



UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Geociências

JEAN CARLOS FERREIRA DOS SANTOS

A CIÊNCIA ABERTA E SUAS (RE)CONFIGURAÇÕES:
POLÍTICAS, INFRAESTRUTURAS E PRÁTICA CIENTÍFICA

CAMPINAS
2019

JEAN CARLOS FERREIRA DOS SANTOS

A CIÊNCIA ABERTA E SUAS (RE)CONFIGURAÇÕES: POLÍTICAS,
INFRAESTRUTURAS E PRÁTICA CIENTÍFICA

TESE APRESENTADA AO INSTITUTO DE
GEOCIÊNCIAS DA UNIVERSIDADE ESTADUAL DE
CAMPINAS PARA OBTENÇÃO DO TÍTULO DE DOUTOR
EM POLÍTICA CIENTÍFICA E TECNOLÓGICA.

ORIENTADOR: PROF. DR. MARKO SYNÉSIO ALVES MONTEIRO

ESTE EXEMPLAR CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELO ALUNO JEAN
CARLOS FERREIRA DOS SANTOS E ORIENTADA
PELO PROF. DR. MARKO SYNÉSIO ALVES
MONTEIRO.

CAMPINAS

2019

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.

ORCID: <https://orcid.org/0000-0001-8354-9127>

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Geociências
Marta dos Santos - CRB 8/5892

Sa59c Santos, Jean Carlos Ferreira, 1988-
A ciência aberta e suas (re)configurações : políticas, infraestruturas e prática científica / Jean Carlos Ferreira dos Santos. – Campinas, SP : [s.n.], 2019.

Orientador: Marko Synésio Alves Monteiro.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Geociências.

1. Ciência Aberta (Movimento). 2. Neurociência. 3. Comunicação na ciência. 4. Infraestrutura de informação. 5. Tecnologia digital. I. Monteiro, Marko Synésio Alves, 1975-. II. Universidade Estadual de Campinas. Instituto de Geociências. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Open science and its (re)configurations : : policies, infrastructures and scientific practice

Palavras-chave em inglês:

Open Science (Movement)

Neuroscience

Science communication

Information infrastructure

Digital technology

Área de concentração: Política Científica e Tecnológica

Títuloção: Doutor em Política Científica e Tecnológica

Banca examinadora:

Marko Synésio Alves Monteiro [Orientador]

Paula Xavier dos Santos

Sarita Albagli

Germana Fernandes Barata

Diego Jair Vicentin

Data de defesa: 22-02-2019

Programa de Pós-Graduação: Política Científica e Tecnológica



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE GEOCIÊNCIAS**

AUTOR: Jean Carlos Ferreira dos Santos

**A CIÊNCIA ABERTA E SUAS (RE)CONFIGURAÇÕES:
POLÍTICAS, INFRAESTRUTURAS E PRÁTICA CIENTÍFICA**

ORIENTADOR: Prof. Dr. Marko Synésio Alves Monteiro

Aprovado em: 22 / 02 / 2019

EXAMINADORES:

Prof. Dr. Marko Synésio Alves Monteiro - Presidente

Profa. Dra. Sarita Albagli

Profa. Dra. Paula Xavier dos Santos

Prof. Dr. Diego Jair Vicentin

Profa. Dra. Germana Fernandes Barata

**A Ata de defesa com as respectivas assinaturas dos membros, encontra-se disponível no
SIGA - Sistema de Fluxo de Tese e na Secretaria de Pós-graduação do IG.**

Campinas, 02 de abril de 2019.

Dedico à equipe do NeuroMat

AGRADECIMENTO

Em primeiro lugar agradeço aos meus pais, Maria Neide Ferreira dos Santos e Edwaldo Ferreira dos Santos. Não fosse o carinho, o incentivo e a força dos meus pais quase nada na minha vida seria possível.

À Coordenação de Pessoal de Nível Superior (CAPES), pela bolsa de doutorado concedida no período de 2014 a 2015, fundamental para a minha permanência na pós-graduação e para o desenvolvimento da pesquisa de doutorado.

A todos os profissionais da Universidade Estadual de Campinas, pelo seu comprometimento com a ciência, a educação pública, gratuita e de qualidade.

Aos docentes e funcionários do Departamento de Política Científica e Tecnológica e do Instituto de Geociências.

Aos funcionários e ex-funcionários da Secretaria de Pós-Graduação do IG, em especial à Val e à Gorete.

A toda a equipe do NeuroMat, por tornar este trabalho possível. Obrigado pela disponibilidade, colaboração e acolhida durante o período de realização do trabalho de campo. Em especial, agradeço ao João, Evandro e Célio pelos e-mails respondidos, agendamentos, reuniões. Agradeço ao Prof. Galves, pela gentileza e disponibilidade.

Ao meu orientador Marko Monteiro, pela parceria, compreensão, paciência e ensinamentos ao longo do mestrado e do doutorado.

Aos membros da banca examinadora: Sarita Algabli, Paula Xavier, Germana Barata e Diego Vicentin. Obrigado pela atenção e valiosas contribuições à melhoria do meu trabalho.

Aos membros suplentes da banca: Daniela Manica, Marco Antônio Almeida e Henrique Parra.

Ao Prof. Marcelo Buzato pelos comentários e sugestões no exame de qualificação.

Aos amigos de todas as horas: Ana Paula, Vinicius, Daniela, Diego Canabarro e Nathália. Sou muito grato pela amizade de vocês. Obrigado pelo ombro amigo, risadas, cafés, viagens. Agradeço imensamente à Ana e ao Vinicius, pela leitura e sugestões fundamentais para o fechamento da tese.

À Ana Carolina e ao Uelinton, pela amizade valiosa que tornou a minha vida em São Paulo bastante agradável.

À querida Mônica Frigeri, pelas ideias, planos e risadas compartilhadas.

À querida Ednéia, pela amizade, suporte e coautorias.

Ao grande amigo Gustavo Grandini, pelos diálogos e sessões de desabafo.

A todos os companheiros e amigos de todas as horas do NIC.br: Salete, Paula, Carlinhos, Juca, Jaque, Marcelo, Gil, Chico, Cordeiro, Bruno e Vera.

Aos tripulantes do Submarino, por me religarem com a escrita e com a literatura.

Às minhas irmãs, Jessika e Anna Paula, pela torcida e carinho de sempre.

À minha sobrinha Lívia, pela luz e felicidade que trouxe para as nossas vidas.

O meu muito obrigado e eterna gratidão a todos os familiares e amigos que, cada um à sua maneira, de longe ou de perto, me incentivaram e me ofereceram alento nos momentos mais difíceis ao longo desses anos de doutorado. Obrigado a todos!

Abacateiro
Acataremos teu ato
Nós também somos do mato
Como o pato e o leão
Aguardaremos
Brincaremos no regato
Até que nos tragam frutos
Teu amor, teu coração

Abacateiro
Teu recolhimento é justamente
O significado
Da palavra temporão
Enquanto o tempo
Não trazer teu abacate
Amanhecerá tomate
E anoitecerá mamão

Abacateiro
Sabes ao que estou me referindo
Porque todo tamarindo tem
O seu agosto azedo
Cedo, antes que o janeiro
Doce manga venha ser também

Abacateiro
Serás meu parceiro solitário
Nesse itinerário
Da leveza pelo ar
Abacateiro
Saiba que na refazenda
Tu me ensina a fazer renda
Que eu te ensino a namorar

Refazendo tudo
Refazenda
Refazenda toda
Guariroba

(Refazenda, Gilberto Gil)

RESUMO

Esta tese tem por objeto o tema da Ciência Aberta. As práticas de Ciência Aberta podem envolver o acesso aberto e gratuito a dados de pesquisa, artigos e ferramentas de pesquisa e estão associadas à transparência nos processos de pesquisa e à reprodutibilidade experimental; à democratização do acesso ao conhecimento científico; e, frequentemente, ao desenvolvimento e avanço dos processos de descoberta e inovação por meio do compartilhamento aberto de ferramentas e resultados de pesquisa. O objetivo geral da tese é analisar o movimento recente de Ciência Aberta em sua relação com as infraestruturas de informação e com as práticas científicas. Busca-se também compreender de que maneira as práticas institucionalizadas são reconfiguradas em contextos de Ciência Aberta ou até que ponto se desenvolvem novas práticas de pesquisa, novas dinâmicas e formas de organização do trabalho científico no contexto da Ciência Aberta. Para tanto, empreendeu-se uma pesquisa do tipo etnográfica como forma de identificar o objeto inserido nas práticas do Centro de Pesquisa, Inovação e Difusão em Neuromatemática da Universidade de São Paulo (NeuroMat), um centro multidisciplinar que atua no desenvolvimento de modelos matemáticos para a modelagem de dados experimentais produzidos pela pesquisa em neurociência. Durante o trabalho de campo desenvolvido no período de 2015-2017, buscou-se compreender a produção e as configurações da Ciência Aberta nesse espaço, rastreando as dinâmicas e os significados produzidos em torno das práticas de compartilhamento e de ferramentas abertas desenvolvidas e implementadas pelo centro, no contexto do seu envolvimento com iniciativas de Ciência Aberta. As escolhas teóricas e metodológicas da tese inserem-se no campo interdisciplinar dos Estudos Sociais da Ciência e da Tecnologia, mais especificamente na corrente teórica que aborda a ciência no seu aspecto localizado e socialmente construído a partir da noção de “culturas epistêmicas”. Recorreu-se ainda aos Estudos de Infraestrutura, corrente dos ESCT que objetiva analisar as infraestruturas de informação como uma rede de relações entre elementos materiais e imateriais que sustentam o trabalho científico. Conclui-se que a Ciência Aberta no NeuroMat adquire diferentes significados e integra-se a diferentes processos. Percebe-se que as ferramentas abertas utilizadas em processos de compartilhamento dos resultados de pesquisa e colaboração têm um significado importante na implementação das pesquisas desenvolvidas no centro. Nesse aspecto, a tese defende que as escolhas em relação à Ciência Aberta nesse local inserem-se em um processo mais amplo e complexo de institucionalização e reconhecimento do NeuroMat em sua área de atuação, à medida que as práticas de compartilhamento de dados e as ferramentas abertas são colocadas como viabilizadoras da produção de conhecimento no domínio da pesquisa em neurociência.

Palavras-chave: Ciência Aberta (Movimento). Neurociências. Comunicação na ciência. Infraestrutura de informação. Tecnologia digital.

ABSTRACT

This thesis has the aim of studying Open Science. The Open Science practices may involve the open and free access to research data, scientific papers and research tools, and are linked to the accountability of research processes, as well as to the experimental replication; the democratization of the access to scientific knowledge; and often to the development and advance of discovery and innovation processes, fostered by the open sharing of research resources and results. The goal of this study is to analyze the recent open science movement in its linkages with information infrastructures and with scientific practices. There is also an aim of understanding in which ways the institutionalized practices are reconfigured in contexts of open science or even how new research practices are developed, new dynamics and ways to organize the scientific work in the context of open science practices. Therefore, this study put on an ethnographic-based research as means of identifying the object embedded in the practices of the Research, Innovation and Dissemination Center for Neuromathematics (NeuroMat), a multidisciplinary center that creates mathematical models to represent and help to understand the nervous system phenomena based on the analysis of experimental data. Within the fieldwork developed throughout two years, the goal was to understand the production and configurations of open science in this field, tracking the dynamics and meanings produced around the sharing practices and the open tools developed and implemented by the Center in the context of its involvement with open science initiatives. The theoretical and methodological choices made for this thesis are bound to the interdisciplinary field of Science and Technology Studies (STS), especially the trend that frames science in its own local context characteristic and also its nature of being socially constructed, through the notion of “epistemic cultures”. Additionally, the discussions refer to the Infrastructure Studies, an STS trend which analyzes information infrastructures as a network of interrelated material and immaterial elements that support the scientific work. The main conclusion is that NeuroMat’s open science gains different meanings and integrates itself to different processes. The open tools used for the sharing of research results and in collaboration processes have an important role in the implementation of research developed within the Center. In this aspect, the thesis argues that the choices regarding open science practices are bound to a broader and more complex process of the institutionalization and recognition of NeuroMat in its own field, to the extent that data sharing practices and open tools are framed as enablers of the knowledge production in the neuroscience research domain.

Keywords: Open Science (Movement). Neuroscience. Science communication. Information Infrastructure. Digital technology.

LISTA DE FIGURAS

Figura 1: Taxonomia da Ciência Aberta.	39
Figura 2: Apresentação da linha de pesquisa em neuromatemática durante workshop.	85
Figura 3: Técnicos instalam o supercomputador do NeuroMat.	87
Figura 4: Cartaz da IV Maratona de edição.	105
Figura 5: Animação ilustrativa do conceito matemático "Lei dos Grandes números".....	112
Figura 6: Nervos do plexo braquial no braço humano. Fonte: Wikimedia Commons.....	120
Figura 7: Visualização em 3D dos disparos neurais. Fonte: NeuroMat.....	136
Figura 8: Apresentação do banco de dados durante o II Encontro de Ciência Aberta.....	145
Figura 9: Pesquisador exhibe o modelo que simula o <i>E. Elegans</i>	156
Figura 10: Componentes desenvolvidos para o <i>OpenWorm</i>	157

LISTA DE QUADROS

Quadro 1: Correntes de Ciência Aberta.	40
Quadro 2: Eventos.....	73
Quadro 3: Pesquisadores Entrevistados.	75
Quadro 4: Equipe Técnica do NeuroMat.	76

LISTA DE ABREVIATURAS E SIGLAS

ABRAÇO	Ação NeuroMat para a Lesão do Plexo Braquial
AMPARO	Rede de Apoio NeuroMat a Amigos e Pessoas com Doença de Parkinson
ANT	Teoria Ator-Rede
CCSL	Centro de Competência em Software Livre da Universidade de São Paulo
CEPID	Centro de Pesquisa, Inovação e Difusão
C&T	Ciência e Tecnologia
EEG	Eletroencefalografia
EMG	Eletromiografia
EOSC	European Open Science Cloud
ESCT	Estudos Sociais da Ciência e da Tecnologia
FAIR	Findable, Accessible, Interoperable and Reusable
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FLOSS	Free and Open Source Software
fMRI	Imagem por ressonância magnética funcional
GOFAIR	Global Open FAIR Implementation Networks
GRF	German Research Foundation
ICe/UFRN	Instituto do Cérebro da Universidade Federal do Rio Grande do Norte
IME/USP	Instituto de Matemática e Estatística da Universidade de São Paulo
INCF	International Neuroinformatics Coordinating Facility
INDC/UFRJ	Instituto de Neurologia Deolindo Couto da Universidade Federal do Rio de Janeiro
MEG	Magnetoencefalografia
NES	Neuroscience Experiments System
NIF	Neuroscience Information Framework
NIH	National Institutes of Health
NPNR	Núcleo de Pesquisa em Neurociências e Reabilitação
NSF	National Science Foundation
NUMEC	Núcleo de Apoio à Pesquisa em Modelagem Estocástica e Complexidade
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
PLOS	Public Library of Science
SciELO	Scientific Electronic Library Online
SGC	Structural Genomics Consortium

TMS	Estimulação magnética transcraniana
UNICAMP	Universidade Estadual de Campinas
USP	Universidade de São Paulo

SUMÁRIO

INTRODUÇÃO	17
1 CIÊNCIA ABERTA: APROXIMAÇÕES CONCEITUAIS	29
1.1 O <i>ethos</i> científico e a privatização do conhecimento	29
1.2 Ciência, código aberto e compartilhamento	34
1.3 Ciência Aberta: múltiplas abordagens	37
1.4 Abertura de dados de pesquisa: políticas e infraestruturas	44
2 ESTUDANDO A CIÊNCIA ABERTA: ABORDAGEM TEÓRICA E METODOLÓGICA.....	53
2.1 Culturas epistêmicas e prática científica.....	53
2.2 Estudos de Infraestruturas.....	58
2.2.1 Infraestruturas de informação e as práticas científicas emergentes	61
2.3 Etnografia da Ciência e da Tecnologia	66
2.4 Detalhamento da Pesquisa de Campo	70
2.4.1 <i>Entrevistas</i>	74
3 A CIÊNCIA ABERTA NO CENTRO DE PESQUISA, INOVAÇÃO E DIFUSÃO EM NEUROMATEMÁTICA	78
3.1 Adentrando o campo	78
3.2 Neuromatemática: a nova ciência do cérebro	83
3.3 Estrutura do NeuroMat	90
3.4 Eixos de atuação do NeuroMat	91
3.4.1 <i>Eixo de Transferência de Conhecimento e Tecnologia</i>	92
3.4.2 <i>Eixo de Educação e Difusão Científica</i>	93
3.5 Configurando a Ciência Aberta no NeuroMat	94
3.6 Licenças, recursos informacionais e profissionais da Ciência Aberta.....	99
3.7 Abrindo a ciência na Wikipédia.....	103
3.8 Ferramentas abertas e a organização da neuromatemática	106
4 PRÁTICAS DE COMPARTILHAMENTO DE DADOS: FRONTEIRAS DISCIPLINARES E CULTURAS EPISTÊMICAS.....	114
4.1 Neurocientistas e práticas de compartilhamento	114
4.1.1 <i>Neurocientistas teóricos</i>	118
4.1.2 <i>Neurocientistas experimentalistas</i>	123
4.2 Modeladores e práticas de compartilhamento	129
5 DINÂMICAS EM TORNO DE UM SOFTWARE DE GERENCIAMENTO E COMPARTILHAMENTO DE DADOS.....	143
5.1 Neuroscience Experiments System.....	143
5.2 Padrões e infraestruturas de compartilhamento	147

5.4 O software como produto da atividade acadêmica: dilemas de abertura e fechamento	160
CONSIDERAÇÕES FINAIS.....	169
REFERÊNCIAS BIBLIOGRÁFICAS	179
APÊNDICE A – Roteiro de questões e tópicos utilizado nas entrevistas	195
APÊNDICE B – Fontes consultadas disponíveis online	196

INTRODUÇÃO

Esta tese tem como objeto de estudo a Ciência Aberta. O objetivo da tese é compreender a Ciência Aberta a partir das práticas científicas e infraestruturas de informação. A pesquisa utiliza-se do enfoque teórico e analítico dos Estudos Sociais da Ciência e da Tecnologia (ESCT) e se baseia em uma etnografia desenvolvida em um centro de pesquisa da área de modelagem de dados experimentais em neurociência. Antes do detalhamento dos objetivos e questões que nortearam a tese, apresento a seguir alguns antecedentes que influenciaram o seu desenvolvimento.

Antecedentes

A tese aqui apresentada é, de muitas maneiras, resultado de uma trajetória acadêmica iniciada na graduação e que ganhou contornos mais concretos ao longo da pós-graduação. Na pós-graduação pude direcionar meus interesses de pesquisa para refletir, especificamente, sobre os processos de comunicação científica no contexto das Tecnologias de Informação e Comunicação (TIC). No contexto de transformações múltiplas e amplas associadas às TIC, no qual as comunidades científicas depararam-se com diferentes possibilidades oferecidas pelo ambiente aberto, colaborativo e transfronteiriço da Internet, a formação de uma agenda cada vez mais voltada para o “*openness*” ou “abertura” na comunicação científica era um tema que me instigava.

