade Federal da Bahia, Salvador, 2013.

PESCUMA, Derna; CASTILHO, Antonio Paulo F. de. **Projeto de pesquisa**: O que é? Como fazer? Um guia para sua elaboração. São Paulo: Olho d'Água, 2005. 96 p. (Método; 3).

PEZZINI, Flávia Fonseca et al. The Brazilian Program for Biodiversity Research (PPBio) Information System. **Biodiversity & Ecology**, v. 4, 2012, p. 265–274. Disponível em: <http://www.biodiversity-plants.de/biodivers_ecol/publishing/b-e.00083.pdf>. Acesso em: 14 nov. 2014

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBIO). 2015. Disponível em: <<http://ppbio.inpa.gov.br> >. Acesso em: 16 mar. 2015.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBIO). **Documento básico**. 2005. 52 p. Disponível em: < http://www.mct.gov.br/upd_blob/0011/11190.pdf>. Acesso em: 2 ago. 2014.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBIO). 2014a. **Núcleos regionais**. Disponível em: <<http://ppbio.inpa.gov.br/nregionais>>. Acesso em: 14 dez. 2014.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBIO). **Sobre**. 2014b. Disponível em: <<http://ppbio.inpa.gov.br/sobreppbio>>. Acesso em: 14 dez. 2014.

PROGRAMA DE PESQUISA EM BIODIVERSIDADE (PPBIO). **Termo de compromisso**. [2011]. Disponível em: <<http://ppbio.inpa.gov.br/sites/default/files/TERMO%20DE%20COMPROMISSO%20e%20Normas%20de%20conduta.pdf>>. Acesso em: 14 dez. 2014.

PRIMO, Alex. O que há de social nas mídias sociais? Reflexões a partir da Teoria Ator-Rede. **Contemporanea: comunicação e cultura**, v. 10, n. 3, p. 618-641, set./dez. 2012. Disponível em:

<<http://www.portalseer.ufba.br/index.php/contemporaneaposcom/article/view/6800/4681>>. Acesso em: 12 jan. 2015.

RICHARDSON, Roberto Jarry. **Pesquisa social: métodos e técnicas**. 3. ed. rev. e ampl. São Paulo: Atlas, 2010. 334 p.

RONAN, Colin A. **História ilustrada da ciência: da Universidade de Cambridge: volume I: das origens à Grécia**. Rio de Janeiro: Círculo do Livro, 1987. 136 p.

RUDIO, Franz Victor. **Introdução ao projeto de pesquisa científica**. Petrópolis: Vozes, 1978. 144 p.

SALES, Luana Farias. **Integração semântica de publicações científicas e dados de pesquisa**: proposta de modelo de publicação ampliada para a área de ciências nucleares. 2014. 264f. Tese (Doutorado em Ciência da Informação)–Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

SALVI, Rosana Figueiredo; BATISTA, Irinéa de Loudes. A análise dos valores na Filosofia da Ciência e na educação científica. In: ENCONTRO NACIONAL DE PESQUISA EM EDUCAÇÃO EM CIÊNCIAS, 6., 2007, Florianópolis. **Anais...** Florianópolis: ABRAPEC, 2007. Disponível em: <<http://www.nutes.ufrj.br/abrapec/vienpec/CR2/p1077.pdf>>. Acesso em: 4 maio. 2013.

SÁNCHEZ VÁSQUEZ, Adolfo. **Ética**. 10. ed. Rio de Janeiro: Civilização Brasileira, 1987. 267 p.

SANTOS, Manuella. **Direito autoral na era digital: impactos, controvérsias e possíveis soluções**. São Paulo: Saraiva. 2009. 182 p.

SEMIDÃO, Rafael Aparecido Moron. **Dados, informação e conhecimento enquanto elementos de compreensão do universo conceitual da Ciência da Informação**: contribuições teóricas. 2014. 198f. Dissertação (Mestrado em Ciência

da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

SETZER, Valdemar W. Dado, informação, conhecimento e competência. **DataGramZero**, Revista de Ciência da Informação, n. 0, dez. 99. Disponível em: <http://www.dgz.org.br/dez99/Art_01.htm>. Acesso em: 18 mar. 2015.

SILVA, Terezinha Elisabeth da Silva; TOMAÉL, Maria Inês. Repositórios institucionais e o modelo open. In: TOMAÉL, Maria Inês (Org.). **Fontes de informação na internet**. Londrina: EDUEL, 2008. p. 123- 149.

SILVA FILHO, Waldomiro José da. Considerações pós-éticas sobre a ação e a racionalidade. In: TEIXEIRA, Angélica. (Org.). **Especificidades da ética na Psicanálise**. Salvador: Associação Científica Campo Psicanalítico, 2005. p. 86-97.

SILVA FILHO, Waldomiro José da. Ética sem fundamentos. In: GOMES, Henriette Ferreira; BOTTENTUIT, Aldinar Martins; OLIVEIRA, Maria Odaisa Espinheiro. (Org.). **A ética na sociedade, na área da informação e da atuação profissional: o olhar da Filosofia, da Sociologia, da Ciência da Informação e da formação e do exercício profissional do Bibliotecário no Brasil**. Brasília: Conselho Federal de Biblioteconomia, 2009. p. 9-29.

SONNENWALD, Diane. H. Scientific collaboration. **Annual Review of Information Science and Technology**, v. 42, n. 1, p. 643-681, 2008.

SOUZA, Bianca Kelly. Que importa quem fala? – o desaparecimento do autor segundo Michel Foucault. **intuitio**, Porto Alegre, v. 4, n. 2, nov. 2011, p. 123-132. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/intuitio/article/view/9676>>. Acesso em: 9 dez. 2014.

TARGINO, Maria das Graças. Comunicação científica: uma revisão de seus elementos básicos. **Informação & Sociedade: Estudos**, v. 10, n. 2, p. 1-27, 2000. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/download/326/248>>. Acesso em: 4 jan. 2013.

TORRES-SALINAS, Daniel; MARTÍN-MARTÍN, Alberto; FUENTE-GUTIÉRREZ, Enrique. An introduction to the coverage of the Data Citation Index (Thomson-Reuters): disciplines, document types and repositories. **EC3 Working Papers**, n. 11, jun. 2013.

TORRES-SALINAS, Daniel; ROBINSON-GARCÍA, Nicolás, CABEZAS-CLAVIJO, Álvaro. Compartir los datos de investigación em ciência: introducción al data sharing. **El profesional de la información**, v. 21, n. 2, p. 173-184, mar./abr. 2012.

TRANNIN, Maria Cecília; PEDRO, Rosa Maria Leite Ribeiro. Cartografando comunicabilidades em redes sócio-técnicas: um estudo da rede Instituto Fábrica do Milênio. In: CONGRESSO BRASILEIRO DE SOCIOLOGIA, 13., 2007, Recife **Anais...** Recife: UFPE, 2007. Disponível em: <http://www.sbsociologia.com.br/portal/index.php?option=com_docman&task=doc_download&gid=1377&Itemid=171>. Acesso em: 14 ago. 2013.

WITTER, Geraldina Porto. Ética e autoria na produção científica. **Informação & Informação**, Londrina, v. 15. n. esp., p. 131-144, 2010. Disponível em: <<http://dx.doi.org/10.5433/1981-8920.2010v15nespp131>>. Acesso em: 14 jun. 2014.

ZIMAN, John. Michael. **Conhecimento público**. Belo Horizonte: Itatiaia; São Paulo: EDUSP, 1979. 164 p. (O homem e a ciência; 8)

APÊNDICES

APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

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

Pesquisa: Autoria de dados digitais de pesquisa: teias de uma rede em construção

Objetivo: Analisar a autoria científica a partir do compartilhamento de dados de pesquisa.

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

O Sr.(a) está sendo convidado(a) a participar como entrevistado(a) da pesquisa de doutorado acima citada. A sua participação na entrevista é voluntária e seu nome será mantido em sigilo. A entrevista será gravada para posterior transcrição e análise.