Por essa razão, durante o mestrado em Política Científica e Tecnológica, realizado no período de 2012 a 2014, realizei uma análise qualitativa sobre a percepção dos pesquisadores de uma área específica do conhecimento acerca do movimento pelo Acesso Aberto, buscando com isso compreender de que maneira esse movimento impactava nas práticas de publicação científica desses pesquisadores¹.

Pretendendo dar continuidade e aprofundamento a essa compreensão do movimento pelo Acesso Aberto, ingressei no doutorado em Política Científica e Tecnológica em 2014 com um projeto que propunha descrever o desenvolvimento desse movimento no Brasil do ponto de vista das políticas públicas, e analisar a percepção de pesquisadores em diferentes áreas do conhecimento, ampliando, assim, o escopo analítico e metodológico da

1 Cf. Santos (2014).

pesquisa desenvolvida no mestrado. O aprofundamento bibliográfico e principalmente alguns acontecimentos relacionados ao tema da abertura na ciência, me colocaram diante de um novo objeto e de novos problemas que me instigaram a rever a proposta de pesquisa de doutorado. Descrevo, a seguir, alguns desses acontecimentos com o intuito de ilustrar como as minhas principais percepções sobre o objeto foram se transformando e as questões de pesquisa mudaram até desaguarem no trabalho que o leitor tem em mãos.

Alguns dos fatos marcantes para esta tese ocorreram ainda no mestrado. Em uma etapa da pesquisa de mestrado, na qual a minha leitura e compreensão do movimento pelo Acesso Aberto ainda era, de certo modo, limitada, atentei-me para uma recorrência do termo “*Open Science*”, ou Ciência Aberta, na literatura científica, documentos governamentais, em agências de fomento e entre ativistas envolvidos com o movimento pelo Acesso Aberto.

Entre esses marcos bibliográficos, refiro-me sobretudo à leitura do relatório da *Royal Society* da Inglaterra “*Science as an open enterprise*”². Lançado em 2012, o relatório, que repercutiu amplamente na comunidade envolvida com o movimento pelo Acesso Aberto e comunidade científica em geral, põe ênfase nos desafios criados pelo “dilúvio de dados”, especialmente a abertura dos dados e todos os processos e tecnologias subjacentes à sua produção e análise e como a abertura dos dados (não apenas dos artigos científicos) pode garantir, na atualidade, o caráter aberto da ciência. Com ênfase nos dados e nos desafios subjacentes à sua abertura, o relatório aponta uma série de orientações para que os cientistas possam alcançar a abertura dos resultados de pesquisa.

Outro acontecimento foi o lançamento da organização norte-americana sem fins lucrativos “*Center for Open Science*” em 2013. A organização foi fundada com o propósito de dar suporte à Ciência Aberta nos EUA e principalmente para o estímulo à reprodutibilidade da pesquisa acadêmica, disponibilizando uma infraestrutura tecnológica para armazenamento e compartilhamento de dados de pesquisa³.

Já em 2014, uma série de eventos realizados na Universidade Estadual de Campinas (UNICAMP) dava início à aproximação da universidade com a rede de laboratórios “*Structural Genomics Consortium*”, que se define como uma organização de pesquisa que utiliza a Ciência Aberta na pesquisa e desenvolvimento de fármacos, abrindo todos os resultados na Internet.

2 Cf. Royal Society (2012).

3 The Center For Open Science. Disponível em: <<https://cos.io/>>. Acesso em: 24 jan. 2019.

Concorrem também para a escolha do tema as recorrentes repercussões no meio acadêmico – e mesmo na mídia tradicional – de políticas mandatórias e iniciativas de abertura dos resultados de pesquisa implementadas pelos principais órgãos de apoio à pesquisa dos EUA e da Europa, contribuindo significativamente para a discussão em torno da Ciência Aberta⁴.

O início da incursão para entender o que era propriamente Ciência Aberta se deu por meio da análise de uma série de documentos e sites de organizações de apoio à pesquisa. Em um primeiro momento, o termo soou como uma espécie de evolução ou de virada no movimento pelo Acesso Aberto, como se representasse uma passagem do nível das publicações para abranger outros produtos e processos da atividade de pesquisa. Essa leitura me permitiu formar uma compreensão inicial de Ciência Aberta como uma espécie de evolução, no sentido em que reivindicava que o acesso aberto aos artigos científicos tão somente não era suficiente.

Um aspecto importante dessa virada era que a “nova” abordagem parecia colocar maior ênfase na transparência, na reprodutibilidade dos processos subjacentes à produção dos resultados de pesquisa e, sobretudo, nos dados de pesquisa. Mas a constatação de que havia uma complexidade de abordagens e de atores preocupados com a abertura dos resultados de pesquisa e, principalmente, de que muitas das ideias não eram tão novas assim. Percebi que a Ciência Aberta não podia ser tratada meramente como uma evolução do movimento pelo Acesso Aberto, mas talvez como um modo de nomear práticas e modos diversos de abrir os processos e resultados do trabalho de pesquisa científica no contexto das TIC, configurando novas formas de produção e comunicação do conhecimento científico.

Durante esse tateamento do tema, interessei-me por adentrar um aspecto mais específico do movimento de Ciência Aberta relacionado à propriedade intelectual nas políticas e iniciativas de abertura. Tal motivação surgiu a partir de uma avaliação de que os organismos de financiamento da ciência – que, historicamente, foram e são responsáveis por estimular a cultura empreendedora voltada para a produção de pesquisas com resultados comercializáveis, o que frequentemente significou (e significa) o “fechamento” por meio do

4 Entre esses acontecimentos, nas esferas das políticas de C&T estão as principais políticas dos Institutos Nacionais de Saúde (NIH – National Institutes of Health) e da Fundação Nacional da Ciência (NSF – National Science Foundation) voltadas cada vez mais para a abertura de dados e ferramentas de pesquisa. Destacam-se também as diretrizes estabelecidas pela Comissão Europeia – como parte de um conjunto de medidas para acesso à informação científica – que visam a melhorar e aumentar o acesso e a reutilização dos dados de pesquisa desenvolvidos no âmbito do projeto Horizonte 2020. Destaco ainda a atuação cada vez maior de organizações privadas na promoção do acesso aberto, principalmente aos dados de pesquisa, tais como a Wellcome Trust e a Fundação Bill e Melinda Gates.

licenciamento de tecnologias através do patenteamento ou transferência de conhecimento para a indústria – vinham crescentemente apoiando formas de disseminação aberta dos resultados de pesquisa. Ou seja, apesar dos incentivos e de várias pressões em direção ao desenvolvimento de práticas que tendem a restringir o acesso aos resultados de pesquisa, atualmente observa-se uma tendência à abertura dos processos e produtos da pesquisa científica de maneira ampla – em um *ethos* aparentemente oposto (LEVIN et al., 2016; CAULFIELD et al, 2012).

A partir dessas premissas, inicialmente interessei-me em analisar o desenvolvimento da Ciência Aberta como contraponto a um modelo de organização da pesquisa que valoriza o patenteamento e outras formas de propriedade intelectual que podem ser vistas como “fechamento” da ciência e – sem perder de vista a percepção dos pesquisadores – pretendia também mapear a compreensão de diferentes atores sobre esses dois modelos. Foi a partir da preocupação com a percepção dos atores que passei a mapear alguns espaços acadêmicos que de algum modo estavam inteirados dos debates e práticas envolvendo Ciência Aberta.

Ao me lançar em uma busca por esses espaços, obtive contato com pesquisadores de diferentes áreas e com formas variadas de produção de conhecimento, o que ajudou a reformular as questões do projeto de tese e, sobretudo, lançar um outro olhar para a Ciência Aberta e ampliar o foco da pesquisa – até então voltado para os aspectos da propriedade intelectual. Os espaços visitados durante essa fase exploratória do campo foram o Clube de Biologia Sintética da Universidade de São Paulo, o *Structural Genomic Consortium* (SGC), o Centro Nacional de Energia e Materiais e o Centro de Pesquisa, Inovação e Difusão em Neuromatemática da Universidade de São Paulo.

O que me levou até o Grupo de Biologia Sintética da Universidade de São Paulo foram algumas postagens na comunidade do grupo no Facebook sobre Ciência Aberta⁵. O grupo apresentava um profundo relacionamento com o compartilhamento e reutilização de dados, apresentando uma abordagem a partir do movimento DIYbio – “*Do it yourself biology*” ou “Faça você mesmo biologia” – e do *biohacking*⁶. As práticas desse movimento

5 O Grupo de Biologia Sintética da USP é formado por estudantes de graduação e pós-graduação que realizam atividades relacionadas ao desenvolvimento de soluções em bioengenharia, utilizando a abordagem da biologia sintética, tais como a implementação de novos sistemas gênicos, modificação de sistemas biológicos naturais e padronização de códigos. Cf. Site do grupo: <<http://synbiobrasil.org/>>. Acesso em: 23 set. 2018.

6 Conversei com alguns participantes do grupo e acompanhei uma atividade presencial durante a Campus Party Brasil em 2014.

envolviam o acesso a repositórios de dados abertos de sequenciamento genético, hardware aberto e software livre, e onde frequentemente os membros traziam a ideia de “subversão” da estrutura acadêmica tradicional por meio de suas práticas ou de produção de uma ciência “independente”.

No decorrer das observações nas listas de e-mail, fóruns e participação em uma atividade do grupo, percebi que os elementos que remetiam à Ciência Aberta se mostravam muito fugidios e me conduziam para temáticas como *biohacking* e cultura *maker*, que embora fossem temas instigantes e que de algum modo se relacionavam ao tema central da pesquisa, considerei que mudariam drasticamente o rumo da tese.

Ainda, naquele ano, participei da conferência “*Chemical Probe-based Open Science: Uncovering New Human and Plant Biology*”, realizada na Faculdade de Ciências Médicas da UNICAMP⁷. A conferência deu início a uma série de atividades que faziam parte do acordo de cooperação da UNICAMP com a organização sem fins lucrativos *Structural Genomic Consortium*⁸. A organização reúne laboratórios de pesquisa em proteômica e produção de medicamentos e utiliza um modelo de organização da pesquisa baseado em Ciência Aberta, publicando todos os artigos e dados de pesquisa em acesso aberto na Internet.

Logo em seguida, por sugestão de uma professora do Programa de Pós-graduação em Política Científica e Tecnológica, entrei em contato com pesquisadores do Centro Nacional de Pesquisa em Energias e Materiais (CNPEM), especializado em desenvolvimento de pesquisa de ponta e disponibilização de diversas instalações de pesquisa abertas à comunidade científica. A tentativa de aproximação com o CNPEM era motivada sobretudo pela constatação de que os laboratórios nacionais que compõe o CNPEM geram grandes quantidades de dados, mostrando-se um local interessante no qual seria possível investigar as práticas de compartilhamento de uma grande infraestrutura de pesquisa utilizada por diferentes tipos de atores.

Durante visitas às instalações do CNPEM, dois pesquisadores ligados a um dos laboratórios nacionais trouxeram alguns pontos de vista importantes que passei a considerar a partir dali. O primeiro foi o panorama sobre a importância dos repositórios públicos de estruturas biomoleculares para a pesquisa em genômica e bioquímica. Um aspecto importante trazido pelos pesquisadores foi o papel histórico desses repositórios no processo de

7 Evento realizado nos dias 28 e 29 de abril de 2014.

8 Cf. Sugimoto (2014).

institucionalização de algumas áreas de pesquisa⁹, afirmando que essas infraestruturas estão no “coração” dessas áreas e que se houver “patenteamento” de determinados resultados, algumas pesquisas “morrem”.

Outro ponto importante do diálogo com os pesquisadores do CNPEM (e que conectava com o evento do SGC citado anteriormente) surgiu quando comentei sobre a minha participação no evento do “*Structural Genomics Consortium*” na UNICAMP. Nesse momento, os pesquisadores relataram que o laboratório ao qual estavam vinculados havia sido procurado por representantes do SGC, que apresentaram a proposta de integrar o CNPEM à rede de laboratórios do Consórcio. Os pesquisadores relataram que havia interesse do Consórcio em ter um parceiro no hemisfério sul, entretanto, não houve interesse do laboratório do CNPEM integrar-se ao SGC naquele momento. Os informantes justificaram a recusa em participar do Consórcio afirmando que “a Ciência Aberta do SGC é outra” e que o Consórcio “não é tão aberto assim” como se apresentava, pois caso os pesquisadores participantes do Consórcio chegassem a resultados “revolucionários” com potencial de comercialização pela indústria de medicamentos, os resultados poderiam ser patenteados pelas empresas envolvidas.

Essa interpretação dos pesquisadores do CNPEM de que o SGC fazia “outra” Ciência Aberta, diferente do que eles imaginavam ou compreendiam como “fazer” Ciência Aberta, instigou-me a olhar mais de perto a proposta do SGC. Ao começar a acompanhar algumas atividades do SGC na UNICAMP, de fato me pareceu que a Ciência Aberta que o projeto propunha não se configurava como algo contrário à propriedade intelectual, como me levou a crer em um primeiro momento. Isso ficou mais evidente quando acompanhei um workshop¹⁰ do SGC para pesquisadores da UNICAMP realizado no Instituto de Química.

Durante o workshop, o então Diretor Presidente do Consórcio, o pesquisador canadense Aled Edwards, explicou que o SGC comporta uma fase chamada de “pré-competitiva”, na qual os pesquisadores disponibilizam todos os resultados em acesso aberto na Internet, sem se preocupar com o patenteamento. Segundo ele, isso ocorria porque a publicação aberta dos resultados se mostrava fundamental na descoberta de novas moléculas e desenvolvimento de medicamentos. A utilização de um modelo de Ciência Aberta no Consórcio tinha o papel de acelerar o processo ou encurtar o tempo entre a pesquisa e

9 Um dos repositórios citados foi o RCSB Protein Data Bank para dados de cristalografia, ressonância magnética e microscopia. Disponível em: <<https://www.rcsb.org>>. Acesso em: 23 set. 2018.

10 Evento realizado em 12 de março de 2015.

descoberta de um novo fármaco e sua comercialização que, segundo o Diretor do Consórcio, é um tempo longo diante das demandas por novos tratamentos de doenças. Na fase seguinte das pesquisas, ocorre o desenvolvimento de inovações proprietárias onde entram as competências da indústria e comercialização – fase esta que pode haver patenteamento ou uso comercial dos resultados.

Tanto o caráter complementar entre Ciência Aberta e propriedade intelectual que parecia haver no modo de funcionamento do SGC, como a afirmação dos pesquisadores do CNPEM de que havia uma “outra” Ciência Aberta e de que os repositórios públicos tinham um papel vital para o desenvolvimento da área em que eles trabalhavam, ofuscaram a preocupação inicial com a propriedade intelectual, mostrando uma dimensão mais complexa que merecia atenção.

Foi o contato com o Centro de Pesquisa, Inovação e Difusão em Neuromatemática (NeuroMat) em 2015 que minhas preocupações de pesquisa começaram a tomar outros contornos. O NeuroMat é um dos Centros de Pesquisa, Inovação e Difusão (CEPID), financiados pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e está sediado no Instituto de Matemática e Estatística da Universidade de São Paulo (IME/USP).

O CEPID NeuroMat é um centro multidisciplinar que atua no desenvolvimento e aplicação de ferramentas do domínio da matemática, estatística e probabilística para a modelagem de dados produzidos em diferentes experimentos realizados no campo da neurociência. O primeiro contato com o NeuroMat ocorreu por meio de uma matéria, divulgada na revista Pesquisa FAPESP, sobre um software desenvolvido pelo centro para gerenciamento de dados experimentais¹¹. Conforme era sugerido na matéria mencionada e também no conteúdo do site¹² do centro, o NeuroMat apresentava explicitamente uma questão relativa à Ciência Aberta, envolvendo-se com compartilhamento de dados e ferramentas de software de código aberto.

Como os demais espaços que visitei, no NeuroMat, os elementos que remetiam à Ciência Aberta pareciam adquirir forma e significados particulares. Ali esses elementos remetiam ao desenvolvimento de infraestruturas computacionais para dar suporte ao compartilhamento de dados experimentais e colocava-se a Ciência Aberta como aspecto das atividades e dos objetivos científicos do NeuroMat. As infraestruturas tecnológicas, bem

11 Cf. Toledo (2014).

12 Cf. NeuroMat *Open Science*. Disponível em: <<https://neuromat.numec.prp.usp.br/pt-br/open-science>>.

como os dados de pesquisa, apareciam como centrais na estruturação das práticas. Naqueles contatos iniciais com os informantes desse centro, sugeriu-se, também, a existência alguns conflitos internos em relação à forma como os sujeitos vinculados ao centro interpretavam a Ciência Aberta e, principalmente, o significado do compartilhamento de dados de pesquisa.

Descortinou-se aí uma possibilidade de tratar o objeto a partir de uma abordagem cada vez mais atenta aos significados e às formas que a Ciência Aberta adquiria na prática científica a partir de um olhar para as configurações locais existentes no NeuroMat.

As observações iniciais no NeuroMat e toda a trajetória pelos diferentes sítios descritos acima suscitaram uma série de questões sobre a Ciência Aberta, que têm como ponto de referência a prática científica, a partir das quais conduzi a presente pesquisa: Como os cientistas interpretam e se apropriam dos princípios e ferramentas frequentemente associados à ideia de Ciência Aberta? Como se constitui a Ciência Aberta na prática científica? A introdução de uma abordagem de Ciência Aberta produz rearranjos e novos ordenamentos na prática científica? Em que medida as (novas) práticas da Ciência Aberta contrastam, complementam ou interagem com práticas de pesquisa institucionalizadas/sedimentadas ou com os objetivos de cada ator envolvido na atividade científica?

A forte presença das ferramentas computacionais no NeuroMat permite ainda uma questão: qual o papel dessas infraestruturas tecnológicas abertas na viabilização da Ciência Aberta e, portanto, como (re)configuram práticas e fluxos de informação no contexto da pesquisa acadêmica?

Objetivos

Considerando as questões colocadas acima, o objetivo geral da tese é:

- ✓ Analisar e compreender o movimento recente de Ciência Aberta em sua relação com as infraestruturas de informação e com as práticas científicas.

Os objetivos específicos da tese são:

- ✓ Compreender a produção e as configurações locais da Ciência Aberta no Centro de Pesquisa, Inovação e Difusão em Neuromatemática (NeuroMat);
- ✓ Identificar as fronteiras e as divisões culturais que se estabelecem em relação à Ciência Aberta nesse espaço;
- ✓ Analisar as possibilidades de novas práticas de circulação do conhecimento científico, bem como possíveis novos papéis e as formas de trabalho e práticas que se estabelecem ao redor da Ciência Aberta no centro de pesquisa analisado.

Ao perseguir esses objetivos, defendo que a Ciência Aberta adquire significado quando inserida em uma rede de relações, sendo apropriada de diferentes formas nos espaços das práticas científicas, o que complexifica a adoção e generalização dessa abordagem no âmbito das diferentes comunidades científicas. Os potenciais e impactos da abertura de resultados e processos científicos ainda não foram totalmente mensurados, entretanto pode-se pensar que a abordagem de Ciência Aberta reconfigura a prática científica em muitos aspectos ao “adentrar” o laboratório, principalmente inserindo novas ou ressignificando questões existentes sobre os instrumentos e resultados do trabalho de pesquisa, colaboração científica, reconhecimento acadêmico, autoria, entre outros aspectos.

Um aspecto importante da análise diz respeito à existência de práticas diferentes ou de culturas epistêmicas implicadas, isto é, mecanismos, arranjos, processos e sistemas que produzem e garantem a manutenção do conhecimento nos diferentes campos, e que fazem a Ciência Aberta ser significada de múltiplas formas, seja como uma continuidade de valores universais que orientam o funcionamento da ciência, seja como uma ferramenta particular que viabiliza o trabalho dos cientistas ou mesmo como um elemento estranho ao cotidiano de trabalho dos pesquisadores.

A produção de Ciência Aberta envolve também negociações em torno das infraestruturas tecnológicas utilizadas, dos níveis de abertura, dos tipos de autorizações de uso dos resultados, do que poderá ou não ser disponibilizado abertamente, etc. Nessa rede de negociações, as categorias “aberto” e “fechado”, por exemplo, são constantemente (re)significadas e negociadas, assim como as fronteiras entre essas categorias podem ser também porosas ou fluídas, podendo ocorrer, inclusive, uma hibridização entre ambas.

A metodologia utilizada para o desenvolvimento da pesquisa é qualitativa e exploratória e sustentou-se em um trabalho de campo, do tipo etnográfico, desenvolvido no NeuroMat ao longo de dois anos.

A investigação e análise resultantes da imersão no campo serão exploradas em 5 (cinco) capítulos, considerando a seguinte organização:

O **primeiro capítulo** objetiva fornecer elementos para a compreensão dos conceitos norteadores do movimento e abordagem de Ciência Aberta. O capítulo explora as relações do termo com as normas clássicas que estabelecem a ciência como uma atividade pública e aberta e contrasta com a privatização ou fechamento dos resultados de pesquisa, explorando alguns conflitos que marcam a circulação e a apropriação do conhecimento científico. Em seguida, o capítulo situa a discussão no contexto do Movimento do Software Livre e de Código Aberto, pontuando as principais influências desse movimento nos debates recentes de abertura na ciência. O capítulo aborda as diferentes correntes interpretativas que estruturam atualmente as iniciativas de Ciência Aberta e, em seguida, descreve a recente centralidade dos dados de pesquisa para essas discussões, enfocando as políticas e infraestruturas para compartilhamento e reutilização de dados.

O **segundo capítulo** apresenta a fundamentação teórica, os materiais e métodos que orientaram a pesquisa e as análises presentes na tese. O capítulo está centrado nos conceitos e formas de abordagem oferecidos pelos Estudos Sociais da Ciência e da Tecnologia (ESCT) voltados para uma abordagem centrada nas práticas científicas. Nesse aspecto, abordagem é influenciada pelos estudos de laboratório empreendidos pelos ESCT, utilizando-se de uma abordagem focada nas práticas sociais no espaço do laboratório, pondo foco nas culturas epistêmicas. Utiliza-se do referencial dos Estudos de Infraestrutura para pensar os recursos e ferramentas que sustentam o trabalho científico como uma rede de elementos relacionados. A infraestrutura, nesse aspecto, não se resume aos artefatos tecnológicos, mas envolve o relacionamento entre elementos materiais e imateriais que dão

suporte ao trabalho acadêmico. Discute a etnografia como recurso para adentrar o espaço dos cientistas, possibilitando uma sensibilidade para olhar a ciência em seu caráter processual e construído. O capítulo detalha ainda, em sua última parte, os instrumentos utilizados para coletar os dados que compõe o empírico da pesquisa.

O **terceiro capítulo** adentra a pesquisa de campo realizada no NeuroMat. Como forma de situar o leitor, inicialmente, o capítulo apresenta uma imagem do NeuroMat e da ciência ali produzida e aborda as principais questões científicas que orientam o centro de pesquisa, situando-as no campo da pesquisa sobre modelagem matemática e computacional e neurociência. O capítulo descreve a estrutura do NeuroMat e as atividades desenvolvidas. Em seguida, adentra as iniciativas de Ciência Aberta desenvolvidas, contextualizando-as em relação às atividades do centro.

Nos capítulos seguintes aprofundo os resultados da pesquisa empírica, estabelecendo alguns recortes.

O **quarto capítulo** centra-se na etnografia das práticas de pesquisa de neurocientistas e modeladores. Caracterizo as práticas da neurociência e as questões colocadas por esses pesquisadores em relação ao compartilhamento de dados de pesquisa. O capítulo situa a relação entre modelagem e abertura de resultados experimentais e como o compartilhamento se imbrica nas práticas da modelagem, evidenciando a centralidade das ferramentas abertas e do compartilhamento de dados na estruturação dessa área.

O **quinto capítulo** analisa as dinâmicas em torno do software de código aberto para gestão e compartilhamento de dados experimentais “*Neuroscience Experiments System*” (NES). Esse software é central para os objetivos e missão científica do NeuroMat de produzir modelos neuronais a partir da modelagem de dados produzidos pela área de neurociência, fazendo parte da estratégia de Ciência Aberta implementada pela equipe do centro. Tomando como referência as relações que se estabelecem em torno dessa ferramenta, o capítulo apresenta diferentes papéis assumidos pelo software e discute o significado dos processos de padronização na viabilização dos processos de modelagem de dados imbricados no desenvolvimento da ferramenta. Em seguida, o capítulo aborda os dilemas que permeiam o software como “produto tecnológico”, abordando os conflitos entre abertura e fechamento no decorrer do desenvolvimento da ferramenta.

Por fim, a conclusão retoma os principais resultados da pesquisa. A conclusão parte de uma reflexão da Ciência Aberta como parte do processo mais amplo de configuração de uma nova área de pesquisa e de reconhecimento de um centro de pesquisa. A tese defende que as escolhas em relação à Ciência Aberta, nesse local, inserem-se em processos mais amplos e complexos de institucionalização e reconhecimento do NeuroMat em sua área de atuação. Nesse contexto, a Ciência Aberta é colocada como uma viabilizadora da produção de conhecimento no domínio da pesquisa sobre o cérebro humano desenvolvida no centro de pesquisa estudado.

1 CIÊNCIA ABERTA: APROXIMAÇÕES CONCEITUAIS

Este capítulo partiu de um questionamento que surgiu no início do meu percurso de pesquisa: o que pode ser entendido como Ciência Aberta? A partir de uma revisão bibliográfica, o capítulo fornece um entendimento dos princípios e temas frequentemente relacionados à Ciência Aberta enquanto um movimento e como uma abordagem que se volta para a abertura dos produtos e processos da atividade científica.

Descrevo os princípios normativos que estão na base da comunicação científica, fazendo uma relação com os conflitos entre o *ethos* aberto e a privatização do conhecimento científico. No capítulo apresento a relação dos debates recentes sobre Ciência Aberta com o movimento de Software Livre e de Código Aberto. Em seguida, adentro as principais abordagens de Ciência Aberta, materializadas em diferentes iniciativas. Em sua última parte, discuto a tendência mais recente do movimento de Ciência Aberta, apresentando os aspectos conceituais, políticos e tecnológicos implicadas na abertura de dados de pesquisa.

1.1 O *ethos* científico e a privatização do conhecimento

No início desta pesquisa, o termo “*Open Science*” e Ciência Aberta – como passou a ser empregado no Brasil – parecia algo totalmente novo. Ao longo do levantamento bibliográfico, o que aparentemente parecia ser um termo nascente, revelou-se não tão novo assim.

Diferentes autores já utilizaram *Open Science* em outros contextos como sinônimo do imperativo da comunicação pública dos resultados da atividade científica a que estão submetidos todos os cientistas (CHUBIN, 1985; DAVID 1998; MAURER, 2003; KULCZYCKI, 2016).

Joly et al (2012) argumenta que, de fato, o compartilhamento nos moldes da Ciência Aberta como abertura dos resultados na Internet é uma continuação do “progresso científico”, um conceito que surgiu no século XVI associado com o ideal da publicação livre e aberta de conhecimento por meio da comunicação científica.

O sociólogo Robert K. Merton, que buscou compreender o funcionamento da ciência a partir da análise das dinâmicas internas de comunicação entre cientistas, descreveu as práticas de publicação dos resultados científicos como uma norma obrigatória a que estão submetidos todos os cientistas. A comunicação dos resultados de pesquisa é, segundo a sociologia da ciência de Robert Merton, parte integrante do complexo de valores e normas internas da ciência que constituem uma obrigação moral a ser seguida pelo cientista, ou seja, o *ethos* científico (MERTON, 1942, 1979)¹³.

Ao mesmo tempo, o ato de comunicar o trabalho com os pares assegura a autoria do cientista sobre o seu produto e lhe é outorgado o mérito científico (VELHO, 1997, MERTON, 1988). Para essa sociologia de cunho normativo, a publicização dos avanços e descobertas científicas constitui parte vital na organização social da ciência, de modo que apenas quando os cientistas tornam os resultados das suas pesquisas acessíveis para a sociedade e para os pares é que seus trabalhos se tornam legitimados e ele pode ser considerado, enfim, um cientista.

A noção de abertura como prática pública é tratada pelo economista Paul David sob a ótica da institucionalização das práticas de publicação e colaboração que caracterizaram a Revolução Científica iniciada no Renascimento. O imperativo da comunicação pública e o caráter aberto da ciência, no sentido em que os cientistas são julgados e certificados pelos iguais e pela sociedade, desenvolveu-se durante a Revolução Científica dos séculos XVI e XVII e foi o resultado de uma mudança cultural ampla e paulatina na atividade acadêmica.

Essa transformação foi representada pelo desenvolvimento dos procedimentos de construção de reputações, tendo como referência os colegas especialistas; e pela criação de

13 O *ethos* acadêmico descrito por Merton é formado por um conjunto de quatro imperativos institucionais: o “Comunalismo” - está ligado à disseminação acessível e pública dos resultados para os demais cientistas e a sociedade. O cientista deve ter consciência de que o conhecimento proporcionado pelo trabalho científico é patrimônio comum da humanidade e não propriedade privada de um indivíduo; o “Universalismo” - não há exclusão por nenhum critério exterior à ciência. Os trabalhos científicos devem seguir padrões universais de avaliação. Nesta norma, critérios sociais como raça, religião, gênero, são considerados irrelevantes; o “Desinteresse” - O único objetivo a curto prazo do trabalho científico é a ampliação do conhecimento humano. O imperativo do Desinteresse refere-se à prioridade atribuída ao progresso do conhecimento científico em detrimento de quaisquer outros de natureza pessoal; o “Ceticismo organizado” - que serve de base a polêmicas científicas e à avaliação crítica de cientistas. O cientista deve ser privado de qualquer forma de preconceito e de conclusões precipitadas sobre seus trabalhos.

redes de correspondência entre os cientistas, culminando com a criação do sistema de revisão por pares e dos primeiros periódicos em que passariam a circular os trabalhos científicos. David (2003) argumenta que o surgimento dos canais de comunicação científica nesse período representou uma mudança organizacional revolucionária na institucionalização da ciência uma vez que a cultura baseada no segredo e no mecenato converteu-se em uma cultura fundamentada na reputação profissional construída a partir de uma comunidade de referência e na autonomia científica.

A partir de então, e durante os séculos seguintes, a infraestrutura de circulação do conhecimento científico – formada por periódicos, editores, revisores, bibliotecários, e os próprios cientistas – estruturou-se em torno de instituições intermediadoras especializadas integradas por meio de normas, rotinas, padrões e sistemas técnicos, formando o substrato das funções básicas de certificação, registro, acesso e preservação do conhecimento científico (PLANTIN et al, 2018). Em torno dessa infraestrutura, passou-se a se sustentar também todo um sistema de recompensas sobre o qual passou a se estabelecer o funcionamento da distribuição de crédito entre os cientistas. Esse sistema define que o primeiro a contribuir para a comunidade científica com novas descobertas recebe em troca diferentes formas de reconhecimento simbólico e material (HAGSTROM, 1965).

A história da ciência e da tecnologia é marcada por muitos conflitos em torno da circulação e apropriação dos resultados de pesquisa. Apesar das normas internas que presam pela abertura dos resultados, existe também uma cultura da proteção e retenção dos resultados com objetivos diversos. Em diferentes aspectos, o fortalecimento dessa cultura se deu em um ambiente competitivo que se formou no meio acadêmico no Pós-Guerra.

O período Pós-Guerra marca um importante momento no estabelecimento de um novo contrato social entre Estado, comunidades científicas e indústria, reconfigurando o *ethos* acadêmico tradicional, no sentido em que estabeleceu novos objetivos e funções para a atividade científica cada vez mais vinculados aos interesses e necessidades da sociedade. Em tal contexto, a produção científica acadêmica passa a ser marcada pelo crescimento da participação de outros atores no sistema científico e maior porosidade das fronteiras entre academia e indústria¹⁴.

De acordo com Caulfield et al (2012), no contexto descrito, a atividade científica torna-se, de maneira geral, mais flexível e passa a conviver fortemente com valores de uma cultura empreendedora e com fins industriais¹⁵. Nessa configuração, entre as diferentes possibilidades de disponibilizar os produtos da atividade científica para a sociedade, a transferência de tecnologia, principalmente por meio de projetos de cooperação em pesquisa e desenvolvimento entre universidades e empresas, passou a ter espaço significativo. Em tal contexto, os direitos de propriedade intelectual assumiram, na visão de Oliveira (2011) e Delfanti (2013) uma grande centralidade econômica e organizacional na pesquisa científica, funcionando como mecanismos de apropriação do conhecimento científico e dos produtos tecnológicos produzidos na academia pelo mercado¹⁶.

14 Este novo contexto, que se configurou no Pós-Segunda Guerra, é frequentemente descrito como “Modo 2” ou “ciência pós-acadêmica” – em oposição à forma de organização clássica da pesquisa científica com base nos valores mertonianos (Modo 1). Gibbons (1994) descreve a organização da ciência no modo 2 em termos de uma mudança de valores que conduzem os esforços de pesquisa nas universidades e outras instituições acadêmicas. No modo 2, o conhecimento científico passa a ser produzido no contexto da sua aplicação e com influência direta de diferentes atores sociais (econômicos, políticos e institucionais) – em contraste com o modo 1, no qual a pesquisa é realizada de forma desinteressada e isolado do resto da sociedade (a ciência pela ciência). Alguns sociólogos criticam essa separação entre pesquisa acadêmica e pesquisa industrial, visto que, no sistema científico, as práticas científicas existentes nesses dois ambientes são similares e frequentemente complementares e com objetivos comuns. Cf. Oliveira (2011).

15 Para Ziman (2000), os valores tradicionais do *ethos* científico no Pós-Segunda Guerra passam a ser desafiados com uma ciência cada vez mais voltada para o mercado, pois a produção do conhecimento torna-se cada vez mais proprietária, local, autoritária, comissionada e baseada no papel dos especialistas para resolver problemas.

16 Embora as controvérsias sobre a instituição da propriedade intelectual na ciência perpassem a história da ciência e da tecnologia, a década de 1980 representa um período em que as disputas sobre a apropriação privada do conhecimento científico e da tecnologia tornaram-se bastante evidentes. De acordo com Hope (2008), isso ocorreu em relação à atenção que o licenciamento e patenteamento de tecnologias pelas universidades passa a ganhar nesse período – em razão de diversos estímulos governamentais – e da extensão do patenteamento da proteção do conhecimento para domínios científicos e tecnológicos até então fora do escopo dos direitos de propriedade intelectual. Hope (2008) e Delfanti (2013) apontam a aprovação do Bayh-Dole Act nos EUA como um marco na política de estímulo à proteção e comercialização de tecnologias nas universidades, a partir da qual passou-se a permitir às pequenas empresas e as organizações sem fins lucrativos, incluindo as universidades, serem titulares das invenções obtidas a partir dos resultados dos programas de pesquisas financiados com recursos governamentais. A aprovação da lei, seguida por leis similares em contexto internacional, resultou em um aumento de patentes em contexto universitário, além da

Ao mesmo tempo, esse processo em direção a uma prática científica voltada para os interesses econômicos foi paralelo a uma reação da sociedade e acirramento dos debates acerca da proteção e comercialização dos produtos da atividade científica e emergência dos debates e questionamentos da relação entre academia e mercado, resultando em diversas controvérsias e posicionamentos contra e favor ao patenteamento na ciência, tanto no meio acadêmico quanto fora dele.

As disputas em torno da propriedade intelectual ficaram bastante evidentes no campo da biotecnologia, especificamente com o desenvolvimento das pesquisas em genômica, que resultaram em muitas reações contra o “patenteamento da vida”. Nesse período, e mais fortemente na década de 1990, as controvérsias públicas envolvendo a ciência tornaram-se mais fortes, assim como passaram a ser cada vez mais circundadas por decisões em torno da propriedade intelectual sobre descobertas científicas e inovações tecnológicas¹⁷ (HILGARTNER, 2009, 1997).

Muitos destes debates ocorreram em áreas em que houve uma emergência muito rápida de inovações tecnológicas, como a biotecnologia, os softwares e a Internet, ocorrendo diversas disputas sobre os direitos autorais, o patenteamento de formas de vida, genes, algoritmos, entre outros. A partir desse momento, um crescente número de acadêmicos, organizações não governamentais, ativistas e outros observadores passaram a ver a política de propriedade intelectual como um campo de batalha no qual estão implicadas profundas decisões sobre o futuro da sociedade contemporânea em âmbito local e global (HILGARTNER, 2009). Isso se refletiu na ascensão de movimentos sociais que passaram a contestar a extensão das políticas de propriedade intelectual em domínios como o de tecnologias de informação, tecnologias e recursos genéticos e saúde global, conhecimento tradicional das comunidades, alimentos, a biodiversidade, medicamentos, entre outros (CHAVES; OLIVEIRA, 2007).

Toda essas disputas marcaram, por exemplo, o Projeto Genoma Humano, que entre as muitas polêmicas e debates, o caso envolvendo o geneticista Craig Venter trouxe uma série de questões e controvérsias para a apropriação do conhecimento. Em 1991, no início do Projeto Genoma Humano, o então pesquisador do Instituto Nacional de Saúde anunciou que havia requerido o patenteamento de sequências de DNA associado ao código genético. Após o

criação de escritórios de transferência de tecnologia nas universidades, núcleos de inovação, empresas fundadas por acadêmicos, entre outros.