Eu, _____, declaro estar ciente da pesquisa e de seu objetivo e concordo em participar da mesma, permitindo a gravação da entrevista e sabendo do anonimato garantido.

Porto Alegre, ____ de _____ de 2014.

Assinatura do participante

Assinatura do pesquisador

Informações e/ou esclarecimentos:

Jackson da Silva Medeiros

jackson.medeiros@ufrgs.br

Telefone: (51) 82197331/33085162

Sônia Elisa Caregnato (orientadora)

sonia.caregnato@ufrgs.br

APÊNDICE B – PERGUNTAS PARA ENTREVISTAS

- (1) Poderia falar sobre o seu trabalho realizado com dados científicos e, mais especificamente, com o repositório de dados ecológicos do PPBio?
- (2) Qual o papel dos dados de pesquisa para o seu trabalho no repositório?
- (3) O que o repositório entende por dados de pesquisa? Como eles se apresentam?
- (4) Pela complexidade que os dados podem se apresentar, qual a importância da existência de documentação sobre eles?
- (5) As políticas científicas são complexas. Quem decide sobre criação, organização, manutenção delas no repositório?
- (6) Como saber se os dados são suficientemente consistentes para publicação? Como isso é estabelecido?
- (7) O que de fato é recebido pelo repositório? Existem, por parte do repositório, critérios para seleção e publicação dos dados?
- (8) Como ocorrem os incentivos públicos e/ou privados em relação ao repositório de dados?
- (9) Os modos de compartilhamento, uso, políticas etc., devem ser regulados por instituições, pelo Estado ou pela comunidade de pesquisadores ou pelos próprios autores ou outros? Por quê?
- (10) Consta no Anexo 1, item 6, *Da política de disponibilização de dados vinculados ao uso da infra-estrutura do Termo de Compromisso do PPBio para pesquisa em campo* que “A política de dados não estabelece nenhuma regra sobre autoria dos trabalhos científicos resultantes dos dados coletados, ficando a definição de autores e co-autores a critério do pesquisador”. Embora a responsabilidade sobre a implicação de atribuição de autoria seja do pesquisador, para o repositório, quais os elementos quem definem a autoria? Que efeitos essa atribuição de autoria reflete na citação dos dados de pesquisa?
- (11) Trabalhos apresentados em seminários, conferências, bem como artigos de periódicos científicos, apresentam uma estruturação já estabelecida na autoria. No que se refere aos dados científicos, parece haver uma confusão entre quem é a entidade a ser responsabilizada, se o

repositório ou se o proprietário. Isso pode, de alguma maneira, prejudicar o compartilhamento dos dados? Em que instâncias?

- (12) A quem os dados de pesquisa efetivamente pertencem? Os proprietários são necessariamente autores (e vice-versa)?
- (13) Como o repositório lida com a sua própria responsabilidade em relação aos dados científicos?
- (14) Há algum tipo de licença estabelecida dos dados entre autor e repositório? Se sim, que tipo?
- (15) Quem tem acesso aos dados?
- (16) Como lidar com a preservação de dados? Existem políticas estabelecidas para isso?

ANEXOS

ANEXO A - PORTARIA Nº 693, DE 20 DE AGOSTO DE 2009 – Institui, no âmbito do Programa de Pesquisa em Biodiversidade - PPBio, a Política de Dados

Institui, no âmbito do Programa de Pesquisa em Biodiversidade - PPBio, a Política de Dados.

O MINISTRO DE ESTADO DA CIÊNCIA E TECNOLOGIA, no uso de suas atribuições legais, em especial as que lhe confere o art. 87, parágrafo único, inciso II, da Constituição Federal, e tendo em vista o disposto no Decreto 4.339 de 22 de agosto de 2002 e o cumprimento dos dispositivos da Convenção sobre Diversidade Biológica, aprovada pelo Decreto Legislativo nº 2, de 3 de fevereiro de 1994, e promulgada pelo Decreto nº 2.519, de 16 de março de 1998; e

Considerando que o Programa de Pesquisa em Biodiversidade - PPBio tem entre seus objetivos gerais o fomento à geração e disseminação de informações e conhecimento sobre a biodiversidade brasileira para diferentes segmentos da sociedade;

Considerando que o Programa de Pesquisa em Biodiversidade - PPBio fomenta a criação de sistemas de informação, de bases de dados e gerenciamento de repositórios da informação sobre a biodiversidade brasileira;

Considerando que os dados gerados no âmbito do Programa de Pesquisa em Biodiversidade - PPBio podem ter interesse comercial e o seu uso pode gerar consequências econômicas e ambientais;

Considerando a necessidade de um arcabouço de princípios, regras e orientações para todos os participantes do Programa de Pesquisa em Biodiversidade e usuários das bases de dados geradas no âmbito do Programa, no que diz respeito à abrangência, à coleta, ao armazenamento, à propriedade, à autoria, ao compartilhamento, à citação, ao acesso e uso dos dados e das bases de dados;

Considerando a necessidade de evitar conflitos e de obter compromissos sobre as questões de propriedade intelectual;

Considerando a necessidade de observar o que dispõe a legislação vigente de propriedade intelectual e inovação, especialmente no que concerne à proteção aos direitos autorais, à propriedade industrial e à informação confidencial; resolve:

Art. 1º Instituir, no âmbito do Programa de Pesquisa em Biodiversidade - PPBio, a Política de Dados, com o objetivo de promover o gerenciamento das informações para os dados coletados sobre a biodiversidade brasileira e gerados no âmbito do Programa, seus acessos, usos e disseminação, na forma do anexo a esta portaria.

Art. 2º Esta portaria entra em vigor na data de sua publicação.

SERGIO MACHADO REZENDE

ANEXO

DA POLÍTICA DE DADOS DO PROGRAMA DE PESQUISA EM BIODIVERSIDADE - PPBio

DAS DEFINIÇÕES GERAIS

Metadados: conjunto de informações que acompanham e descrevem as características dos dados biológicos, ambientais, socioambientais e espaciais e as condições de sua coleta, por exemplo: local de coleta, data de coleta, nome do coletor, latitude e longitude, imagens digitais ou fotos, entre outras.

Dados: informações biológicas, ambientais, socioambientais ou espaciais adquiridas com recursos financeiros ou logísticos do PPBio ou por ações amparadas por este. Podem ser caracterizadas como dados digitais ou conjuntos de dados

armazenados e gerenciados por computadores; dados analógicos, oriundos de atividades do PPBio, ainda que não digitalizados, como anotações de campo, planilhas, cadernetas de coleta; e quaisquer relatórios ou mapas produzidos, em formato digital ou analógico, resultado da compilação, análise, reunião ou organização, utilizando como fonte conjuntos de dados do PPBio.

Dados preliminares: são aqueles capazes de fornecer informações básicas descritivas do material biológico coletado (exemplos: morfologia, coloração, tamanho), ou ainda, as informações associadas a esse, sem garantir a identificação taxonômica precisa do mesmo.

Dados consolidados: são aqueles capazes de fornecer informações refinadas e completas e, tanto quanto possível, definitivas sobre o material coletado, incluindo a identificação taxonômica.

Dados ostensivos: são dados preliminares ou consolidados que após respeitado o período de embargo podem ser utilizados sem restrição cujo acesso pode ser franqueado ao público em geral.

Dados sensíveis: são dados preliminares ou consolidados que, se liberados ao acesso público, possam resultar em efeito adverso ao local e/ou às comunidades de origem da mesma e por isso, passível de restrição. Podem ser considerados dados sensíveis (a) a localização de espécies que estejam na lista de espécies ameaçadas de extinção; (b) dados de espécie que possa ser roubada ou traficada por sua raridade ou valor econômico (considerando sua potencialidade: como fornecedora de produtos que venham a ser utilizados na indústria farmacêutica ou química; como agente de controle biológico; entre outras); (c) a localização de habitats e sítios arqueológicos, culturais ou históricos cujo acesso possa ameaçar sua integridade; (d) informações utilizadas em decisões de Política de Estado que possam vir a interferir no alcance das metas e objetivos da mesma. Casos particulares que não estejam listados nessa Política deverão ser encaminhados à Coordenação Executiva para avaliação pelo Comitê Científico e aprovação do Conselho Diretor.

Núcleos de Biogeoinformática: unidades de gerenciamento de sistemas informatizados, aplicativos, bases de dados e metadados, instituídos e mantidos pelos núcleos executores e núcleos regionais.

Comitê Gestor de Informação: colegiado responsável pela deliberação sobre questões técnicas, administrativas, infra-estruturais e operativas que venham a ocorrer durante a operacionalização e gerenciamento de dados e informações do PPBio.

Participantes do Programa: os Núcleos Executores, os Núcleos Regionais, coordenadores de projetos e coordenadores de redes temáticas e todos os pesquisadores, colaboradores, estudantes, técnicos e bolsistas vinculados a esses núcleos que assinarem o termo de compromisso com esta Política.

Núcleos Executores: instituições assim designadas por Termo de Compromisso e Gestão ou Convênio firmado no âmbito do PPBio.

Núcleos Regionais: instituições que trabalham em parceria e de forma coordenada com os núcleos executores.

Indicador de confiabilidade do dado: parâmetro que qualifica o dado quanto à precisão e acurácia da informação e que demonstra a confiabilidade e qualidade dos dados inseridos na base de dados do PPBio.

Divulgação ampla: disponibilização de metadados e dados ostensivos a todos os interessados a partir do portal do PPBio.

Divulgação restrita: disponibilização de dados sensíveis permitida mediante autorização ou senha de acesso.

Período de embargo: período no qual, dados sob restrições de uso e acesso, não são disponibilizados pelo portal, mas são passíveis de visualização pelo Comitê Gestor de Informação.

DAS OBRIGAÇÕES, ATRIBUIÇÕES E COMPETÊNCIAS

2. Sobre as obrigações, atribuições e competências.

2.1. Dos Núcleos de Biogeoinformática.

Seguir normas, padrões e procedimentos estabelecidos para sua atuação pelo Comitê Gestor de Informação;

Implementar mecanismos que garantam a segurança, acessibilidade, qualidade, longevidade, integridade e interoperabilidade dos dados e metadados do PPBio;

Cadastrar os usuários e participantes do programa na instituição na qual estão situados, mantendo seus termos de compromisso;

Monitorar a atualização dos indicadores de qualidade associados aos dados que integram a base de dados;

Respeitar o sigilo de informações consideradas sensíveis;

Criar e manter um portal de acesso aos dados e metadados do PPBio na Internet;

Desenvolver, avaliar e adotar ferramentas computacionais e os aplicativos necessários ao registro, gestão, documentação, análise, integração, busca, acesso, armazenamento, segurança e publicação dos dados do PPBio;

Manter registro das publicações que utilizaram dados do PPBio.

2.2. Do Comitê Gestor de Informação.

Estabelecer normas, padrões e procedimentos para atuação dos núcleos de biogeoinformática;

Revisar e atualizar anualmente, ou quando necessário discutir, juntamente com o Comitê Científico, os termos desta Política de dados;

Aprovar os parâmetros definidos pelos NBGIs em relação às ferramentas computacionais e os aplicativos necessários ao registro, gestão, documentação, análise, integração, busca, acesso, armazenamento, segurança e publicação dos dados e metadados do PPBio;

Estabelecer normas e padrões que garantam a segurança, acessibilidade, qualidade, longevidade, integridade e interoperabilidade dos dados e metadados do PPBio.

Definir e detalhar processos e padrões de armazenagem, segurança, recuperação, análise e publicação dos dados e metadados do PPBio;

Definir normas e procedimentos para o processo de replicação das bases de dados e metadados em servidores de outras instituições;

Promover o acesso e utilização eficiente dos dados por parte dos participantes, observando a legislação vigente;

Decidir, juntamente com o responsável pela inserção dos dados, sobre seu acesso por participantes em período de embargo;

Credenciar os participantes responsáveis pela inserção de dados no banco do PPBio conforme indicação do coordenador do projeto;

Decidir, juntamente com a instituição responsável pela obtenção dos dados, ouvido o Comitê Científico, sobre dados sensíveis que devam ser divulgados de forma restrita, assim como sobre os pedidos de acesso de terceiros;

Controlar o acesso às bases de dados mediante o fornecimento de senhas ou de emissão de autorizações de acesso.

2.3. Do Conselho Diretor.

Aprovar as revisões da Política de Dados propostas pelo Comitê Gestor de Informação e pelo Comitê Científico;

Sugerir mudanças nos aspectos da Política de Dados;

Arbitrar eventuais conflitos, resolver casos omissos, excepcionais ou questões relacionadas a esta Política de Dados.

2.4. Do Comitê Científico.

Orientar o Comitê Gestor de Informação e assessorar o Conselho Diretor;

Definir as categorias de dados sensíveis utilizados nesta Política e deliberar sobre a inclusão ou exclusão de dados da categoria de dados sensíveis para fins de divulgação restrita;

Sugerir ações ao Comitê Gestor de Informação e aos Núcleos de Biogeoinformática;

Revisar e atualizar anualmente ou quando necessário, juntamente com o Comitê Gestor de Informação, os termos desta política de dados;

Sugerir ao Conselho Diretor ações e estratégias de comunicação para disseminar conhecimentos de biodiversidade à sociedade.

2.5. Dos Participantes.

Aceitar o teor dessa política de dados por meio de assinatura de termo de compromisso;

Repassar ao PPBio todos os dados, metadados ou conjunto de dados gerados com recursos do programa, respeitados os prazos e condições estipulados nesta Política;

Classificar os dados gerados em relação às categorias de dados sensíveis definidas pelo comitê científico;

Responsabilizar-se pela qualidade e repasse de todas as informações ao banco de dados do PPBio;

A inclusão de dados obtidos por outros programas nas bases do PPBio deverá respeitar preceitos legais e políticas institucionais, devendo ser acompanhada de autorização escrita da instituição de origem.

DA GESTÃO E AUTORIA DOS DADOS

3. Sobre a gestão e autoria dos dados.

Os dados/metadados ou conjunto de dados/metadados gerados com recursos do PPBio são de interesse público para o desenvolvimento científico - tecnológico e sua gestão é de responsabilidade do Ministério da Ciência e Tecnologia.

O PPBio deverá resguardar a autoria dos dados nas bases de dados e nas publicações resultantes.

A co-autoria ou outras formas de citação da participação na geração, análise e publicação dos dados deverão ser definidas pelas partes envolvidas, refletindo a participação intelectual, de acordo com o código de ética da ciência.

DAS BASES DE DADOS - USOS E ACESSOS

4. Sobre as condições de uso e acesso das bases de dados.

4.1. As bases de dados e metadados do PPBio serão protegidas por mecanismos adequados de prevenção e proteção à acessos não autorizados.

4.2 Os dados coletados, gerados e disponibilizados no âmbito do PPBio são de utilização prioritária para fins educacionais, culturais, científicos, de divulgação e de gestão pública. O acesso e uso com intenção comercial ou de forma que possa resultar na geração de produtos ou processos passíveis de exploração econômica, deverá ocorrer mediante a celebração de contrato entre as partes interessadas, observada a legislação pertinente e as disposições desta Política.

4.3. O acesso aos dados sensíveis e àqueles em período de embargo far-se-á de forma restrita por meio de autorização do Comitê Gestor de Informação mediante consulta ao responsável pela inserção de dados no banco do PPBio e demais partes interessadas.

4.4. O acesso aos dados via portal na Internet deverá ser feito mediante declaração de aceitação das condições de uso e acesso por meio da assinatura de um Termo de Compromisso disponível no portal.