¹⁷ Para uma análise das controvérsias envolvendo o acesso à informação na pesquisa genômica Cf. Hilgartner (2009, 1997).

patenteamento dessas sequências, Venter saiu dos Institutos Nacionais da Saúde (NIH) e fundou o seu próprio instituto de pesquisa, financiado por capital privado. Venter recebeu diversas críticas da comunidade acadêmica internacional, principalmente porque se temia que a falta de acesso público aos sequenciamentos patenteados impactasse o desenvolvimento de pesquisas futuras no campo da biotecnologia (DELFANTI, 2013; HOPE, 2008).

Naquele momento, colocou-se no centro dos debates sobre o desenvolvimento da biotecnologia o papel das infraestruturas de informação públicas e abertas como recursos essenciais para os avanços científicos na disciplina nascente, marca uma “virada informacional” nas ciências biológicas (VERMEULEN et al, 2013), representada em um esforço da comunidade acadêmica no desenvolvimento de repositórios públicos de dados de pesquisa como forma de manter um espaço comum de acesso ao conhecimento.

1.2 Ciência, código aberto e compartilhamento

O movimento de Software Livre e de Código Aberto também são marcos para o estabelecimento de novas formas de produção, apropriação e circulação de produções intelectuais.

A Ciência Aberta contemporânea é altamente influenciada pelos movimentos do Software Livre e Código Aberto (*Free and Open Source Software – FLOSS*)¹⁸, considerados grandes precursores dos debates sobre colaboração, compartilhamento e conhecimento aberto. Os movimentos FLOSS são caracterizados pela existência de uma cultura própria em torno da cooperação e distribuição em rede de código-fonte de programas de computador, sendo, ao mesmo tempo, permeados por um ativismo político em torno da liberdade informacional como liberdade de expressão, da Internet livre e pelos posicionamentos contrários à propriedade intelectual¹⁹ (SILVEIRA, 2014a; DELFANTI, 2013; COLEMAN, 2010).

18 Utilizo a sigla FLOSS para me referir de maneira abrangente aos movimentos Software Livre e *Open Source* e aos elementos sociotécnicos que aproximam esses duas correntes. Reconheço a existência das diferenças conceituais e políticas que marcam os dois movimentos, mas não é objetivo da tese debater essas questões. Para uma discussão sobre essas duas vertentes Cf. Coleman (2010) e Evangelista (2010).

19 De acordo com Lawrence Lessig (2005), o contexto de apropriação de diferentes bens intelectuais viabilizados pelas técnicas de circulação, reprodução e digitalização de conteúdos possibilitou o surgimento de uma “cultura livre” (*free culture*) cujas práticas estão imbricadas em diversos casos com um contraponto aos mecanismos de propriedade intelectual, especificamente leis de direitos autorais e patentes, e se sustentando na liberdade de acesso, (re)uso e compartilhamento da informação.

O FLOSS trouxe novas formas de produção e circulação do conhecimento que de muitas maneiras, subvertem práticas sedimentadas e limites colocados à liberdade de circulação e apropriação da informação e da tecnologia – representada, por exemplo, pelas licenças jurídicas, barreiras técnicas e interesses privados, configura-se, desse modo, como um movimento de caráter político. Nesse sentido, Coleman afirma que o desenvolvimento do software livre não se limita à produção de tecnologia, mas inscreve-se em um processo mais amplo de debate sobre “liberdades civis, propriedade e software, afirmando de novas maneiras que código é uma forma de expressão” (COLEMAN, 2009, p.420 apud ARAÚJO, 2018, p.32).

Uma das principais resultados dessa cultura – e um dos principais instrumentos que estruturam práticas de colaboração e compartilhamento – são licenças jurídicas abertas ou *Copyleft*, que instituem um novo código de conduta para a circulação e a apropriação de produções intelectuais com base nas liberdades de uso, modificação e compartilhamento.

Os movimentos contemporâneas pela abertura na ciência resvalam de muitas maneiras no movimento FLOSS. Alguns dos elementos que conectam esses dois universos ultrapassam os sistemas técnicos – empregados em muitos processos científicos. Silveira (2014b) argumenta que o fundamento das propostas do fundador do Movimento Software Livre, Richard Stallman, estava em liberar o acesso ao conhecimento que constitui o software, tido como um conhecimento comum, pertencente a todos ou a um coletivo ou comunidade. A lógica desse pensamento, consolidado nas práticas de produção do software livre, buscava utilizar a lógica original da ciência moderna, na qual a fonte do conhecimento científico é o próprio conhecimento, de modo que a circulação aberta e livre é colocada como um princípio básico.

Como o conhecimento científico, construído “sobre ombros de gigantes”, a distribuição dos softwares com código fonte aberto é o que permite agregar melhorias ao software e criar outros desenvolvimentos “em cima” do código. Outro elemento que permite identificar uma aproximação entre essas duas culturas de produção refere-se aos mecanismos internos para distribuição de reconhecimento. A lógica interna de distribuição de reconhecimento nas comunidades de software livre aproxima-se da lógica de distribuição do reconhecimento na ciência no sentido em que, em teoria, a posição dos contribuidores ou o

seu reconhecimento dentro da comunidade é feita com base no mérito das contribuições dadas, isto é, na agregação de melhorias aos códigos de software²⁰.

Como o FLOSS, esse movimento na ciência ocorre em um cenário de conflitos. Albagli (2015) afirma que as práticas abertas no contexto da Internet situam-se em um quadro paradoxal, em que, de um lado, encontram-se os novos modos de produção colaborativa e de acesso amplo às redes de compartilhamento e, do outro, estão os mecanismos de captura e de privatização da informação e dos meios de acesso, que atuam cerceando os usos visando a extrair valor da informação. Esse paradoxo pode ser observado quando se reconhece que as mesmas tecnologias que viabilizam práticas descentralizadas de compartilhamento da informação também viabilizam formas de coibir o acesso, bem como formas de vigilância, especialmente quando as práticas abertas rivalizam com modelos de negócio sustentados nos direitos de propriedade intelectual.

Diante das práticas baseadas no compartilhamento, as empresas detentoras dos direitos autorais recorrem a mecanismos sofisticados para monitoramento dos usos e contra-usos do conteúdo que comercializam²¹. Em tal contexto, aponta-se que os modelos de negócio baseados na restrição de acesso à informação impactam negativamente o avanço da ciência, uma vez que há o encarecimento do acesso às fontes de informação necessárias ao desenvolvimento da pesquisa, e também os riscos jurídicos implicados em práticas triviais e cotidianas como fazer *download* de um artigo ou fazer cópias de um livro e disponibilizá-lo online.

20 As disputas conceituais e políticas que marcam a história do Software Livre e, posteriormente, do movimento de Código Aberto também são análogas àquelas envolvendo abertura e privatização da ciência. Acerca dessas disputas destaco o texto “A Catedral e o Bazar” de Eric Raymond (1999) e Evangelista (2010) para compreender as diferenças entre os modelos aberto e fechado na produção de software.

21 Entre as tecnologias empregadas para impedir a produção de cópias piratas estão o *Digital Rights Management* (DRM), que são tecnologias utilizadas para controle de acesso ou uso de obras ou criações protegidas por direitos autorais ou hardwares proprietários. O DRM é comumente utilizado para distribuição de obras intelectuais, como revistas, filmes e livros que são comprados através da Web. Disponível em: <<http://www.w3c.br>>. Acesso em: 20 mar. 2018.

Segundo Paul David (2003, 2008), neste cenário no qual os produtos da ciência são apropriados por poucos, incluindo empresas que buscam exclusividade sobre as fontes de informação, muitos cientistas enfrentam problemas ao longo do desenvolvimento da pesquisa porque os insumos da pesquisa estão protegidos por patentes ou dependem da assinatura de periódicos e bases de dados pagas²².

1.3 Ciência Aberta: múltiplas abordagens

Ciência Aberta é compreendida como um modelo de pesquisa científica fundamentado essencialmente na comunicação aberta dos diversos processos de produção da informação científica de acordo com os princípios da cultura da liberdade, característicos da cultura digital (FECHER; FRIESIKE, 2014). Nesse sentido, as possibilidades de interpretação e direções dessa abordagem são tão amplas e heterogêneas como são as práticas que caracterizam a ciência. Assim, Ciência Aberta vai se referir a muitas coisas, conformando atualmente um termo abrangente que compreende diferentes perspectivas, interesses e pontos de vistas (ALBAGLI, 2015).

Alguns autores vão associar o desenvolvimento de um movimento contemporâneo em direção à Ciência Aberta como uma continuidade das disputas e controvérsias em torno da apropriação dos resultados da pesquisa científica por meio da propriedade intelectual que tradicionalmente marcaram a história da ciência e da tecnologia. Os resultados variados alcançados pelas políticas e pressões para comercialização da pesquisa por meio de patentes, que como evidenciado por diferentes estudos, podem ter um efeito negativo no desenvolvimento da pesquisa científica. David (2008) e Caulfield et al, (2012) enquanto o sistema de recompensas na academia tende a produzir conflitos e competição acima da cooperação científica, o papel de uma abordagem de Ciência Aberta, por outro lado, tem o propósito de aumentar o estoque de conhecimento e informação confiável necessário ao desenvolvimento científico e criar um ambiente de pesquisa científica aberto, colegiado e colaborativo.

22 A insatisfação da comunidade científica com os modelos de negócio das editoras de periódicos científicos é um dos desencadeadores do Movimento *Open Access* ou Acesso Aberto, que advoga pelo acesso gratuito e livre a publicações científicas na Internet. CF. Santos (2014).

Partindo dessa perspectiva, Alessandro Delfanti, ao estudar as práticas contemporâneas de Ciência Aberta em laboratórios DIY (do inglês *Do It Yourself*, “faça você mesmo”), apresenta uma compreensão que coloca a Ciência Aberta como uma alternativa ao modelo de organização da pesquisa acadêmica tradicional – que se baseia no paradigma da propriedade intelectual e na competitividade acadêmica – como forma de contornar os efeitos adversos da propriedade intelectual sobre o desenvolvimento da pesquisa. O autor compreende então a Ciência Aberta como um:

Método para produzir conhecimento científico, divulgando seus resultados e abrindo-os à revisão de toda a comunidade científica, maximizando a circulação de informação e conhecimento e compartilhando. No extremo oposto, há uma “ciência fechada”, como um segredo ou uma ciência em que a dinâmica de comunicação é limitada pelos muros de uma instituição ou sujeita ao pagamento de uma licença, como uma patente ou direito autor (DELFANTI, 2013, p. 27).

Para Delfanti (2013), a Ciência Aberta contemporânea “*remixa*” as propriedades de um *ethos* científico com elementos que são próprios da cultura FLOSS. Nesse *remix*, é reforçada uma visão normativa da ciência enquanto prática desinteressada, cooperativa e centrada na publicização dos resultados, assim como práticas de comunicação e colaboração características da cultura FLOSS – como dispositivos, conceitos, licenças abertas, software de código aberto, entre outros sistemas e padrões tecnológicos abertos.

Há uma prática recorrente de se usar a Ciência Aberta como sinônimo do movimento pelo Acesso Aberto (*Open Access*) – que defende a disponibilização aberta e gratuita de publicações científicas na Internet. Entretanto, é importante demarcar uma distinção entre esses dois termos. O Movimento *Open Access* tem uma história própria e possui uma relativa consolidação de suas iniciativas em torno da economia das publicações científicas. Por outro lado, esse movimento, por meio da consolidação de uma série de ações para publicações de acesso gratuito, criou um contexto propício para se pensar cada vez mais a abertura na ciência mais abrangente.

Essa perspectiva coloca o acesso aberto à literatura científica como uma das vertentes, ou mesmo como um dos pilares, de um processo mais amplo e complexo que vai tratar a abertura na ciência como algo que deve envolver processos, resultados e modos colaborativos e horizontalizados de produção acadêmica, além do engajamento entre a pesquisa científica e a sociedade (SANTOS et al, 2017).

De acordo com as análises da bibliografia recente²³, abertura, na abordagem de Ciência Aberta, engloba práticas diferentes e se direciona a elementos diversos do trabalho de pesquisa e seus resultados: artigos científicos; conjunto de dados (*datasets*) e metadados. imagens; softwares e seus códigos-fonte; revisões; relatórios; apresentações; a abertura dos processos de pesquisa e fluxos de trabalho; a aplicação de licenças jurídicas flexíveis que permitam o acesso e reutilização ampla dos resultados; experimentação em ambientes não formais e a utilização de equipamentos de código aberto (hardware aberto) no ambiente de pesquisa; a interoperabilidade da infraestrutura científica; metodologias abertas e compartilhadas de pesquisa e ferramentas que permitam a mineração de texto e dados, para citar alguns dos processos e produtos abordados.

A figura 1, a seguir, foi desenvolvida pela iniciativa FOSTER Open Science com o propósito de esquematizar os diferentes temas que configuram a Ciência Aberta atualmente. Nesse esquema, representa-se como uma abordagem que se aplica a todo o ciclo da pesquisa, desde os seus momentos iniciais, ao modo como a pesquisa é desenvolvida e aos resultados finais.

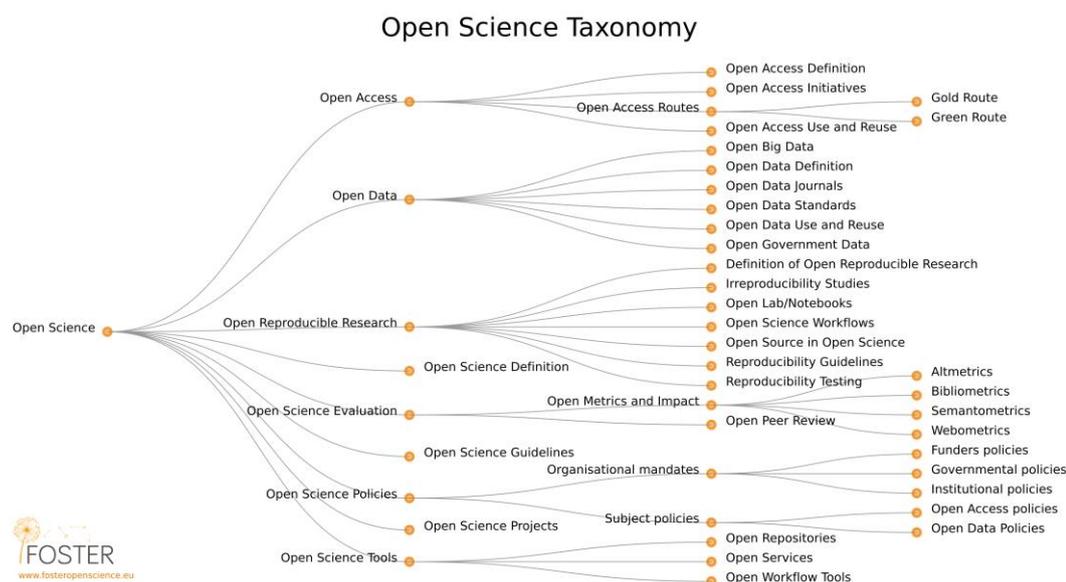


Figura 1: Taxonomia da Ciência Aberta.

Fonte: <<https://www.fosteropenscience.eu>>.

Fecher e Friesike (2014) descrevem a Ciência Aberta como um termo abrangente que engloba uma infinidade de suposições sobre o futuro da produção e a disseminação do conhecimento científico. Por meio da análise de uma série de declarações, manifestos, políticas institucionais, artigos acadêmicos, entre outros documentos que consignam o tema, os autores propõem uma classificação, a partir de cinco “escolas de pensamento” ou “correntes” em tono das quais se estrutura os discursos mais comuns sobre o assunto. No quadro 1, a seguir, apresenta-se uma sistematização das correntes de Ciência Aberta proposta pelos autores. São destacadas as hipóteses centrais de cada corrente, os objetivos e os recursos comumente empregados.

Quadro 1: Correntes de Ciência Aberta.

Corrente	Hipótese Central	Objetivo	Recursos envolvidos
Democratização	O acesso ao conhecimento é distribuído de forma desigual.	Tornar o conhecimento acadêmico disponível para todos.	-Acesso aberto -Dados abertos, -Código aberto
Pragmática	A produção de conhecimento pode ser mais eficiente quando os cientistas trabalham em colaboração.	Abriir os processos de produção de conhecimento.	-Inteligência coletiva. -Redes de compartilhamento -Dados abertos -Código aberto
Público	A ciência precisa estar acessível à população.	Tornar a ciência acessível para os cidadãos	-Ciência cidadã -Relacionamento da ciência com o público -Blogs de ciência
Métricas	A produção científica necessita de formas alternativas de avaliação de impacto.	Desenvolver um sistema de métricas alternativas de avaliação do impacto acadêmico.	- Altmétrias - Revisão por pares -Citações -Fator de Impacto
Infraestruturas	A eficiência da pesquisa depende da disponibilidade de ferramentas e aplicações computacionais.	Criação de plataformas abertas, ferramentas e serviços para os cientistas.	-Plataformas colaborativas -Ferramentas computacionais (ex. Softwares, repositórios)

Fonte: Elaboração própria a partir de Fecher e Friesike (2014).

A corrente da **democratização** caracteriza as iniciativas que estão mais fortemente preocupadas com questões sobre democratização do conhecimento científico, principalmente em contextos em que se origina de pesquisas financiadas com recursos públicos. Enquadram-se nesse grupo as iniciativas específicas do movimento pelo Acesso Aberto, que atua na abertura dos resultados de pesquisa, tradicionalmente tornados públicos através de artigos publicados em revistas científicas.

Os autores mencionam também o movimento de Dados Abertos (*Open Data*), que busca garantir que os dados produzidos pelas pesquisas científicas sejam disponibilizados de maneira aberta e em formatos que possibilitem não apenas a sua consulta, mas seu escrutínio e reutilização em pesquisas posteriores. Em ambos os métodos, a escola democrática tem como alvo de crítica as políticas editoriais que buscam cercear o acesso à literatura científica pela cobrança de assinaturas com preços elevados, pela omissão de dados coletados ao longo da pesquisa e que poderiam ser tornados públicos como informação de apoio para subsidiar novos trabalhos.

A corrente **pragmática** é centrada na colaboração científica por meio do compartilhamento aberto dos processos e resultados de pesquisa. Nessas iniciativas, as tecnologias digitais são empregadas como um meio mais eficiente de disseminar os resultados, possibilitando aos pesquisadores atuarem colaborativamente na resolução de problemas específicos e alcançar resultados e soluções esperadas mais rapidamente. A abordagem pragmática está presente em formas de colaboração científica que requerem a obtenção de resultados imediatos. Nessa abordagem, o compartilhamento de dados brutos e de ferramentas são centrais.