4.5. Recomenda-se a todo usuário, no caso de encontrar um dado que julgue incorreto, informar por meio de formulário próprio disponível no portal, ao pesquisador (ou grupo de pesquisa) responsável pela inserção do dado e ao Comitê

Gestor de Informação, para que estes possam avaliar e, quando couber, providenciar a correção sugerida.

4.6. Os pesquisadores, as instituições participantes do PPBio, assim como o Ministério da Ciência e Tecnologia, não poderão ser responsabilizados em nenhuma hipótese por qualquer dano, consequência ou prejuízo que a utilização dos dados tornados públicos venha eventualmente causar, seja a pessoas físicas, seja a pessoas jurídicas.

4.7. Todos os produtos resultantes da utilização de dados e metadados do PPBio deverão ser acompanhados dos devidos créditos ao PPBio.

DA PROPRIEDADE INTELECTUAL

5. Sobre a proteção e propriedade intelectual.

5.1. Os produtos e processos decorrentes de informações e pesquisas concebidas ou executadas no âmbito do PPBio poderão ser protegidos e/ou patenteados segundo a legislação vigente, desde que seja observado o disposto no subitem 4.2.

DOS PRAZOS

6. Sobre os prazos.

6.1. Os metadados devem ser disponibilizados ao Comitê Gestor de Informação no prazo máximo de 30 dias após a coleta dos dados, tornando-se passíveis de consulta pública via portal na Internet em um prazo máximo de 7 dias após o repasse ao Comitê Gestor de Informação.

6.2. Os dados preliminares devem ser disponibilizados pelos autores ao Comitê Gestor de Informação no prazo máximo de 12 meses após a coleta, podendo ser nesse período visualizado pelo CGI. Findo esse período, os dados serão tratados como ostensivos, exceto se houver solicitação de prorrogação ao Conselho Diretor. Dados consolidados devem ser disponibilizados pelos autores ao Comitê

Gestor de Informação no prazo máximo de 24 meses após a data da coleta. Em casos excepcionais, o prazo poderá ser estendido, desde que autorizado pelo Conselho Diretor do PPBio.

ANEXO B – TERMO DE COMPROMISSO

TERMO DE COMPROMISSO

Eu, _____
_____, CPF nº _____, identidade
nº _____,
email: _____

instituição _____,
solicito autorização para coletar dados para o meu projeto de pesquisa, nível
_____, de
título“ _____
_____”. Para isso, declaro que estou ciente dos
itens abaixo:

- (1) da política de disponibilização de dados vinculados ao uso da infra-estrutura (Anexo 1)
oferecida pelo projeto
“ _____”;
- (2) das normas de conduta nos sítios permanentes de pesquisa (Anexo 2);
- (3) de que para compatibilizar as diferentes pesquisas de curta e longa duração nos sítios de
pesquisa da _____ somente será concedida autorização para pesquisa
após avaliação da versão preliminar dos metadados (Anexo 3) de pesquisa entregue
pelo pesquisador;
- (4) de que toda expedição à _____ deverá ser
informada com pelo menos 5 (cinco) dias de antecedência via e-mail para o gerente dos
sítios de pesquisa da _____ ([mail](#) de contato);

Sendo assim, me comprometo a agir em acordo com os quatro itens anteriores sabendo que
caso contrário estarei sujeito à proibição permanente de acesso às áreas de pesquisa da rede de
parcelas permanentes do PPBio.

(local e data)

(nome e assinatura)

Anexo 1 Da política de disponibilização de dados vinculados ao uso da infra-estrutura

- (1) O projeto segue a política de dados do PPBio e todos os dados coletados usando a infra-estrutura do projeto deverão ser públicos nos prazos estipulados pela política de dados adotada. Coletar dados usando a infra-estrutura do projeto implica que o pesquisador está de acordo com essa política, com as instruções abaixo e prazos estipulados. Para maiores informações sobre a política de dados do PPBio, consulte http://ppbio.inpa.gov.br/Port/docsinternos/politica_dou.pdf ou entre em contato com ppbio@inpa.gov.br;
- (2) Os metadados (informações detalhadas sobre quais dados foram coletados em cada lugar, quando, como e por quem) deverão ser disponibilizados em até 30 dias após cada uma das expedições de campo e deverão seguir o modelo disponível em http://ppbio.inpa.gov.br/Port/dadosinvent/Metadados_modelo.doc/;
- (3) Para fins de segurança, as anotações de campo e dados brutos coletados devem ser digitalizados e encaminhados juntamente com os dados coletados nos sítios permanentes de pesquisa para ppbio@inpa.gov.br em até 1 (um) ano após o término das atividades de campo do projeto. O PPBio disponibiliza um bolsista para auxiliar o pesquisador na digitalização de dados de campo, para maiores informações escreva para ppbio@inpa.gov.br;
- (4) Os dados coletados poderão ser disponibilizados no site por módulo, trilha, parcela ou segmento de parcela, dependendo da unidade amostral escolhida pelo pesquisador para cada dado coletado. Consulte os modelos de tabela de dados disponíveis em <http://ppbio.inpa.gov.br/Port/inventarios/br319> antes de preparar arquivo final para envio;
- (5) As coordenadas geográficas deverão ser coletadas em graus decimais (hddd.ddddd^o) com 5 casas decimais de precisão. Para evitar ambigüidades, os sistema de coordenadas padrão adotado pelo projeto é GCS (Geographic coordinate system - Sistema de coordenadas geográficas) e o *datum* WGS84;
- (6) A política de dados não estabelece nenhuma regra sobre autoria dos trabalhos científicos resultantes dos dados coletados, ficando a definição de autores e co-autores a critério do pesquisador. No entanto, as publicações científicas e qualquer outro meio de divulgação de trabalho de pesquisa, apoiados pelo Edital _____, deverão citar, obrigatoriamente, o apoio prestado pelos financiadores _____, utilizando suas respectivas logomarcas, de acordo com as normas de Uso da Marca da cada instituição.

Anexo 2 Das normas de conduta em nos sítios permanentes de pesquisa

Protocolo de uso das parcelas, trilhas e acampamentos do PPBio

O Programa de Pesquisas em Biodiversidade (PPBio) tem como sítios amostrais grades e módulos, com suas respectivas parcelas permanentes. Tanto as trilhas como as parcelas são unidades amostrais ou experimentais para um grande número de pesquisadores, de diferentes áreas de interesse, mas que tem em comum o interesse na integração de seus dados com os dados de outros pesquisadores trabalhando nas mesmas unidades amostrais. Dado o potencial para conflitos, já que alguns comportamentos e decisões de pesquisa podem resultar em prejuízo para os outros pesquisadores envolvidos, decidiu-se adotar um protocolo de comportamento para o uso da infra-estrutura de pesquisa fornecida pelo PPBio. Este protocolo visa informar aos pesquisadores sobre os diferentes tipos de pesquisa que estão sendo conduzidas nas unidades amostrais/experimentais, e delimitar regras de uso destas e dos acampamentos.

Antes de iniciar qualquer amostragem ou experimento nas áreas de pesquisa do PPBio, o pesquisador deve ler as instruções abaixo e submeter à gerência científica do PPBio (ppbio@inpa.gov.br seus metadados, para que a adequação da proposta às normas de conduta possa ser checada.

1. Parcelas

Parcelas terrestres ou uniformemente distribuídas

Estas parcelas são de 250 m de comprimento e largura variável, dependendo do objeto de estudo (Figura 1). A parcela é demarcada no campo apenas pelo corredor central. Este corredor tem 50 cm de largura e estende-se por todo o comprimento da parcela. O corredor delimita a área que deve ser usada para locomoção dentro da parcela e para algumas modalidades de amostragem destrutiva. O tópico sobre amostragens destrutivas será detalhado a seguir.

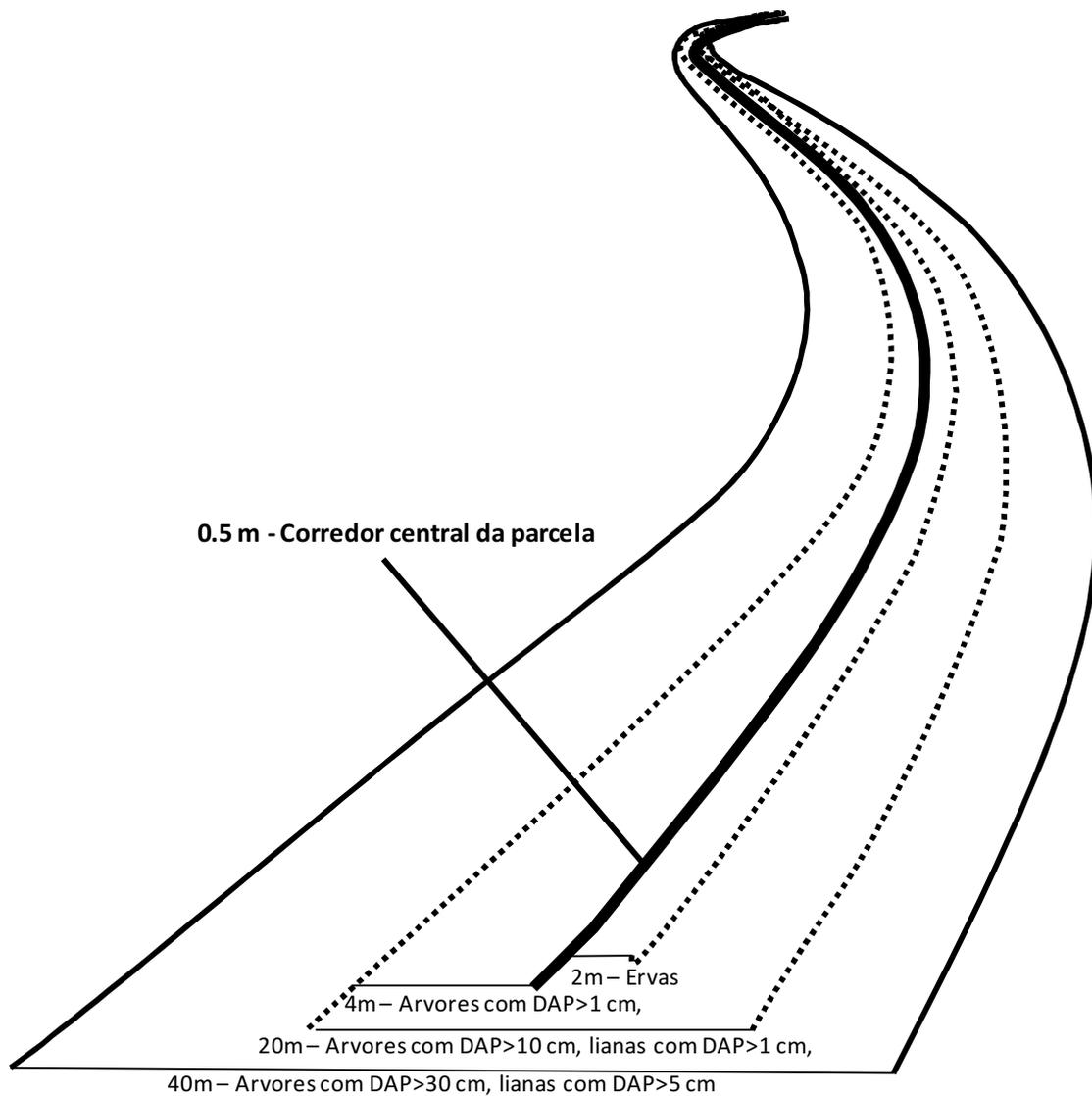


Figura 1: Esquema da parcela terrestre seguindo a curva de nível. A parcela é delimitada apenas pelo corredor central e a largura da parcela é variável de acordo com o objeto de estudo.

Zoneamento da parcela terrestre

Para viabilizar a integração e existência de pesquisas de longo prazo, as parcelas terrestres foram divididas em três zonas ou faixas (Figura 2). Cada uma destas faixas contempla uma das três principais formas de coleta de dados: amostragem de organismos sensíveis, amostragem de impacto, amostragem destrutiva/manipulativa

Parcela terrestre

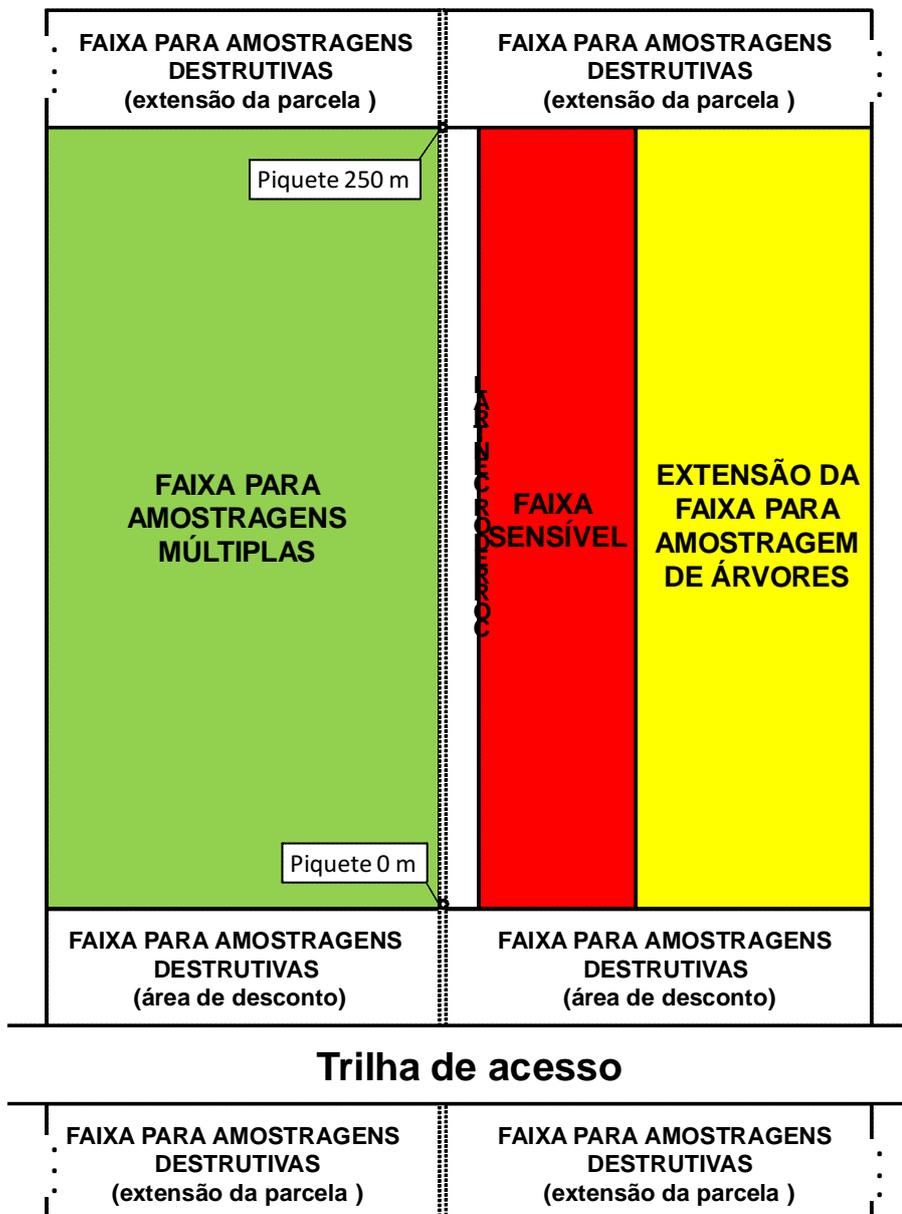


Figura 2. Diagrama ilustrando o zoneamento das parcelas terrestres e localização das faixas definidas para cada modalidade de pesquisa.