Uma estratégia empregada em tais processos podem ser os “cadernos abertos de pesquisa” ou “*open notebooks*” para a publicação de resultados em tempo real e da divulgação da pesquisa em estágios que normalmente não são tornados públicos. Os cadernos abertos de pesquisa são utilizados, em muitos casos, quando os pesquisadores buscam na divulgação em tempo real e na transparência dos processos uma forma de obter contribuições de outros pesquisadores de forma mais rápida, como a identificação de incongruências experimentais²⁴. Desse modo, com a abertura dos cadernos, ao contrário do processo de construção ficar restrito ao laboratório ou a um grupo de especialistas, os pesquisadores

24 Rede de cadernos de pesquisa abertos Cf. <<http://onsnetwork.org/>>. Acesso em: 20 set. 2014.

abrem os resultados de pesquisa em estágio inicial para uma audiência mais ampla para recebimento de comentários e *feedbacks*²⁵.

São empregadas diferentes ferramentas digitais que permitem a colaboração, compartilhamento e disseminação de resultados de pesquisa de maneira descentralizada. São empregadas ferramentas como *wikis*²⁶, redes sociais, blogs e outras plataformas colaborativas online.

A corrente do **público** refere-se a práticas que buscam a aproximação da ciência com grupos externos ao meio acadêmico, seja através da facilitação e diversificação dos canais de comunicação com o público, utilizando blogs, sites, redes sociais ou por meio de ações que permitam a interação e a participação ativa dos não-especialistas interessados em colaborar em projetos de pesquisa (FECHER; FRIESIKE, 2014). As iniciativas mais comuns nessa categoria correspondem ao que vem sendo chamado de “Ciência Cidadã”, também conhecido como *crowd science*. Em projetos de Ciência Cidadã as pesquisas são abertas à colaboração de pessoas que não possuem formação científica formal, mas que possuem *know-how* no assunto do projeto ou que podem contribuir em tarefas de coleta e organização de dados em plataformas online²⁷.

A corrente das **infraestruturas** centra-se nas plataformas e ferramentas tecnológicas que viabilizam práticas de colaboração e compartilhamento. Fecher e Friesike (2014) situam a computação distribuída nessa categoria, incluindo a arquitetura computacional em GRID, que permite o processamento de grandes volumes de dados de forma compartilhada e a estruturação de arranjos de colaboração em pesquisa entre instituições localizadas em diferentes territórios geográficos (CARDOSO et al, 2012)²⁸.

Incluem-se nessa categoria também os ambientes digitais online de armazenamento e compartilhamento de conjuntos de dados, artigos, código-fonte e protocolos de pesquisa. Os ambientes mais comuns são os repositórios digitais. Os repositórios são ambientes digitais que têm a função de arquivar e preservar digitalmente coleções de materiais de pesquisa, como imagens, conjunto de dados, textos, protocolos, modelos

25 Para uma análise da utilização dos cadernos de pesquisa abertos Cf. Clinio (2015) e Clinio e Albagli (2016).

26 *Wiki* é uma coleção expansível de páginas web interligadas em um sistema de hipertexto que permite a edição e armazenamento de informações de maneira colaborativa e rápida. Cf. Alves et al (2016).

27 Para exemplos de projetos de Ciência Cidadã Cf. Franzoni e Sauermann (2014) e Nielsen (2012).

28 Atualmente, um dos principais projetos de computação em Grid é o consórcio Open Science Grid, uma grande infraestrutura computacional nos EUA que oferece suporte a diversas aplicações científicas para a análise intensiva de dados por meio da conexão entre múltiplos computadores de alta performance. Disponível em: <<https://opensciencegrid.org/>>. Acesso em: 12 ago. 2015.

computacionais. Os repositórios digitais são centrais nas práticas contemporâneas de comunicação científica, uma vez que adquiriram um importante papel na viabilização de novas práticas de comunicação dos produtos gerados na pesquisa científica, como dados primários, códigos, versões de artigos científicos entre outros conteúdos.

A corrente das **métricas** engloba as formas de avaliação do impacto e mérito científico. Engloba a transparência nos processos de avaliação e do uso de determinados rankings de avaliação de impacto científico. Estão incluídas aí as métricas alternativas (altmetrias) para medir o impacto e mérito científico utilizando fontes alternativas de dados (número de downloads e visualização de artigos, compartilhamento em redes sociais, incorporação em sites e blogs, entre outros). Essas novas formas de mensuração baseiam-se em uma série de evidências empíricas e no argumento principal de que a disponibilização em acesso aberto na *Web* aberta possui efeito real sobre o impacto das pesquisas, o que se traduz em mais visibilidade para o autor e mais citações para sua produção²⁹.

A corrente das métricas trata também do uso de instrumentos alternativos nos processos de avaliação científica, como a avaliação por pares. Algumas publicações utilizam, por exemplo, a revisão por pares aberta (*Open Peer Review*), que visa a um processo transparente a partir da divulgação da identidade dos revisores³⁰.

Fecher e Friesike (2014) enfatizam que essas correntes não são excludentes ou estanques, mas que, ao contrário, são intrinsecamente interligadas. Pode-se mencionar as infraestruturas tecnológicas e os dados de pesquisa como elementos transversais a todas essas correntes.

O próximo tópico aborda a centralidade dos dados de pesquisa como tendência recente do movimento de Ciência Aberta.

29 Segundo a abordagem alternativa para as métricas, os instrumentos de avaliação – que se baseiam preponderantemente no periódico científico e em seus indicadores derivados, especialmente o Fator de Impacto (FAUSTO, 2016) – são também objeto de preocupação quando se trata de *accountability* e de transparência da ciência. Um movimento recente em direção ao desenvolvimento de métricas alternativas de avaliação de impacto científico tem se voltado contra o uso do Fator de Impacto como métrica hegemônica da ciência. Uma dessas iniciativas foi a “*San Francisco Declaration on Research Assessment*” (DORA), lançada em 2012 durante o encontro da *American Society for Cell Biology*. Disponível em: <<https://sfdora.org/read/>>. Acesso em: 20 mar. 2018.

30 A Nature e alguns periódicos da Editora Elsevier vêm utilizando, de maneira experimental, algumas formas de revisão por pares aberta. Periódicos como a PeerJ e a *Public Library of Science* (PLOS), ambas da área de ciências da vida, utilizam o modelo de acesso aberto, empregam diferentes formas de transparência de revisão por pares, como a pós-publicação e publicação dos pareceres. Para um aprofundamento sobre essas formas de revisão por pares Cf. Ford (2015) e Ross-Hellauer (2016).

1.4 Abertura de dados de pesquisa: políticas e infraestruturas

Para Leonelli (2013), a ênfase crescente na abertura dos dados de pesquisa reflete um interesse social e econômico nos dados que caracteriza o atual estágio da sociedade da informação. Nesse contexto, a análise de grandes volumes de dados tem sido cada vez mais integrada aos processos de descoberta, produção de conhecimento e geração de produtos, conferindo aos dados de pesquisa científica um valor social e econômico que ultrapassa as fronteiras da ciência.

A recente virada para a abertura dos dados representa uma mudança significativa no sistema de comunicação científica. De acordo com alguns autores, até recentemente, com algumas exceções, os dados permaneciam essencialmente fora da comunicação acadêmica formal (PLANTIN et al, 2018; LEONELLI, 2013; KOWALCZYK; SHANKAR, 2013; BORGMAN, 2015).

A prática científica e o próprio sistema de comunicação historicamente funcionou, direta ou indiretamente, no sentido de dificultar o acesso aos dados de pesquisa e o sistema acadêmico raramente foi favorável à disponibilização dos dados que não fosse no contexto da literatura formal, embasando análises e inferências ou compartilhando porções de dados dentro de comunidades isoladas. Leonelli (2013) defende que a vasta maioria dos dados científicos produzidos na segunda metade do Século XX está disponível – e é acessada – por um pequeno grupo de especialistas. Uma parte reduzida desses dados foi tornada pública somente por meio da publicação em periódicos científicos – ou seja, em quantidade muito reduzida em relação ao volume gerado durante o desenvolvimento da pesquisa.

Segundo Kowalczyk e Shankar (2013), a infraestrutura de comunicação acadêmica não estabeleceu, ao longo do seu desenvolvimento, formas para os dados serem registrados formalmente, serem citados, publicados ou submetidos à curadoria. Em parte, isso se deu devido ao predomínio do regime de publicação impressa, que encarecia e tornava praticamente inviável a comunicação dos dados associados aos artigos.

Borgman (2015) e Fecher e Friesike (2014) atribuem ainda essa tendência ao fechamento dos dados às restrições aplicadas pelos editores de periódicos ou pelos provedores dos dados por meio de licenças e, mais recentemente, utilizando formatos que dificultam a extração e reutilização – seja porque foram mal descritos ou porque foram “escondidos” em documentos não-modificáveis, como o formato PDF. Os dados também encontram-se sob um

grau de controle relacionado às disputas por prioridade na pesquisa, amplificadas pelo ambiente acadêmico competitivo.

A Ciência Aberta centrada nos dados aponta para discussões mais amplas que passam por questões envolvendo ética e integridade científica, sobretudo em contextos em que se afirma que a pesquisa científica vive uma “crise da reprodutibilidade”³¹. A principal justificativa aqui seria a existência de uma grande quantidade de pesquisas cujas especificações necessárias à criação das condições adequadas para a obtenção de resultados similares são insuficientes ou não são divulgadas, o que representa um grande problema a ser enfrentado pela ciência contemporânea (BOULTON et al, 2012; LEONELLI, 2013).

Para Munafò et al (2017), o atual sistema de reconhecimento acadêmico promove a publicação de “narrativas limpas”, que podem incentivar a comunicação incompleta dos procedimentos e dos resultados do estudo. Boulton et al (2012) compreende que a publicação aberta dos dados auxilia na identificação e controle de fraude nos resultados, bem como na meta-análise a partir de dados “brutos” em vez de resultados já trabalhados e sintetizados em artigos ou relatórios. Adicionalmente, uma série de atores vêm apontado a necessidade de abertura dos resultados como um imperativo ético necessário ao desenvolvimento da pesquisa em áreas que impactam a vida social, como saúde e meio ambiente (SANTOS et al, 2017).

O crescente interesse na abertura dos dados de pesquisa é representada por uma série de declarações lançadas nos últimos anos visando a incentivar e ampliar as práticas de compartilhamento de dados no meio acadêmico. Entre os principais incentivadores, encontram-se as agências de financiamento norte-americanas e europeias. Essas agências passaram a adotar uma série de diretrizes que, em muitos casos, torna obrigatória a disponibilização pública e aberta dos conjuntos de dados gerados em pesquisas financiadas com recursos governamentais.

As principais agências europeias e norte-americanas vêm atuando na promoção da abertura na ciência com o apoio ao desenvolvimento de infraestruturas de compartilhamento de dados e por meio de instrumentos que estimulem o arquivamento da produção científica

31 A comunicação da pesquisa com fins à reprodutibilidade compreende o processo de descrever na íntegra o desenho do estudo e os dados coletados que fundamentam os resultados relatados, em vez de uma versão curada do projeto ou um subconjunto dos dados coletados. Cf. Munafò et al (2017).

em repositórios pelos próprios pesquisadores ou a apresentação de planos de gestão de dados de pesquisa³².

Nos EUA, os Institutos Nacionais de Saúde se destacam nessa esfera. A declaração de “Compartilhamento de dados de pesquisa” explica a posição dos institutos de que a publicação de dados é essencial para a tradução rápida da pesquisa em conhecimento e para alcançar melhorias na saúde humana. A declaração também exige que os pesquisadores incluam em suas solicitações de financiamento um plano explicitando como os dados produzidos serão compartilhando ou especificar quando houver restrições que impedem a disponibilização de dados³³.

Entre as agências europeias, a *Wellcome Trust* exige, desde 2005, que os pesquisadores candidatos a receber financiamento da agência submetam um plano de gerenciamento e compartilhamento de dados ou que os disponibilizem em repositórios abertos sempre que a pesquisa for suscetível a gerar dados que tenham um valor significativo para a comunidade de pesquisa em geral. A agência estabeleceu uma estratégia de “pesquisa aberta”³⁴ criando uma plataforma para que os pesquisadores apoiados que recebem fundos da agência disponibilizem, em tempo real, os artigos e conjunto de dados de pesquisa.

A Comissão Europeia consolidou uma série de diretrizes para estimular a disponibilização de dados de pesquisas realizadas no âmbito do projeto-piloto Horizonte 2020. O projeto enfatiza principalmente o impacto positivo do compartilhamento de dados para melhorar a visibilidade da pesquisa e o impacto dos resultados no desenvolvimento econômico por meio da facilitação do acesso aos resultados de pesquisa pelas empresas. Foi criada uma série de iniciativas que buscam implementar estes objetivos.

O *Open Research Data Pilot*³⁵ consiste em ferramentas para tornar os dados acessíveis por meio de uma abordagem de dados abertos (livres para acessar, reutilizar, redirecionar e redistribuir), incluindo diretrizes que estabelecem como obrigatória a disponibilização dos dados e metadados em repositórios científicos, a aplicação de licenças *Creative Commons* e informações sobre as ferramentas e instrumentos necessários à validação e replicação dos resultados (COMISSÃO EUROPEIA, 2013). Destaca-se também a iniciativa

32 Para um levantamento das políticas de Ciência Aberta e dados abertos em diferentes países Cf. Santos et al (2017).

33 Cf. National Institutes of Health (2003).

34 Cf. WELLCOME TRUST. **Open research** | Wellcome. Disponível em: <<https://wellcome.ac.uk/what-we-do/our-work/open-research>>. Acesso em: 27 jan. 2019.

35 Cf. Koulocheri e Koulocheri (2018).

“*Fostering the Practical Implementation of Open Science in Horizon 2020 and Beyond*” (FOSTER), voltado para o treinamento e preparação de pesquisadores e outros, visando ao desenvolvimento de recursos e habilidades para a abertura dos resultados de pesquisa³⁶.

Foi criado também um grupo de trabalho E-IRG – *Infrastructure Reflection Group*, com representação em vários países da União Europeia. O grupo é dedicado ao estudo e implementação de boas práticas e recomendações para acesso, preservação e curadoria de grandes volumes de dados e padronização de metadados que possibilitem a utilização “transversal” entre disciplinas (MIRANDA, 2015). Como parte desse esforço, foi criada a *Global Open FAIR Implementation Networks* (GOFAIR) com apoio dos governos holandês e alemão, com estratégias internacionais voltadas para impulsionar o desenvolvimento de serviços de dados de acordo com os princípios FAIR (dados encontráveis, acessíveis, interoperáveis e reutilizáveis³⁷) na Europa e no mundo.

A rede GOFAIR foi, por sua vez, estabelecida para impulsionar a implementação do projeto da Comissão Europeia de criação da *European Open Science Cloud* (EOSC)³⁸. Trata-se de um projeto que pretende criar uma grande infraestrutura virtual para os dados das pesquisas financiadas pela União Europeia.

Por meio da *German Research Foundation* (GRF), a Alemanha também tem atuado com o financiamento para a criação de infraestruturas para o compartilhamento de dados. A GRF criou a ferramenta *Registry of Research Data Repositories* (Re3data)³⁹ para registrar e dar visibilidade a repositórios de dados de pesquisa.

Além da atuação dessas organizações governamentais, recentemente uma série de organizações globais e grupos emergiram como atores importantes para o estabelecimento de uma infraestrutura para a abertura de dados de pesquisa. Entre as organizações recentes, destaca-se a Aliança de Dados de Pesquisa (RDA – *Research Data Alliance*⁴⁰), lançada pela Comissão Europeia, *National Science Foundation* (NSF) e Departamento de Inovação do Governo Australiano, e, atualmente, integrada por diferentes países, instituições e áreas do conhecimento. A RDA é organizada em torno de grupo de trabalho e grupos de interesse para discutir e elaborar recomendações para dados de pesquisa em nível global.

36 FOSTER. Home Page. Disponível em: <<https://www.fosteropenscience.eu/>>. Acesso em: 24 jan. 2019.

37 FAIR é a sigla para “*Findable, Accessible, Interoperable and Reusable*”.

38 EOSC. European Open Science Cloud. Home Page. Disponível em:

<<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>>. Acesso em: 18 jan. 2019.

39 Re3data.org. **Home Page**. Disponível em: <<https://www.re3data.org/>>. Acesso em: 19 jan. 2019.

40 Research Data Sharing without barriers. Home Page. Disponível em: <<https://www.rd-alliance.org/>>. Acesso em: 22 jan. 2019.

A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) por exemplo, estabeleceu uma série de recomendações para o acesso a dados produzidos em pesquisas financiadas com recursos públicos. A declaração adotada pela organização aponta a importância do acesso a dados de pesquisa no desenvolvimento econômico e social, destacando que o valor dos dados está em sua utilização. O acesso aberto aos dados nesse sentido deve ser uma norma internacional associada à pesquisa financiada com recursos públicos (OCDE, 2007). O documento chama a atenção, entre outros pontos, para o papel do compartilhamento de dados como forma de evitar a duplicação dispendiosa e desnecessária de coleta de dados, melhorando assim a eficiência da pesquisa financiada com recursos públicos (PILAT; FUKASAKU, 2007).

Dados de pesquisa são compreendidos como representações de observações, objetos e outros materiais utilizados como evidência de um fenômeno e para a obtenção de recursos para a pesquisa e reconhecimento dos pares. Os dados podem ser representados em uma variedade de formatos. Dependendo dos estudos que os produzem, podem ser números, textos, áudio, vídeo ou imagens fixas. Esses formatos variam de acordo com o domínio da pesquisa e da comunidade científica (KOWALCZYK; SHANKAR, 2013).

Curty e Aventurier (2017) afirmam que grande parte das discussões e ações direcionadas à abertura de dados é orientada por um pressuposto de que os dados podem ser úteis para os outros – tanto dentro de uma mesma disciplina como de maneira interdisciplinar. Nesse sentido, os crescentes investimentos políticos e tecnológicos na abertura de dados justificam-se não somente pelo potencial de disponibilidade, acessibilidade e análise dos dados de pesquisa, mas na sua efetiva reutilização em outros contextos e aplicações.

A noção de dados abertos, nesse aspecto, diz respeito não somente a tornar os dados publicamente acessíveis, mas referem-se também a um padrão de disponibilização dos dados que permita a efetiva reutilização. A definição mais difundida define os “dados abertos” como aqueles que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa – sujeitos, no máximo, à exigência de atribuição da fonte e compartilhamento pelas mesmas regras. Esta definição foi proposta pela *Open Knowledge*⁴¹, que estabeleceu os requisitos necessários para que dados e conteúdos sejam considerados abertos. Os requisitos de dados abertos, citados em Lóscio et al (2018, p.13), são:

41 A *Open Knowledge* é uma organização sem fins lucrativos que atua em iniciativas voltadas para dados abertos, recursos educacionais, bens culturais, entre outros. Disponível em: <<https://br.okfn.org/>>. Acesso em: 07 abr. 2015.