1) Faixa para amostragem de organismos sensíveis

Existem organismos mais sensíveis, que são rapidamente destruídos quando a área em que vão ser amostrados e monitorados não é isolada da passagem dos pesquisadores. O mero pisoteio continuado é suficiente para matar a maior parte das ervas, fungos e plântulas das espécies arbustivas e arbóreas. Outros organismos de pequeno porte que ainda não são estudados frequentemente podem também ser sensíveis. Portanto, as áreas reservadas para o acompanhamento destes organismos (e outros que venham a ser considerados sensíveis) devem ser evitadas ao máximo por todos os outros pesquisadores.

A área reservada para o estudo de organismos sensíveis é uma faixa de 2m do lado direito do corredor central.

Apenas quando é necessário coletar ou medir algum organismo que está do lado direito é que se deve entrar nesta faixa, de preferência pelo menor número de pesquisadores possível. Todas as atividades que podem causar maior impacto devem ser realizadas do lado esquerdo da linha central, ou em áreas especiais que serão definidas abaixo.

2) Faixa para amostragens múltiplas

A amostragem de organismos que não se encontram na lista de organismos sensíveis deve ser feita obrigatoriamente dentro desta faixa, que fica localizada do lado **esquerdo da parcela**. Atividades de pesquisa continuada (p.ex. fenologia, monitoramento de animais com busca ativa), que implicam em pisoteio da área, devem ser concentradas no **lado esquerdo** da parcela. Algumas vezes a área necessária é maior do que uma faixa de 20 m para o lado esquerdo, e assim parte da amostragem pode ser feita também do lado direito. Neste caso, a amostragem deve **evitar ao máximo o pisoteio na faixa para organismos sensíveis**. Nos casos em que for estritamente necessária a amostragem nesta faixa (p. ex. para medição de árvores), os organismos dentro desta faixa devem ser acessados pelo lado de fora da faixa. Quando não for possível acessar os organismos sem entrar na faixa sensível, o menor número de pesquisadores possível deve entrar nesta faixa.

3) Faixa para amostragens destrutivas

Qualquer amostragem que implique em destruição – remoção de solo, corte de plantas, reviramento de serrapilheira, entre outras - deve ser conduzida no corredor central ou fora da parcela permanente. A opção por um destes locais vai depender das características da amostragem.

- corredor central: deve ser usado para atividades que não causem grandes perturbações ou atrapalhem o deslocamento dentro da parcela;
- área de desconto (10m entre a trilha e piquete zero da parcela): deve ser usada para atividades pouco destrutivas e que requerem monitoramento freqüente (p. ex. piezômetros, extratores de água, armadilhas para invertebrados, monitoramento de umidade do solo, etc.);
- extensão no final da parcela ou do lado oposto da trilha: deve ser usada para atividades muito destrutivas (p.ex. monólitos de raízes, pitfalls para herpetofauna e outras

armadilhas de grande porte). Extensões da parcela podem ser facilmente estabelecidas com um clinômetro, seguindo as instruções dadas no protocolo para estabelecimento de parcelas (<http://ppbio.inpa.gov.br/Port/instalacao/>).

Amostragens destrutivas fora destas posições são expressamente proibidas. Lembre-se que as parcelas são permanentes, e muitos estudos de dinâmica são conduzidos. Mudanças nas características das parcelas por conta destas atividades destrutivas podem levar a conclusões errôneas nos estudos.

Alguns exemplos de atividades destrutivas e instruções para sua realização:

- *Coleta de solo, coleta de raízes* – estas atividades são feitas na faixa central, na área de desconto ou nas extensões da parcela, dependendo do tamanho da amostra coletada;
- *Instalação de Pitfalls (armadilhas de queda)*, especialmente os de grande tamanho usados para coletar répteis – recomenda-se que sejam instalados em extensões da parcela. É importante ter cuidado para que o solo removido na instalação não se espalhe e invada a parcela, pois isso modifica a vegetação e também os invertebrados do solo;
- *Instalação de redes de neblina para pássaros e morcegos* – a instalação deve ser feita na faixa central, removendo o mínimo possível os galhos e folhas que atravessam a faixa acima de 1,5 m de altura. É importante minimizar a remoção da vegetação, pois os estudos da estrutura da vegetação com sensores de laser (LIDAR) são realizados na mesma faixa. Se houver necessidade de maior remoção de vegetação, deve-se optar por fazer uma extensão da parcela para o lado oposto da trilha.
- *Coleta total de flores e/ou frutos*: coletas botânicas são permitidas nas parcelas, mas a coleta do indivíduo inteiro deve ser feita fora da parcela. A instalação de coletores de sementes e frutos deve ser feita na faixa para atividades de impacto intermediário.

Experimentos: É importante consultar a gerência científica do PPBio antes de iniciar qualquer experimento, para checar se a localização pretendida está de acordo com as normas de conduta e se o impacto é aceitável.

Experimentos que causam destruição ou modificação intensa do ambiente devem seguir as mesmas instruções dadas acima para as amostragens destrutivas, ou seja, devem estar antes do início ou depois do fim da parcela, ou em uma extensão desta. Se o experimento não causar modificação do ambiente (p.ex. experimentos de monitoramento), eles podem ser conduzidos dentro da parcela, mas devem ser alocados preferencialmente no lado esquerdo da faixa central. Caso precisem ocupar alguma área do lado direito, devem estar **depois da faixa de 2m reservada aos organismos sensíveis**. Experimentos com monitoramento constante devem ser instalados preferencialmente na área de desconto ou em na extensão da parcela para o lado oposto da trilha para evitar trânsito desnecessário pela parcela.

Parcelas aquáticas

As parcelas aquáticas representam um trecho de 50m do canal do igarapé e a área de entorno de suas margens até o início das vertentes. As parcelas são identificadas por canos marcando o início e final do trecho e são acessadas a partir da trilha principal por um corredor de acesso. O corredor de acesso é a única área que deve ser usada para deslocamento dentro da parcela aquática, pois o pisoteio e abertura de picadas em outras áreas do baixio pode afetar a configuração das poças e outros ambientes marginais monitorados.

No anexo 2 estão listados os estudos que normalmente são conduzidos nas parcelas aquáticas, sua localização dentro da parcela, os possíveis impactos sofridos e causados durante a coleta de dados. Os estudos foram divididos de acordo com a sensibilidade dos organismos estudados e grau de perturbação causado pela amostragem.

Zoneamento de parcelas aquáticas

Os igarapés são ambientes muito sensíveis, com baixa produtividade primária e que dependem basicamente da entrada de material externo para a manutenção de sua integridade. Qualquer atividade que aumente ou reduza a entrada de materiais nos igarapés é potencialmente impactante e pode levar a conclusões equivocadas nos estudos de monitoramento. É importante que os pesquisadores respeitem os trechos e faixas definidos para a execução de cada tipo de atividade. Esta atitude permite a integração dos estudos com os dados das parcelas aquáticas e evita alterações não-naturais dentro das parcelas.

1) Faixa sensível

O baixio é a principal fonte de entrada de material para o canal do igarapé. Durante as chuvas, este ambiente é conectado diretamente ao canal do igarapé e alterações nas suas características podem afetar organismos que usam o baixio. Vários grupos biológicos (p. ex. odonatas, peixes, girinos, coleópteras, etc.) estudados nestas parcelas tem parte do ciclo de vida associado à área alagada do baixio. O pisoteio continuado pode mudar a configuração dos ambientes alagados marginais e com isso a sua dinâmica hidrológica. Por estes motivos, o baixio da parcela aquática foi delimitado como área sensível, onde apenas os pesquisadores que irão trabalhar no próprio baixio devem andar. Os demais pesquisadores devem usar somente o corredor de acesso para transitar na parcela aquática.

Atividades no entorno dos igarapés, como revolvimento de serrapilheira, retirada de solo ou raízes, devem ser evitadas nas parcelas aquáticas. Caso sejam estritamente necessárias, devem ser feitas de forma cautelosa, evitando o carreamento de materiais para os igarapés. Muito cuidado deve ser tomado para não afetar a vegetação ripária durante estas atividades. A retirada desta vegetação pode afetar a estrutura das margens e a entrada de sedimento no canal do igarapé.

É expressamente proibida a retirada da vegetação marginal aos igarapés em parcelas aquáticas, independente da finalidade.

Parcela aquática

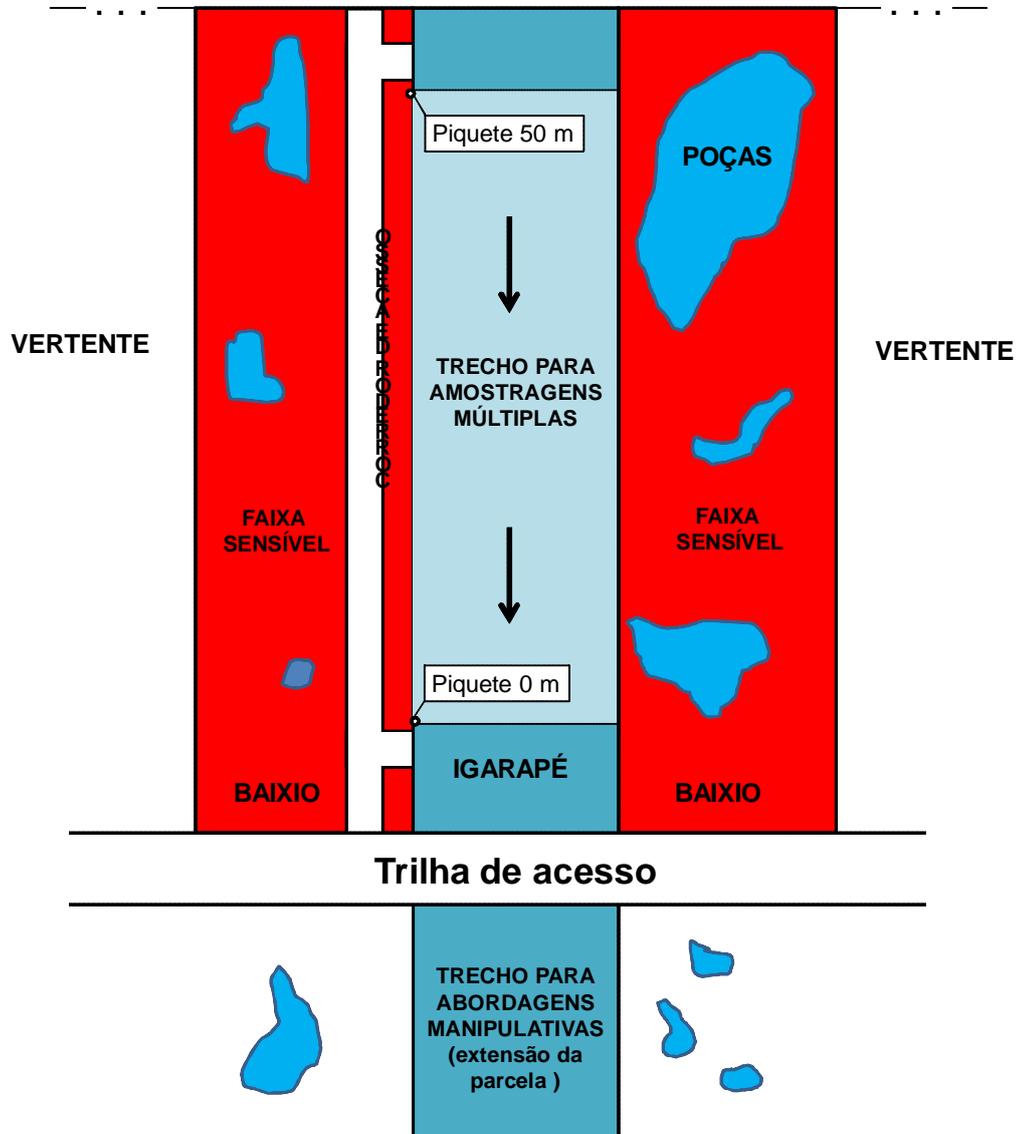


Figura 3. Diagrama ilustrando o zoneamento das parcelas aquáticas e localização dos trechos definidos para cada modalidade de pesquisa.

2) Trecho para amostragens múltiplas

Este trecho da parcela aquática deve ser usado para inventários iniciais e monitoramento da biota aquática. Quaisquer atividades que resultem em impacto de longo prazo devem ser conduzidas fora deste trecho. Armadilhas só são permitidas neste trecho quando a sua permanência não atrapalhar as demais atividades de amostragem no trecho. Caso contrário, as armadilhas de coleta devem ser montadas a jusante do piquete zero ou no trecho para abordagens manipulativas.

A retirada ou revolvimento intensivo de materiais dos igarapés, como areia, troncos ou folhas, não deve ser realizada dentro ou nas proximidades das parcelas aquáticas. Caso estritamente necessária, deve ser feita no trecho para abordagens manipulativas.

3) Trecho para abordagens manipulativas

Toda a parcela aquática (canal + baixio) é monitorada permanentemente, e qualquer atividade potencialmente impactante deve ser desenvolvida numa extensão imediatamente a jusante (abaixo) da parcela, acessado pelo lado oposto da trilha principal (Figura 3).

Seguem abaixo exemplos de atividades que podem alterar as condições ambientais das parcelas aquáticas que devem ser desenvolvidas dentro do trecho para abordagens manipulativas:

- Experimentos que incluam material externo aos igarapés (p.ex. experimentos de colonização de folhiço) ou que se baseiem no uso de equipamentos (p.ex. gaiolas para exclusão de predadores, armadilhas para captura de quelônios ou jacarés);
- Experimentos manipulativos nos baixios, que impeçam ou modifiquem o curso da água da chuva (p.ex. experimentos de isolamento de poças, gaiolas para exclusão de mamíferos);
- *Coleta zoológica* – coletas zoológicas são permitidas neste trecho, mas devem ser reduzidas ao mínimo necessário. A coleta de peixes deve ser evitada dentro das parcelas aquáticas, exceto em módulos/grades onde estão sendo realizados levantamentos iniciais. Sugere-se que os peixes coletados nestas áreas sejam fotografados vivos para auxiliar a identificação em campo nas próximas amostragens, reduzindo a necessidade de futuras coletas.

Na Reserva Ducke, já existe um guia preliminar de fotos para identificação de peixes em campo. Para maiores informações visite o site (<http://ppbio.inpa.gov.br/Port/guias/>)

2. Trilhas

As trilhas que delimitam as grades e módulos do PPBio também são áreas de pesquisa e precisam ser conservadas. Tipos de pesquisa normalmente conduzidas nas trilhas incluem os seguintes:

- levantamentos visuais e auditivos de mamíferos, aves e sapos
- levantamentos de árvores
- monitoramento da fenologia de plantas
- estrutura da vegetação
- monitoramento de necromassa (troncos caídos)

Ao usar as trilhas é preciso evitar desvios de seu trajeto, que possam perturbar as áreas laterais. Quando as pessoas passam a usar sistematicamente desvios das trilhas, a trilha muda de conformação, o que pode acabar afetando os resultados dos estudos.

Não é permitido cortar a vegetação na margem das trilhas. Pesquisadores e seus assistentes não devem tirar troncos ou galhos da trilha, pois estes fazem parte de estudos de queda da liteira lenhosa. Sempre que árvores caídas obstruindo a passagem forem encontradas, deve-se anotar sua posição e passar esta informação ao gerente de logística do PPBio (logistica.ppbio@gmail.com), para que este providencie a limpeza da trilha.

Se o seu estudo envolve a marcação de indivíduos ou locais ao longo das trilhas, faça esta marcação com material devidamente identificado com seu nome, para evitar confusões com marcações de outros pesquisadores. Um mapa das marcações deve ser entregue junto com os metadados do estudo, para ajudar os gerentes de logística e científico a localizá-las. Isto é importante para informar a outros pesquisadores que vão trabalhar na área para que respeitem os objetos de estudo marcados. Todas as marcações devem ser retiradas ao fim do estudo.