I. Disponibilidade e acesso: o dado precisa estar disponível por inteiro. Deve estar num formato conveniente e modificável;

II. Reuso e redistribuição: o dado precisa ser fornecido em condições de reuso e redistribuição podendo ser combinado com outros;

III. Participação universal: todos podem usar, reusar e redistribuir o dado sem restrições de áreas, pessoas ou grupos.

Sem os requisitos acima o dado não pode ser considerado um dado aberto. O compartilhamento de dados (*data sharing*) e a publicação de dados (*data publishing*) não têm o mesmo significado que dados abertos nesses termos conceituais, que visam a ampla reutilização dos dados. De acordo com Womack (2015), a publicação de dados refere-se a um processo mais formal de tornar dados publicamente disponíveis, enquanto o compartilhamento de dados não necessariamente é algo público, podendo se referir à troca de dados entre pesquisadores de maneira informal e dentro de contextos fechados, assim como dados públicos não necessariamente são dados abertos. O compartilhamento e a publicação de dados são compreendidos como etapas em direção aos dados totalmente abertos, entretanto, são modos incompletos que não garantem que os dados estejam bem documentados, formatados e não estejam sujeitos a retribuições envolvendo direitos de uso (WOMACK, 2015).

Os dados de pesquisa apresentam uma série de características específicas que tornam sua abertura complexa. Segundo Molloy (2012), o contexto e características particulares que circundam os dados de pesquisa, sua produção e uso refletem a necessidade de princípios específicos para o campo científico com objetivo de orientar a disponibilização e uso de dados no âmbito da Ciência Aberta, sobretudo porque em alguns casos os dados podem estar acessíveis, mas não estão licenciados para serem reutilizados, ou podem estar acessíveis e livres para serem reutilizados, mas a forma como são disponibilizados cria barreiras à extração e modificação (Ex. Formato PDF).

O relatório *Science as an Open Enterprise*, publicado pela *Royal Society* em 2012, enfatiza que a mera divulgação de dados tem muito pouco valor por si só para a ciência, requerendo uma abertura “mais inteligente” na qual os dados sejam efetivamente comunicados, seguindo requisitos fundamentais que os tornem acessíveis, inteligíveis, avaliáveis e reutilizáveis. Nesse aspecto, a definição recente dos chamados “princípios FAIR”

representou um marco importante nas políticas de gerenciamento de dados de pesquisa visando a reutilização.

Os princípios FAIR são um conjunto de diretrizes orientadoras para a gestão de dados de pesquisa buscando torná-los encontráveis, acessíveis, interoperáveis e reutilizáveis (FAIR – *Findable, Accessible, Interoperable and Reusable*). Os princípios FAIR voltam-se principalmente para as práticas de reutilização dos dados de pesquisa e para o aprimoramento da capacidade dos mecanismos computacionais de encontrarem e utilizarem dados automaticamente (WILKINSON et al, 2016).

Esses princípios reforçam sobretudo o papel dos metadados como agentes fundamentais da encontrabilidade, interoperabilidade e da reutilização de dados de pesquisa. Os metadados se referem aos dados sobre os dados, permitindo sua identificação e a contextualização. Os metadados incluem, geralmente, informações sobre o produtor dos dados, as formas de coleta, período de coleta, métodos, assim como detalhes técnicos sobre a utilização (ROYAL SOCIETY, 2012).

Um aspecto relevante da configuração desse infraestrutura para a reutilização dos dados são os desenvolvimentos que possibilitem que a publicação de dados seja cada vez mais incorporada ao sistema de recompensas existente na academia. Os periódicos desempenham uma importante função neste momento. A crescente preocupação com os dados de pesquisa se refletiu em mudanças nas políticas editoriais de importantes periódicos científicos em relação à forma como abordam os dados de pesquisa.

O desenvolvimento de formas de publicação científicas destinadas a publicar dados de modo a ampliar as condições de reutilização e atribuição aos seus produtores é uma das transformações significativas no sistema de publicações verificada atualmente. De acordo com Curty e Aventurier (2017), a publicação de dados nessa perspectiva tem se estabelecido a partir de diferentes abordagens. Os repositórios representam a forma mais disseminada e relativamente tradicional de disponibilizar dados e, de modo sistemático, essas ferramentas têm incorporado recursos que possibilitam a recuperação e citação dos dados.

Com esse propósito surgiram também as chamadas Publicações “Expandidas” ou “Ampliadas”. Essas publicações disponibilizam – agregados aos artigos publicados – os materiais e métodos, bem como o conjunto de dados obtidos no processo de investigação, proporcionando aos leitores maior transparência, verificação e análise dos dados no momento da leitura e interatividade (CURTY; AVENTURIER, 2017). Algumas das principais

publicações acadêmicas, entre eles periódicos da *Public Library of Science* (PLOS), grupo Nature e *Royal Society* passaram a exigir que os autores disponibilizem, no momento da submissão dos artigos, todos os dados que estão na base das conclusões do trabalho ou indiquem o local onde esses dados podem ser acessados pelos leitores, editores, revisores e possíveis interessados, e de maneira que permitam sua reutilização, bem como a disponibilização de ferramentas estatísticas, protocolos, ferramentas de software, metadados e demais instrumentos que auxiliem na compreensão e ou reuso dos dados (BOULTON et al, 2012).

Mais recentemente, verifica-se o desenvolvimento de novas categorias de publicação científica voltadas exclusivamente para a publicação de dados: “*Data Papers*” e “*Data Journals*”. Conceitualmente, os *Data Papers* descrevem uma ou mais coleções de dados, dedicando-se ao detalhamento dos métodos para a coleta dos dados, a descrição da composição, estrutura e formato da coleção sem, no entanto, adentrar à interpretação dos dados. Os *Data Journals*, por sua vez, são uma modalidade de periódicos científicos dedicados a publicar e descrever conjuntos de dados de pesquisa, utilizando-se de uma série de recursos voltados para o depósito, apresentação e descrição dos dados e podem dispor de um repositório de dados próprio⁴².

Estes formatos de publicação científica visam à visibilidade e aperfeiçoamento dos mecanismos de reutilização e referência aos dados. Nesse contexto, Platin et al (2018) apontam ainda a importância de mecanismos de citação de dados associados às publicações de dados. Tais mecanismos – como o *Digital Object Identifier* (DOI)⁴³, atribuído a artigos de periódicos e conjuntos de dados – possibilitam a formalização da citação dos dados, permitindo atribuir referência aos dados e reconhecimento de seus produtores.

Considerações Finais

Este capítulo tratou das bases conceituais da Ciência Aberta, objeto de estudo da presente tese. Ciência Aberta é um conceito multifacetado, o que traz certa dificuldade em conferir uma unidade analítica. Longe de ser um movimento uniforme com pautas e formas de atuação comuns, a Ciência Aberta atualmente é marcada por uma diversidade de atores

42 Curty e Aventurier (2017) apresentam uma sistematização detalhada dos conceitos e características das publicações de dados, com foco em Publicações Ampliadas, *Data Papers* e *Data Journals*.

43 O DOI é um identificador único para objetos digitais.

envolvidos e por diferentes reivindicações. Inclusive, chamar de “movimento” pode gerar certos problemas considerando as várias intersecções com outras iniciativas que se desenvolveram em paralelo com histórias e objetivos próprios.

Como uma abordagem para o desenvolvimento da pesquisa científica, a Ciência Aberta constitui-se em muitos aspectos como uma alternativa aos modelos de pesquisa centrados no fechamento e à apropriação privada dos resultados científicos. Em tal aspecto, essa abordagem toma como base as normas sociais que historicamente marcam a ciência como uma organização pautada pela comunicação pública dos resultados da pesquisa. A Ciência Aberta também possui interface com a cultura do Movimento do Software Livre e de Código Aberto, estabelecendo formas de produção de conhecimento que se baseiam na colaboração e no compartilhamento aberto.

As políticas e reivindicações recentes que pautam as discussões e desenvolvimentos relacionados à Ciência Aberta têm enfatizado o compartilhamento de dados de pesquisa como centrais para a constituição de uma ciência verdadeiramente aberta. A crescente centralidade dos dados nesse movimento chama a atenção para os potenciais de reutilização dos dados para os mais diferentes objetivos, o que tem levado ao desenvolvimento de novos suportes de comunicação, ferramentas de análise e gestão dos dados, novas formas de avaliação científica, papéis e práticas de pesquisa voltados para a comunicação e reutilização dos dados. A ênfase nos dados redireciona não só as pautas do movimento de Ciência Aberta, como reconfigura a forma como a ciência se organiza e como é praticada. Um aspecto fundamental dessas transformações, compreender as reconfigurações que permeiam as práticas científicas.

Esta tese busca investigar tais reconfigurações, utilizando uma abordagem teórica e uma pesquisa etnográfica. O próximo capítulo delinea as bases teóricas e a metodologia da pesquisa desenvolvida.

2 ESTUDANDO A CIÊNCIA ABERTA: ABORDAGEM TEÓRICA E METODOLÓGICA

Este capítulo explicita o ferramental teórico e analítico, bem como os procedimentos metodológicos que orientam a pesquisa de campo que forma a base empírica das análises desenvolvidas nos demais capítulos da tese. O capítulo descreve os pressupostos teóricos dos Estudos Sociais da Ciência e da Tecnologia que tratam das comunidades epistêmicas, vinculado a ideia de que as comunidades científicas são divisões culturais. Em seguida, apresento o referencial teórico dos Estudos de Infraestrutura como instrumentos teóricos que possibilitam analisar as infraestruturas que dão suporte ao trabalho acadêmico. Em seguida, explico a importância do método etnográfico para o tipo de pesquisa empreendida. Ao final, detalho os instrumentos utilizados no trabalho de campo, que consistem em observações *in loco*, entrevistas, entre outras interações com o NeuroMat.

2.1 Culturas epistêmicas e prática científica

Estudar as práticas científicas nas recentes correntes teóricas dos Estudos Sociais da Ciência e da Tecnologia significa dirigir a análise para as formas de interpretação e para os aspectos relacionais e situados das práticas dentro do laboratório. Quando falamos em “formas de interpretação”, estamos desconsiderando a ciência como um empreendimento unitário no qual as mesmas estruturas e normas funcionam universalmente para todas as disciplinas e domínios científicos, tratando a ciência e o saber dos peritos em sua dimensão cultural, o que significa considerá-los como territórios de “divisões culturais” nas quais habitam grupos de peritos, delimitados por conjunto de padrões e dinâmicas que estão em jogo na prática especializada e que variam de acordo com a disposição de cada expertise (KNORR-CETINA, 1999b). A autora vai insistir na desunião da ciência e do conhecimento e na necessidade de nos defrontarmos com essa separação, reconhecendo aquilo que chama de “culturas epistêmicas”. O epistêmico é aquilo que se relaciona com a verdade, a facticidade e a objetividade na ciência.

As culturas epistêmicas referem-se às maquinarias e estratégias para estabelecer verdades, o que envolve uma diversidade de artifícios e estratégias retóricas de persuasão por meio da comunicação:

As culturas epistêmicas são culturas da criação e da confirmação do conhecimento. São as culturas dos cenários de conhecimento: os padrões agregados e as dinâmicas que se exibem na prática dos peritos e que variam consoante os diferentes cenários de saberes de peritos (KNORR-CETINA, 1999b, p. 377).

Nesse sentido, as culturas epistêmicas referem-se ao conjunto de padrões e dinâmicas que estão em jogo na prática especializada e que variam de acordo com a disposição de cada expertise, caracterizando-se em termos dos objetos que estão sendo constituídos. Portanto, podem ser entendidas como a reunião dos arranjos e mecanismos interligados pela afinidade, necessidade e coincidência histórica, que em determinado domínio formam “como sabemos o que sabemos”.

Tomando como exemplo as ciências naturais para refletir sobre o aspecto cultural da ciência, Knorr-Cetina (1999a) argumenta que essas ciências – tidas como empreendimentos homogêneos – são constituídas, na realidade, como monopólios epistêmicos diversificados e independentes que adotam múltiplos procedimentos científicos e produzem uma sucessão variada de produtos, resultados e tipos de conhecimentos unificáveis. Na prática, as diferenças entre tais monopólios ocorrem por meio da diversidade de relações estabelecidas por essas culturas com seus objetos de estudo e instrumentos, pelo significado assumido pelo empírico e as relações sociais estabelecidas na atividade científica em um determinado campo (KNORR-CETINA, 1999a). As divisões são constituídas nos diferentes níveis da formação dos peritos, por meio das organizações de pesquisa, na seleção da carreira e nos sistemas de organização e classificação que diferenciam as várias disciplinas e domínios do conhecimento científico.

Toda cultura epistêmica guarda uma relação particular com a prática. Desse modo, as divisões se tornam visíveis quando olhamos para a ciência como uma prática ou experiência e comparamos os processos de conhecimento em diferentes domínios. Essas são práticas situadas, e não necessariamente equivalentes às metodologias dos manuais, podendo mudar conforme os contextos (WOUTERS; BEAULIEU, 2006).

Ao trazer esta perspectiva para analisar a circulação do conhecimento científico, por exemplo, os contrastes em relação aos pressupostos de que existe uma norma da comunicação pública da ciência praticada por todos se mostram visíveis. O enquadramento das práticas de comunicação na ciência em termos binários, isto é, entre um pesquisador (a origem) e uma coletividade abstrata (a comunidade científica), tem sido posto em discussão pelas correntes atuais dos ESCT, sobretudo pelas correntes que buscam revelar a complexidade das redes em que a ciência contemporânea se insere.

Knorr-Cetina (1999a) afirma que a comunicação está implicada em todas as dimensões do trabalho científico e, em particular, podemos identificar a sua relevância nos processos de formação de consenso, o que inclui as formas de circulação da informação por meio de um processo de interpretação e mecanismos de persuasão e negociação através da comunicação oral ou escrita, ou, de maneira latente, nas formas de retenção seletiva da informação. A comunicação não se reduz somente a uma relação entre duas partes (cientista e sua comunidade), mas está inserida em um conjunto de relações em que há diversos atores humanos com interesses e práticas diversos, e também não-humanos que inserem dinâmicas e propriedades específicas no espaço no qual a ciência é performada.

De acordo com Hilgartner (1997), nos regimes de comunicação da ciência, as formas de assegurar a exclusividade e prioridade da pesquisa passam por um processo dialético de revelação e ocultação dos resultados da pesquisa, em outras palavras, de abertura e fechamento de tais resultados, existindo compreensões diferentes entre os cientistas não somente sobre o que abrir, mas também como, quando e com quem abrir. Além disso, há também perspectivas diversificadas do que significa de fato abrir a pesquisa. Ocorrem também relações porosas e complementares entre as categorias “aberto” e “fechado”, de maneira que tais categorias frequentemente se misturam, mesclando em momentos determinados o total secretismo e o acesso aberto seletivo aos produtos e processos da ciência (HILGARTNER; BRANDT-RAUF, 1994; HILGARTNER, 2012, 1997).

Pensemos então os dados de pesquisa. Dificilmente poderemos afirmar que os dados tornados abertos e públicos gerados em um desenvolvimento científico representam a totalidade do que poderia ser revelado e publicado. O mesmo pode ocorrer em relação a desenvolvimentos científicos supostamente secretos ou fechados.

Hilgartner (1997) propõe um modo de captar as formas de relacionamento dos cientistas com o compartilhamento de dados é utilizar uma perspectiva atenta aos “fluxos”

desses dados. A ideia de fluxo significa algo que pode escoar de um ponto a outro, pode ser “caudaloso”, “fino” ou pode ser interrompido por diferentes razões e em momentos variados. Nesse aspecto, a função de quem estuda essas práticas é tentar identificar as partes de um determinado fluxo e como elas são normalmente distribuídas, para quem, com quais condições e como as “porções” de dados são demarcadas. Dito de outro modo, buscamos identificar como os dados, geralmente entidades “reclusas” ou “discretas”, nas palavras do autor, são extraídos dos fluxos em constante evolução da produção científica e posteriormente inseridos em relações de intercâmbio e quais as relações estabelecidas em torno deles.

Considero que esse modo de olhar permite compreender e problematizar argumentos centrais que permeiam o movimento de Ciência Aberta, como avaliações otimistas sobre a abertura de resultados de pesquisa. Ainda que avaliações que olhem para os aspectos positivos sejam parcialmente corretas sobre os potenciais da Ciência Aberta, muitas vezes essas visões perdem a noção de questões cruciais sobre a complexidade dos processos envolvidos. Um dos problemas das visões triunfalistas é que muitas vezes acabam por reforçar uma percepção de que a ciência progride em bloco e funciona a partir de mecanismos universais homogêneos, ofuscando o caráter contextual das práticas científicas.

Algumas avaliações – como a ideia reiterada de que a abertura nos termos da Ciência Aberta se traduz em benefício para sociedade e, portanto, os cientistas devem disponibilizar os resultados de pesquisa abertamente, muitas vezes a qualquer custo – parecem desconsiderar que os aspectos individuais dos pesquisadores, as amarras sociais, técnicas e organizacionais frequentemente sobrepujam o que supostamente é de interesse da sociedade (FRIESIKI et al, 2015). É importante também ponderar que a decisão por abrir a pesquisa pode ocorrer por motivações nem sempre altruístas ou de interesse coletivo, assumindo um caráter instrumental/utilitarista e circunstancial para o pesquisador (LEVIN et al, 2016).

Acerca disso, Hilgartner (1999, p. 31), em uma interlocução com Knorr-Cetina, complementa:

Cada campo científico tem suas próprias convenções sobre o que constitui um documento publicável e o que constitui um resultado interessante. Dentro de comunidades particulares de pesquisadores, essas convenções podem ser claramente compreendidas e, de fato, parecem óbvias para os participantes. No entanto, as convenções não são idênticas em todos os campos nem são totalmente estáveis. Para fornecer informações gerais sobre como os dados são compartilhados ou, de outro modo, trocados entre cientistas, não se pode simplesmente assumir que as convenções de uma única comunidade de pesquisa constituem a única maneira de conduzir a ciência; Em vez disso, as convenções de diferentes comunidades de pesquisa se tornam fenômenos para os analistas sociais explicarem.

Algo que reforça essas afirmações é que mesmo diante da proliferação de manifestos e políticas voltadas para a Ciência Aberta (e o tom muitas vezes apologético de seus entusiastas), as avaliações têm mostrado que os pesquisadores não aderem facilmente e na proporção desejada às propostas de abrir os resultados nos termos do movimento⁴⁴.

Não é intenção da tese adentrar somente motivações dos pesquisadores em disponibilizar os resultados de pesquisa, até porque as razões que levam a esse cenário são bastante variadas. Mesmo as afirmações recorrentes de que a aceitação das diferentes formas de abertura (como as publicações de artigos acesso aberto) passa por reformulações do sistema de incentivos, recompensa e pela capacidade desse modelo de oferecer reputação, prestígio e capital simbólico não parecem dar conta da complexidade e da multiplicidade de relações que se estabelecem em torno das formas de circulação do conhecimento científico (DELFANTI; PITRELLI, 2015; CAULFIELD et al, 2012).

Portanto, em vez de ter um significado “incontestável”, os indivíduos diferem na maneira como interpretam a norma de abertura dos resultados de pesquisa e sobre os melhores procedimentos e formas de praticá-la. Assim, as escolhas sobre o que, como, quando e para quem tornar a pesquisa aberta é uma questão altamente situada (DASTON, 1995; LEVIN; LEONELLI, 2016).