3. Acampamentos

Os acampamentos são de uso comum. Portanto, todos os usuários são responsáveis por sua manutenção. Antes de entrar em uma grade ou módulo do PPBio, cheque com o gerente de logística ou com os usuários mais frequentes, quais são os equipamentos e materiais disponíveis no acampamento, e quais precisam ser levados.

É proibido retirar materiais ou equipamentos do acampamento ou impor outro uso a eles. Caso observe ausência, alteração no uso ou deterioração dos equipamentos e materiais durante a sua estadia no acampamento, o gerente de logística do PPBio deve ser informado.

A área de entorno do acampamento deve ser mantida em boas condições. Lembre-se que você está em uma área de pesquisa, e qualquer alteração ambiental pode afetar os resultados da pesquisa. Assim, **não é permitido cortar árvores, caçar ou pescar, ou desbastar a vegetação que esteja a mais de 5 m do acampamento.**

O **lixo inorgânico** produzido durante a excursão deve ser todo recolhido e trazido para fora da grade ou módulo, e depositado em local adequado. Não é permitido enterrar ou queimar

lixo inorgânico dentro das grades ou módulos. O **lixo orgânico** deve ser enterrado em um local fixo a 10 m do acampamento. **É expressamente proibido jogar restos orgânicos nos igarapés.** Também não se deve jogar lixo orgânico sem enterrar, pois isto atrai animais para o acampamento, o que pode ser prejudicial aos próprios usuários.

Caso tenha a necessidade de deixar alguma coisa no acampamento, comunique ao gerente de logística imediatamente após a volta do campo.

O **banho** deve ser tomado nos chuveiros ou áreas específicas para banho instaladas no acampamento. Caso os chuveiros estejam com problemas e seja necessário usar algum igarapé, deve-se usar um balde, o mínimo de sabão e tomar o banho “de cuia” pelo menos 10 m de distância do igarapé. Sugere-se que o sabão para banho seja sempre biodegradável. É permitido “tomar banho” no igarapé sem produtos químicos (sabão, sabonete, xampu, detergente), mas deve-se evitar entrar em locais sujeitos a erosão das margens e lugares com substrato não consolidado (p. ex. bancos de folhas, grande quantidade de sedimento lodoso)

É proibido lavar roupas ou utensílios culinários diretamente nos igarapés, pois os produtos químicos usados para isso contaminam as águas e vão afetar todas as pesquisas das áreas aquáticas. Use baldes para lavar as roupas, descartando a água com sabão diretamente no solo, longe do igarapé.

O PESQUISADOR É RESPONSÁVEL PELO COMPORTAMENTO DA SUA EQUIPE. Informe a todas as pessoas que estão sob a sua responsabilidade sobre as regras de comportamento em áreas de pesquisa permanente. *A integração das pesquisas a curto e longo prazo depende da atitude correta de cada um.* Caso tenha alguma dúvida ou sugestão entre em contato com ppbio@inpa.gov.br.

Anexo 3 Versão preliminar dos metadados

AS INSTRUÇÕES PARA PREENCHIMENTO ESTÃO EM CAIXA ALTA

Os itens em itálico estão preenchidos como exemplo

Título:
Responsável (is):
<p>1. (RESPONSÁVEL) nome: e-mail: link para o lattes:</p> <p>2. (CO-RESPONSÁVEL, SE HOVER) nome: e-mail: link para o lattes:</p> <p>3. (AJUDANTES DE CAMPO ETC..)</p>
Endereço:
<i>Instituto Nacional de Pesquisas da Amazônia – INPA, Coordenação de Pesquisas em Ecologia – CPEC, Núcleo de Pesquisas de Roraima - NPRR, Rua Coronel Pinto 315 - Centro, 69301-150 Boa Vista – Roraima - Brasil</i>
Telefone:
<i>55 99 9999 9999 / 8888 8888</i>
Financiamento:
COLOCAR NOME E O NUMERO DO PROJETO BOLSAS DE MESTRADO/DOUTORADO/PCI ETC.
Resumo:
Palavras-chave:
Licença e Direitos de Uso:
<p>Todos os dados do PPBio serão públicos em no máximo 1 ano após sua coleta, desde que creditados os responsáveis pela coleta em qualquer publicação que os utilizem. Recomendamos aos interessados em utilizar esses dados que entrem em contato com os responsáveis para discutir o interesse e possibilidade de co-autoria.</p> <p><i>Fulano de Tal (fulano.de.tal@email.com.br) foi o responsável pela coleta destes dados.</i></p>
Abrangência Geográfica:
<i>Os dados foram coletados nas parcelas permanentes na grade de savana do Campo Experimental Agua Boa (Embrapa-RR), localizada em Boa Vista, RR.</i>
CASO NÃO COLETE EM TODAS AS PARCELAS, ESPECIFIQUE QUAIS PARCELAS UTILIZOU.
Coordenadas Geográficas:
<p>Campo Experimental Água Boa (Embrapa-RR): Oeste: 736267,3973 UTM (região 20N) Leste: 741510,0093 UTM (região 20N) Norte: 295159,8497 UTM (região 20N) Sul: 292551,7735 UTM (região 20N)</p>
Abrangência Temporal:
<i>A coleta de dados foi realizada entre os dias 15 e 20 de dezembro de 2005.</i>
IMPORTANTE: NÃO SE ESQUEÇA DE ANOTAR A DATA DE COLETA E DE INCLUIR ESSA INFORMAÇÃO

NA TABELA DE DADOS

Informações sobre Métodos de Coleta:

Os artrópodes foram coletados na copa das árvores que atingiam, no máximo, 5 m de altura como adaptação do método do batimento, onde o ramo ou galho era inteiramente fechado por uma rede entomológica antes da coleta (batimento fechado). Foram tomadas as medidas da árvore para o cálculo do volume da coleta e do volume total da copa, e foi adotado um escore (0-4) para estimar dados fenológicos (flores, frutos, folhas).

Instrumentos Utilizados:

Rede entomológica, trena, etc...

Informações sobre a tabela de atributos:

ATRIBUTOS SÃO COLUNAS DA TABELA DE DADOS. OS ATRIBUTOS GERAIS SÃO OBRIGATÓRIOS. CASO NÃO SEJA POSSÍVEL ESPECIFICAR TODOS OS NÍVEIS DE UNIDADE AMOSTRAL LISTADOS PARA OS SEUS DADOS, PREENCHA COM "NA" O VALOR DO ATRIBUTO PARA CADA OBJETO (LINHA DA TABELA DE ATRIBUTOS).

ATRIBUTOS GERAIS

Nome do Atributo: MÓDULO ou GRADE

Definição: NOME DO MÓDULO ou GRADE ESPECIFICADO DE ACORDO COM O CÓDIGO ADOTADO PELO PPBIO

Nome do Atributo: TRILHA

Definição: NOME DA TRILHA ESPECIFICADO DE ACORDO COM O CÓDIGO ADOTADO PELO PPBIO

Nome do Atributo: PARCELA

Definição: NOME DA PARCELA ESPECIFICADO DE ACORDO COM O CÓDIGO ADOTADO PELO PPBIO

Nome do Atributo: SUB-PARCELA

Definição: NOME DO SEGMENTO DE PARCELA ESPECIFICADO DE ACORDO COM O CÓDIGO ADOTADO PELO PPBIO

Nome do Atributo: DATA

Definição: DATA DE COLETA

ATRIBUTOS ESPECÍFICOS (ESPECIFIQUE AQUI OS DADOS QUE VOCÊ COLETOU)

Nome do Atributo: Argila

Definição: Porcentagem de argila

Nome do Atributo: Silte

Definição: Porcentagem de silte

Nome do Atributo: Areia

Definição: Porcentagem de areia