Alguns autores dos ESCT têm se dedicado a compreender e teorizar sobre as práticas contemporâneas de compartilhamento na ciência. Esses estudos sugerem que a forma como os cientistas se relacionam com a abertura ou o compartilhamento dos seus resultados vai depender dos objetivos, preferências, restrições, configurações institucionais, dos aspectos econômicos envolvidos nos custos de manutenção de infraestruturas técnicas, da existência ou

⁴⁴ Cf. Levin et al (2016).

não de incentivos para compartilhar os resultados, da incerteza em relação à confiabilidade das contribuições produzidas por colegas e interesses incongruentes, normas, conceitos, padrões tecnológicos, ou mesmo formas e níveis de conhecimento entre participantes de uma comunidade (LEVIN; LEONELLI, 2016; MILLERAND, 2011, 2012; HINE, 2006; EDWARDS et al, 2011; HILGARTNER; BRANDT-RAUF, 1994; LEONELLI, 2016; BEAULIEU, 2001; BRUYNINCKX, 2015; DASTON, 1995)⁴⁵.

Sem deixar de lado tendências otimistas, inclusive porque essas abordagens otimistas também “fazem” a Ciência Aberta como fenômeno (WOUTERS; BEAULIEU, 2006), outros significados precisam ser analisados, inclusive reconhecimento a Ciência Aberta, como vários outros movimentos com propostas semelhantes, trilha seu caminho com muitas contradições (DELFANTI; PITRELLI, 2015; CAULFIELD et al, 2012).

2.2 Estudos de Infraestruturas

As análises e teorizações desenvolvidas nesta pesquisa dialogam com os teóricos dos Estudos Sociais da Ciência e da Tecnologia que buscam compreender como as infraestruturas organizam a circulação do conhecimento na sociedade (EDWARDS, 2010; BOWKER et al, 2010; BORGMAN et al, 2013).

O termo infraestrutura é empregado com uma ampla variedade de significados. Frequentemente refere-se a grandes serviços, equipamentos coletivos e sistemas duráveis, que funcionam “bem”, desde ferrovias, estradas, telecomunicações, sistema de transporte e abastecimento. Aparecem também com grande frequência nos discursos políticos e econômicos associadas a soluções abrangentes, de grande porte e indispensáveis a vida das pessoas. Ao mesmo tempo, é associado a problemas onipresentes, que demandam sempre soluções e desenvolvimentos adicionais e permanentes (EDWARDS et al, 2009; BOWKER et al, 2010).

⁴⁵ Destaco a importância dos trabalhos desenvolvidos por Sabina Leonelli (2013, 2017) Florence Millerand (2011, 2012) e Alessandro Delfanti (2013) para análise de práticas de Ciência Aberta em uma perspectiva ESCT.

Para Star e Ruhleder (1996), as definições mais comuns apresentam as infraestruturas como um substrato, isto é, uma coisa sobre a qual algo “corre” ou “opera”, algo construído e mantido, que emerge de um substrato invisível, como no caso do sistema de trilhos de trem sobre os quais correm os vagões. Entretanto, essa imagem não é útil ou suficientemente precisa para uma compreensão acurada das infraestruturas.

Os estudos sociais de infraestruturas ampliaram o significado e formaram uma compreensão teórica das infraestruturas como sistemas sociotécnicos que possibilitam a circulação de bens, conhecimento, significados, pessoas e formas de poder (LOCKREM; LUGO, 2012). Tal compreensão é grande devedora do trabalho do historiador da tecnologia norte-americano Thomas Hughes que, estudando o sistema elétrico, contribuiu para uma compreensão das infraestruturas como arranjos materiais e heterogêneos que convivem em uma espécie de relação intrínseca e inseparável:

Os sistemas tecnológicos contêm componentes confusos e complexos de solução de problemas. Ambos são socialmente construídos e modeladores da sociedade. Entre os componentes dos sistemas tecnológicos estão os artefatos físicos, como os turbogeradores, transformadores e linhas de transmissão em sistemas de energia elétrica e luz elétrica. Os sistemas tecnológicos também incluem organizações como empresas de manufatura, empresas de serviços públicos e bancos de investimento, e incorporam componentes geralmente rotulados como científicos, como livros, artigos e programas universitários de ensino e pesquisa. Os artefatos legislativos, como as leis reguladoras, também podem fazer parte dos sistemas tecnológicos. Por serem socialmente construídos e adaptados para funcionar em sistemas, os recursos naturais, como as minas de carvão, também se qualificam como artefatos do sistema (HUGHES, 1987, p. 51).

Hughes chama atenção principalmente para esse caráter interdependente entre diferentes artefatos físicos e não físicos que compõem um sistema, os quais contribuem diretamente ou por meio de outros componentes para o objetivo comum do sistema, de forma que se um componente for removido ou se suas características se modificarem, os outros elementos do sistema reconfigurarão, conseqüentemente, suas características (HUGHES, 1987).

Nesta perspectiva, a infraestrutura é sempre e fundamentalmente produto de relações complexas e heterogêneas, nunca uma coisa ou “o que”, mas é sim sempre uma relação que sustenta o trabalho diário de alguém (STAR, 1999). Dessa forma, infraestrutura não implica apenas o substrato por onde os bens circulam, mas todos os elementos

mobilizados para que eles possam circular: pessoas, sistemas técnicos, padrões, leis, processos burocráticos, manuais, documentação, entre outros.

Em razão desse caráter relacional, a abordagem analítica reforça o caráter modular, de múltiplas camadas e de elementos heterogêneos, estáticos e dinâmicos, que se interligam no momento em que a infraestrutura é acionada:

As infraestruturas não são sistemas, no sentido de processos totalmente coerentes, deliberadamente projetados, de ponta a ponta. Em vez disso, as infraestruturas são ecologias ou sistemas adaptativos complexos; elas consistem em inúmeros sistemas, cada um com origens e objetivos únicos, que são feitos para interoperar por meio de padrões, camadas de conectores, práticas sociais, normas e comportamentos individuais que suavizam as conexões entre eles (EDWARDS et al, 2013, p.5).

A compreensão da natureza da infraestrutura proposta aqui envolve desdobrar as escolhas políticas, éticas e sociais que são feitas ao longo do seu desenvolvimento, o que, analiticamente, consiste em ir aos bastidores, mudando a ênfase dos componentes materiais para as relações que formam tais infraestruturas (BOWKER et al, 2010).

Esta opção de aproximação pode ser alcançada por meio de um processo fundamental de “inversão infraestrutural” (STAR, 1999). A inversão de infraestrutura consiste em um modo de olhar que não vai enfatizar as coisas ou as pessoas como as únicas causas das dinâmicas, mas procura concentrar-se nas relações heterogêneas que formam a infraestrutura. Nesse sentido, inverte-se as explicações históricas tradicionais e revela-se como as escolhas e políticas embutidas nesses sistemas se tornam componentes articulados (a exemplo das ferrovias, horários e estruturas administrativas nas burocracias formando o sistema ferroviário).

Star e Ruhleder (1996, p.253) afirmam que com esta maneira de olhar “Substrato se torna substância” e ao contrário de nos perguntarmos o que é realmente uma infraestrutura, o processo de inversão direciona a análise para examinar um fenômeno *como* infraestrutura, isto é, como algo relacional e fazendo um questionamento fundamental: o que está apoiando o trabalho de outro e quem (ou o que) está sustentando esses relacionamentos? Perceberemos que haverá uma sucessão de múltiplas relações de elementos.

2.2.1 Infraestruturas de informação e as práticas científicas emergentes

Nas análises e reflexões desenvolvidas nesta tese a preocupação se volta principalmente para as infraestruturas de informação que dão apoio ao trabalho acadêmico. Essa abordagem analítica trata as infraestruturas acadêmicas de produção de conhecimento e informação como uma rede heterogênea de atores – ou para utilizar a terminologia comum dos ESCT, de atores humanos e não-humanos – que viabilizam a produção, documentação e a circulação da informação sobre o mundo humano e natural.

Como as demais infraestruturas, elas são compostas de muitos sistemas técnicos e redes interligadas que representam elos e camadas de sistemas e estruturas, cada uma com sua própria dinâmica única, servindo a uma função específica, integrada com base em indivíduos, organizações, rotinas, normas compartilhadas, padrões e sistemas técnicos (EDWARDS et al, 2013; EDWARDS, 2007).

Os estudos das infraestruturas de informação – fortemente inspirados no trabalho seminal de Susan Leigh Star (STAR; RUHLEDER, 1996; STAR, 1999) – voltam-se para os aspectos materiais dos sistemas informacionais. As infraestruturas de informação são compreendidas de maneira ampla, desde as instalações físicas, como bibliotecas, arquivos, museus e, atualmente, os serviços digitais, geralmente associados à Internet, os serviços computacionais, repositórios de dados, assim como os sistemas de informação e comunicação a nível global que processam e transportam dados dentro e fora das fronteiras nacionais (BOWKER et al, 2010).

Historicamente, a ciência funciona com uma organização altamente distribuída, dependente de lugares, objetos, instituições e pessoas: universidades, órgãos de fomento à pesquisa científica e tecnológica, sociedades científicas, periódicos, livros, amostras, editores científicos, laboratórios, equipamentos, estações de pesquisa, bibliotecas, arquivos de dados, além das tecnologias de informação em rede que apoiam atividades de pesquisa científica, como colaboração, compartilhamento de dados e disseminação de descobertas, assim como os profissionais associados envolvidos (RIBES; LEE, 2010). Essa variedade de entidades e suas práticas têm sido ligadas por meio de diversos tipos de infraestruturas de informação – variando das redes científicas do século XVIII baseadas na circulação impressa para formas computacionais avançadas existentes atualmente. Como vários estudiosos argumentam, são essas infraestruturas de informação – não relatos ilusórios de um “método científico

unificado” – que diferenciam as ciências de outras formas de produzir e organizar o conhecimento (BOWKER et al, 2010; RIBES et al, 2013).

As práticas de produção científica contemporâneas são fortemente dependentes da disponibilidade das infraestruturas de informação, estando associadas a uma dimensão coletiva, representada pelos grupos e redes de pesquisa globais, ao surgimento de novas disciplinas, aos desenvolvimentos científicos e tecnológicos de fronteira (KNORR-CETINA, 1999a).

As infraestruturas de informação contemporâneas, por sua vez, têm representado uma mudança importante na organização da ciência – apoiadas por políticas governamentais de ciência e tecnologia e por financiadores nacionais e internacionais – que empregam intensamente tecnologias para alavancar recursos computacionais para avançar em escala, natureza e eficiência (RIBES et al, 2013). Podemos citar os supercomputadores, as redes computacionais compartilhadas (GRID), os biobancos, repositórios e bibliotecas digitais de conjunto de dados, códigos-fonte e textos acadêmicos, computação em nuvem e os laboratórios virtuais.

As atuais infraestruturas de informação, como a *e-science*, impactam de maneira significativa as formas de produção de conhecimento, principalmente os fluxos de informação e a colaboração científica, à medida que muitos dos principais processos e funções organizacionais da pesquisa são transferidos para essas infraestruturas tecnológicas, possibilitando, por exemplo, a reunião de grupos de pesquisadores dispersos geograficamente e também concedendo aos pesquisadores acesso a vastos conjuntos de dados federados ou centros de computação distribuída sem passar por controle humano ou processo administrativo (RIBES et al, 2013).

Por outro lado, os elementos que permitem colocar as infraestruturas de *e-science* como um agente de transformação organizacional da ciência não se resumem aos sistemas técnicos avançados, associados a serviços digitais e computacionais, mas sobretudo à associação de elementos, geralmente invisibilizados, que possibilitam que essas infraestruturas funcionem da maneira “revolucionária”. Isso fica evidente quando analisamos as recentes infraestruturas para compartilhamento de dados.

Como parte dos novos métodos de produção de conhecimentos baseados, por exemplo, na modelagem com a finalidade de predição, simulação, representação ou recriação dos sistemas naturais por meio de uma série de novas formas de coleta, compilação e análise de dados, as infraestruturas de compartilhamento demandam desenvolvimentos, preocupações e formas de organização distintos, novos nodos, atores, papéis e competências para viabilizar essas práticas.

Quais compreensões ou interpretações podemos oferecer sobre a Ciência Aberta a partir dos estudos de infraestruturas de informação? Levin et al (2016) vai argumentar que o desenvolvimento científico intensivo em dados e sua natureza computacional faz com que o compartilhamento de infraestrutura seja parte de normas e competências institucionalizadas das práticas científicas emergentes.

Leonelli (2016), analisando o desenvolvimento de infraestruturas de dados na área de biologia vegetal, argumenta que a tendência ao compartilhamento de dados se fortalece cada vez mais em razão de que os grupos de pesquisa isolados dispõem de recursos financeiros, humanos e tecnológicos insuficientes para tratar o grande volume de dados gerados pela ciência. Além disso, argumenta a autora, a relevância e o status que os dados adquiriram como mercadoria tanto dentro quanto fora das comunidades científicas demandam que eles estejam amplamente disponíveis, o que requer uma infraestrutura robusta, articulada e “à mão” para viabilizar seus usos em diferentes contextos.

Nesse cenário, a constituição de bases comuns de compartilhamento de dados de pesquisa emerge como recurso essencial para a pesquisa acadêmica ou como “ferramenta de pesquisa”, configurando-se como infraestrutura a partir da relação entre os repositórios disciplinares públicos online, políticas de C&T voltadas para a construção desses ambientes, desenvolvimento e disponibilização de ferramentas computacionais em código aberto, normas, processos de padronização, modelos de compartilhamento, modelos de compartilhamento e reutilização, e trabalhadores do conhecimento necessários para a manutenção dessa infraestrutura.

Edwards et al (2011) afirma que os dados de pesquisa se movem entre pessoas, organizações, laboratórios equipamentos, de uma disciplina para outra e de determinados formatos para outros, produzindo fricções, consumindo esforços, gerando conflitos e discordância, produzindo extensos processos de negociação e soluções (socio)técnicas que buscam reduzir esse atrito – por exemplo, padrões de metadados.

As infraestruturas reforçam também a atuação dos indivíduos que ocupam funções existentes ou emergentes, geralmente viabilizando o seu funcionamento. Entre esses profissionais estão os designers e desenvolvedores, usuários e mediadores, gerentes, administradores de sistemas, curadores de dados, bibliotecários e outros atores que viabilizam as infraestruturas que sustentam o trabalho acadêmico (LEONELLI, 2016; BOWKER et al, 2009; HINE, 2006, 2008). Esses profissionais trabalham geralmente em associação com os cientistas, em funções de desenvolvimento, implementação e suporte para que a infraestrutura funcione. Dessa forma, tornam-se fundamentais para as práticas científicas habilitadas pelas infraestruturas de informação (BORGMAN, 2012; LEVIN; LENELLI, 2016; MILLERAND, 2012).

É importante pontuar também que, analiticamente, a infraestrutura aparece como uma propriedade relacional, não como uma coisa despojada de uso (STAR; RUTHLEDER, 1996). Nesse sentido, compreender como algo associado ao trabalho possibilita descrever as dinâmicas sobre compartilhamento científico que se assentam em infraestruturas de informação, assim como um olhar para as próprias infraestruturas como objetos que são construídos e performados na e pela prática científica. Em vez de olharmos para o técnico e o social como partes independentes ou como dicotomias, buscamos entendê-los como elementos que se associam, assim como as relações de contradições que existem entre ambos, resultado das “ações” que condicionam o comportamento da infraestrutura.

Essa maneira de olhar exige pensar para além das linguagens instrumentais de utilidade e função das ferramentas, que tendem a tratá-las como instrumentos neutros, reduzindo seus significados a efeitos negativos e positivos e apenas em relação à sua operacionalização e uso (correto ou errado). A esse respeito, Hine (2006) aponta que os efeitos das ferramentas de compartilhamento, tais como banco de dado, dependem dos processos sociais específicos, circunstâncias particulares e contextuais em que as ferramentas são concebidas e de como são colocadas em funcionamento (HINE, 2006). Essa chave de análise passa a considerar as infraestruturas de informação em toda a gama de efeitos e dimensões, não somente como algo que acarreta ganhos, mas também perdas, custos e ajustes decorrentes dos limites e tensões que sua instalação e uso geram. Algumas das tensões mais “espinhosas” ao redor, através e dentro da infraestrutura não são estritamente de ordem técnica, mas também de natureza legal, econômica e tudo mais que pode ser taxado como “social” (EDWARDS et al, 2009).

Ao contrário de assumirmos um valor intrínseco ou neutro da infraestrutura, nos movemos buscando entender seus efeitos não como coisas neutras e uniformes, ou na dicotomia positivo-negativo, mas procurando identificar as (re)configurações das formas de produção de conhecimento científico, especificamente as repercussões sobre o comportamento dos atores, modos de sociabilidade/interação entre os cientistas, as formas de organização e os fluxos de informação, assim como as associações de elementos que formam a sua substância.

Por meio da perspectiva teórica aqui explicitada, conforme propõe Edwards et al (2009), é importante olhar para essas tensões e limites como um lado das possibilidades/configurações associadas às infraestruturas e os potenciais que emergem com seu uso. Assim como consideramos que quando novas formas de produção de conhecimento “melhores” e “maiores” emergem, outras podem ser diminuídas ou substituídas, assim como os esforços em direção às outras formas de acesso ao conhecimento para um grupo podem representar a diminuição ou restrição do acesso para outro – ainda assim podem representar uma evolução em termos de produção de resultados “estimulantes” e avanço do ponto de vista do trabalho científico tradicional.

Um outro aspecto se refere às mudanças organizacionais associadas a essas infraestruturas e às (re)configurações tanto do ponto de vista do comportamento social como da organização do trabalho científico associados a elas (LYNCH, 1993).

Ao passo que são inseridas na prática, as ferramentas que surgem da cultura da Internet e da computação, por várias de suas características, produzem novas mediações, podendo reconfigurar os padrões de produção de conhecimento tornando espacialmente e socialmente complexa a produção de ciência, especialmente no que diz respeito às formas de comunicação do conhecimento científico, as práticas de investigação, a organização espacial da pesquisa, a distribuição de papéis, a representação do conhecimento, o controle de qualidade, o reconhecimento e alocação de crédito acadêmico, os aspectos econômicos e jurídicos em torno da produção de conhecimento, etc (RIBES et al, 2013; HINE, 2006).

2.3 Etnografia da Ciência e da Tecnologia

A pesquisa combinou a análise bibliográfica com uma pesquisa de campo do tipo etnográfica para abordar as práticas de Ciência Aberta no contexto do Centro de Pesquisa, Inovação e Difusão em Neuromatemática (NeuroMat). A etnografia objetivou detalhar e compreender as condições locais de produção de Ciência Aberta no NeuroMat, atentando para as práticas de uma variedade de atores humanos e elementos relacionados à produção da “abertura” nesse espaço.

A etnografia é um recurso metodológico e analítico que incorpora uma variedade de métodos participativos e observacionais, concentrando-se em descobrir conhecimentos e comportamentos geralmente negligenciados, mundanos e cotidianos que possam representar o funcionamento interno de uma comunidade.

Entende-se que a tarefa de entender como opera uma comunidade e como os membros sabem sobre seu funcionamento, sua linguagem e normas, e os recursos de contextualização dessas representações dentro de uma estrutura é um desafio que apresenta limitações, o que requer o emprego de determinados instrumentos para a aquisição e processamento de uma quantidade considerável de dados (JOHNSON, 2018). Nesse aspecto, Hine (2016) afirma que empreender um esforço de criar narrativas representativas dos grupos exige uma experiência que permita acessar o universo dos sujeitos de maneiras únicas. Nesse aspecto, o “ir a campo da etnografia” possibilita uma imersão nesse universo.

Star (1999) compreende que a etnografia permite uma sensibilidade para a coleta de dados e análise, considerando que os sujeitos produzem significados com base nas suas circunstâncias, e que tais significados são inscritos em suas ações e nas avaliações que fazem em relação à infraestrutura. Esta sensibilidade possibilita sobretudo trazer à tona as vozes silenciadas, as diferentes estratégias discursivas e compreender as diferenças entre palavras e as ações⁴⁶.

46 São evidentes os pontos de contato entre os Estudos de Infraestrutura com a perspectiva das “redes sociotécnicas” ou, mais precisamente, com a corrente teórica da Teoria Ator-Rede (ANT). Para a ANT não há uma explicação definida de antemão para as redes, mas sim uma busca constante por identificar os elementos que a compõe por meio de observação *in loco*. Assim, deve-se seguir os atores, acompanhá-los por onde quer que eles se desloquem, saber com quem eles andam, questioná-los incansavelmente (LATOURET, 2012). Geiger e Ribes (2011) discorrem sobre essa relação apontando que os teóricos da ANT frequentemente realizam um tipo similar de inversão de infraestrutura ao aceitar a reconhecida injunção de Bruno Latour de “seguir os atores”, na qual a categoria “atores” é definida de maneira bastante ampla. Susan L. Star, principal representante dos Estudos de Infraestrutura, desloca o modelo de interesses presentes na ANT apontando a necessidade de uma abordagem “ecológica” de sistemas tecnológicos, no

O método etnográfico ocupa um espaço importante nos ESCT contemporâneos⁴⁷, sendo colocado como um poderoso instrumento teórico e analítico para rastrear e revelar os intrincados padrões de conexão da ciência contemporânea e permitindo uma sensibilidade para a complexidade e significações das práticas que não estão evidentes no discurso ou em documentos (HINE, 2007). Especialmente nos estudos de laboratório⁴⁸, como ferramenta metodológica e analítica, a etnografia procura enfatizar o processo da produção e estabilização de controvérsias na produção do conhecimento científico. Nesse aspecto, busca compreender como os atores da ciência “manuseiam” os objetos, realizam experimentos e refletem sobre práticas e, principalmente, como o social é constitutivo do processo de produção da ciência – que tende a se apresentar como neutro ou livre de influência externa (PREMEBIDA et al, 2011).

No contexto dos estudos de laboratório, o trabalho etnográfico conduz o analista a olhar de perto o conhecimento científico e os artefatos técnicos em seu caráter processual e construído, seu contexto e sua materialidade negociada entre uma diversidade de atores humanos e não-humanos (MONTEIRO, 2012).

Analisar a infraestrutura e as práticas científicas requer a navegação entre diferentes camadas: material, humana e a dos dados coletados. Um dos desafios para o estudioso das infraestruturas e das práticas científicas é encontrar formas de transitar entre essas diferentes camadas, quais instrumentos utilizar e que recortes estabelecer. É necessário considerar essas diferentes camadas no momento da implementação dos estudos de infraestrutura tendo em vista que o trabalho científico continua em muitos lugares ao mesmo tempo, ramificando-se além das paredes do laboratório (LATOURET; WOOLGAR, 1997).

A reticulação do trabalho acadêmico parece ainda mais evidente e complexa quando são consideradas as práticas científicas sustentadas em infraestruturas computacionais. Nesse sentido, Geiger e Ribes (2011) e outros teóricos das infraestruturas como Star (1999) reforçam que a navegação pela infraestrutura envolve uma metodologia de decodificação ou inversão, colocando em primeiro plano os elementos verdadeiramente dos bastidores. Isso significa tornar visível o invisível e explorar aquilo que parece naturalizado, o

sentido em que a rede é formada por uma heterogeneidade de atores e objetos, cada um configurando um universo de mundos sociais distintos, existindo de “n-modos”. Nessa rede de relações, a colaboração entre atores heterogêneos e mundos sociais distintos somente se torna possível a partir da padronização ou estandardização de métodos e do desenvolvimento de “*boundary objects*” ou objetos de fronteira, isto é, elementos adaptáveis a diferentes pontos de vista e robustos o suficiente para manter a identidade entre esses mundos sociais (STAR; GRIESEMER, 1989, p.389).

47 Para uma discussão sobre a etnografia como método nos ESCT Cf. Monteiro (2012).

48 Cf. Premebida et al. (2011); Monteiro (2012).

não dito, ou como Star (1999) escreve: estudar o não estudado, o que nos parece entediante, “chato” da infraestrutura.

Ao realizar esse esforço de inversão, em vez de nos restringir ao que está visível, a análise se desloca entre diferentes camadas visíveis e invisíveis, variáveis dependentes e independentes e “relações entrelaçadas com o pensamento e trabalho”, nos possibilitando assim formas de apurar qualitativamente os atores individualmente na sua atuação dentro de ambiente social mais amplo (STAR, 1999).

Desse modo, a etnografia como instrumento metodológico e analítico permite o trânsito entre essas diferentes camadas, servindo a esse processo de inversão infraestrutural, uma vez que o etnógrafo coloca em primeiro plano elementos dos bastidores, visibilizando elementos escondidos por detrás da infraestrutura, concentrando-se em desvendar conhecimentos e comportamentos negligenciados, vozes silenciadas, pontos de vista, os elementos da prática, do suporte tecnológico, do “chão de fábrica” – geralmente ignorados – que sustentam as atividades cotidianas. Essa tarefa exige do etnógrafo, portanto, a disposição para transitar por múltiplos contextos e atento às diferentes situações (MARCUS, 1995; HINE, 2007; LATOUR; WOOLGAR, 1997), o que também requer uma combinação de análise histórica e literária, ferramentas tradicionais como entrevistas e observações, análise de sistemas, estudos de usabilidade e outros (STAR, 1999).

A etnografia que sustentou as análises e reflexões desenvolvidas por esta tese incorporou uma variedade de métodos participativos e observacionais que expusessem ou retratassem o funcionamento interno do centro de pesquisa em questão (JOHNSON, 2018). Entretanto, ao contrário de tentar estar em toda parte na rede de atores – conforme propõem as abordagens mais tradicionais dos ESCT, como a Teoria Ator-Rede – e reconhecendo certas limitações de tempo e de acesso a alguns dos sítios enfrentadas ao longo do trabalho de campo, utilizei a noção de “copresença” proposta por Beaulieu (2010), que descentraliza a noção de espaço físico e volta-se para a produção de conhecimento não baseada no laboratório, destacando o papel do etnógrafo como autor, participante-observador e estudioso que “faz o campo”.

Utilizando essa noção de copresença, situei-me deliberadamente nos momentos e lugares específicos em que a Ciência Aberta estivesse sendo projetada, construída, contestada, quebrada ou reparada. Dessa forma, dialogo com trabalhos como Geiger e Ribes (2011) sobre “etnografia dos rastros”. O recurso da etnografia dos rastros consiste em combinar a observação participante com a exploração dos registros ou rastros documentais em sistemas altamente mediados tecnologicamente (ambientes online, plataformas, softwares), permitindo transformar os “traços” deixados pelos atores em documentos e ambientes sociotécnicos em descrições densas sobre suas práticas e eventos que normalmente não estão visíveis em outras formas de pesquisa qualitativa, uma vez que ambientes não apenas documentam eventos e interações, como também são utilizados pelos indivíduos como locais de interação e *accountability* de suas atividades para o grupo⁴⁹.

Uma característica da etnografia dos rastros proposta por Geiger e Ribes (2011) – e que para a mim foi especialmente importante porque me chamou atenção para as muitas atividades do NeuroMat que envolvem interações em repositórios de código-fonte, Wikipédia, banco de imagens, nos sites, blog – é que essa abordagem vai realçar as práticas documentais dos atores nesses ambientes que representam protocolos ou padrões por meio dos quais as comunidades lidam com sua dispersão (geográfica) ou com práticas dispersas, harmonizando práticas e orientando comportamentos constitutivos dessa comunidade (vocabulários controlados e terminologias, ontologias, arquitetura de páginas web, etc).

Como será possível verificar na etnografia, as práticas de Ciência Aberta no NeuroMat passam pelo desenvolvimento de uma série de ferramentas e uso de padrões que, direta ou indiretamente, envolvem tais objetivos. Nesse sentido, essa abordagem dialoga com as propostas presentes em Hine (2000) que colocam os documentos estáticos como materiais a serem integrados à investigação à medida que possibilitam compreender as visões dos atores. Desse modo, os guias, manuais, biografias, artigos científicos e códigos podem ser concebidos como materiais etnográficos na forma como apresentam e moldam a realidade e são incorporados nas práticas.

49 Conforme propõe Geiger e Ribes (2011), a etnografia dos rastros representa, na prática, uma extensão da inversão infraestrutural que vai concentrar o trabalho etnográfico nos rastros deixados nas infraestruturas técnicas – *logs* que incluem histórico de versões, registro de data e hora, registros institucionais, transcrição de conversas e código-fonte alterações, revisões, comentários, cópias de backup. Ao colocarmos essas práticas documentais em primeiro plano, e uma vez decodificados sistematicamente, os rastros produzidos nesses ambientes auxiliam o pesquisador a reconstruir retroativamente ações específicas com razoável nível de granularidade, identificando as práticas de coordenação, os fluxos de informação, rotinas situadas e outros fenômenos sociais e organizacionais em uma variedade de escalas e, assim, obter detalhes sobre o modo de funcionamento de uma organização.

2.4 Detalhamento da Pesquisa de Campo

A pesquisa combinou uma extensa revisão bibliográfica com uma pesquisa de campo realizada no NeuroMat no período de fevereiro de 2015 a outubro de 2017. O NeuroMat é um centro de pesquisa localizado na Universidade de São Paulo e financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo. Caracteriza-se como um centro multidisciplinar, reunindo matemáticos, probabilistas, estatísticos, cientistas da computação e neurocientistas, além de uma equipe de profissionais que dá suporte às mais diferentes atividades. O centro é dedicado ao desenvolvimento de modelos matemáticos para a compreensão das bases funcionais do sistema nervoso.

O percurso de levantamento de dados contou com a participação em eventos, acompanhamento de reuniões internas e interação com os membros do centro de pesquisa por meio de entrevistas e diálogos em intervalos das atividades realizadas pelo centro. As informações coletadas nos eventos e entrevistas foram complementadas com informações e declarações veiculadas na mídia tradicional e em revistas especializadas em divulgação científica, assim como de páginas especializadas, juntamente com um vasto material disponibilizado no site do NeuroMat, como projetos de pesquisa e relatórios de atividades, seguindo o método da etnografia dos rastros descrita acima⁵⁰. Destaco ainda a leitura minuciosa de artigos e trabalhos acadêmicos produzidos pelos pesquisadores do centro ou citados durante as apresentações. Apesar da dificuldade em me apropriar e entender as linguagens especializadas desses sujeitos, a análise dessa produção também serviu à composição do relato etnográfico.

Desde 2015, também acompanho as postagens no site e na página do NeuroMat no Facebook, onde são divulgados os relatórios de pesquisa, artigos, apresentações, vídeos, eventos e outras informações. A *newsletter* divulgada no site do Centro também constituiu uma importante fonte de informações sobre as atividades desenvolvidas – passadas e futuras; divulgação da participação dos pesquisadores em atividades diversas; relação das publicações recentes e entrevistas; o perfil de pesquisadores do próprio centro ou de pesquisadores visitantes. Também recorri aos vídeos de divulgação e gravações das atividades disponibilizados no canal do centro no YouTube para rever alguns eventos que acompanhei

50 Também utilizei alertas do Google com as palavras “NeuroMat” e “neuromatemática”, o que possibilitou compor um arquivo de notícias e produções acadêmicas vinculadas ao centro.

pessoalmente como para coletar informações de atividades que não consegui participar in loco.

As fontes de informação online utilizadas estão documentadas no ANEXO B e serão citadas no decorrer do texto com uma codificação (F+número de identificação da fonte).

As discussões e entrevistas não somente apontaram nomes de pessoas que poderiam contribuir com a pesquisa, como fizeram referência a atores “não-humanos” como sites, softwares, bases de dados, repositórios, artigos, entre outros recursos de pesquisa. As práticas das áreas estão fortemente imbricadas com ferramentas computacionais. Em razão disso, busquei compreender a utilização das ferramentas nas práticas e também suas características e funcionalidades – por meio de acesso às páginas na *Web*, leitura de artigos especializados e da documentação das ferramentas com a finalidade de integrar essas informações à análise.

As buscas na *Web* e as consultas em recursos online mencionados nas entrevistas, palestras, relatórios, artigos e nas apresentações (*slides*) foram essenciais para o acesso, ainda que virtualmente, a determinados artefatos, como equipamentos experimentais, e visualização das representações dos processos de pesquisa a que não consegui ter acesso pessoalmente, além de facilitar o entendimento de termos que para mim eram bastante herméticos e, algumas vezes, incompreensíveis⁵¹.

Ao longo do trabalho de campo participei de workshops sobre modelagem de dados neuronais, eventos de difusão, seminários de pesquisa e apresentações de pesquisadores e de outros membros da equipe em diferentes tipos de eventos. A minha participação em algumas dessas atividades se deu a partir de recomendação ou convite dos próprios entrevistados que, no decorrer das entrevistas, ainda destacavam nomes de outros colegas cujas opiniões consideravam relevantes para o meu trabalho⁵².

Esse tipo de engajamento foi fundamental para um maior entendimento da dinâmica do centro de pesquisa, por possibilitar interagir com os sujeitos e, principalmente, porque representaram momentos em que a temática da Ciência Aberta emergiu como parte

51 As apresentações das pesquisas nos workshops são feitas geralmente com auxílio de imagens, que ilustram os protocolos, equipamentos e representações gráficas de organismos vivos. A interação dos pesquisadores com essas imagens foi de grande importância para o entendimento das práticas, possibilitando visualizar, ainda que com limites, as formas de experimentação, as circunstâncias em que esses dados são coletados e suas características.

52 A participação em alguns eventos e reuniões se deu mediante autorização de membros da equipe, visto que algumas dessas atividades não eram abertas à participação de pessoas externas.

dos debates em curso. Nessas reuniões, geralmente realizadas na sede do NeuroMat em São Paulo, a equipe discute as pesquisas em andamento; um pesquisador que acaba de ser integrado à equipe é apresentado; ou mesmo discute-se e avalia-se o andamento do projeto NeuroMat no contexto do Programa CEPID. Destaco ainda os workshops que tratam de pesquisa e de ferramentas de modelagem, que geralmente contam com a participação de pesquisadores estrangeiros considerados referências na área.

Entre os eventos realizados no período em que desenvolvi a etnografia destaco os workshops de Jovens Pesquisadores do NeuroMat, voltados para alunos de pós-graduação, pesquisadores de pós-doutorado e iniciação científica. O objetivo desses encontros era viabilizar espaços de formação desses pesquisadores dentro da missão e objetivos científicos do projeto do Centro, o que incluiu a apresentação das pesquisas em andamento e interação com pesquisadores seniores⁵³. A participação como observador no I Workshop de Jovens Pesquisadores foi de especial importância para o trabalho de campo uma vez que explorou a temática da Ciência Aberta nas discussões e possibilitou uma série de aproximações e interações com diversos atores que foram fundamentais para a orientação dos passos seguintes do trabalho de campo.

O quadro 2, a seguir, apresenta a relação de eventos acompanhados no período de maio de 2015 a outubro de 2017. Os eventos foram documentados por meio de diário de campo, gravação de áudio e fotografia.

53 A categoria “**jovens pesquisadores**” é utilizada pelos membros do NeuroMat para se referir a pesquisadores em estágio de formação, isto é, mestrandos, doutorandos e alunos de iniciação científica, ou recém-doutores vinculados ao centro, como os bolsistas de pós-doutorado. Já os “**pesquisadores seniores**” correspondem aos pesquisadores que possuem maior tempo de carreira acadêmica e ocupam níveis superiores na organização hierárquica do projeto.

Quadro 2: Eventos

Eventos Realizados pelo NeuroMat	Data de realização
1st NeuroMat Young Researchers Workshop	5 – 7 de maio, 2015
Workshop “Random Graphs in the Brain”	23 a 27 de novembro, 2015
Seminário de Pesquisa NeuroMat	7 de outubro, 2015
Seminário de Pesquisa NeuroMat	14 de abril, 2016
Seminário de Pesquisa NeuroMat	24 de agosto, 2016
II Maratona de Edição Neurociência e Matemática	29 de agosto de 2016
Seminário Rede Amparo (streaming)	24 de outubro, 2016
Seminário Rede Amparo (streaming)	30 de janeiro, 2017
Workshop “High-Performance Computing, Stochastic Modeling and Databases in Neuroscience”	24 – 29 de abril, 2017
Encontro “Cultura Matemática no Brasil: diagnósticos e perspectivas”	16 de maio, 2017
Workshop “Random Structures in the Brain”	16 – 20 de outubro, 2017
Eventos com participação de pesquisadores do NeuroMat	Data de realização
II Encontro anual do grupo de trabalho em Ciência Aberta – Apresentação de pesquisador do NeuroMat sobre as ferramentas de Ciência Aberta do centro.	28 de novembro, 2015
<u>BIN@BRAZIL</u> – Apresentação de pesquisadores do NeuroMat sobre uso das ciências matemáticas como recurso para aplicação em neurociências.	7 de novembro, 2016

Fonte: Elaborado pelo autor.

Além da participação nos workshops e seminários, foram feitas visitas regulares à sede do NeuroMat nesse período do trabalho de campo para entrevistas e algumas reuniões. Destaco a reunião de preparação do I Workshop de Jovens Pesquisadores mencionado acima, realizada em abril de 2015. Nesta reunião pude acompanhar a preparação da programação do evento e um debate sobre a atividade envolvendo Ciência Aberta que faria parte das atividades do workshop. Também participei de reuniões regulares com a equipe de difusão e também com a equipe técnica responsável pelas ferramentas computacionais relacionadas às estratégias de Ciência Aberta.

Por fim, destaco que durante o trabalho de campo também foi feita uma visita ao Instituto Deolindo Couto da Universidade Federal do Rio de Janeiro (INDC/UFRJ), onde está localizado o Núcleo de Pesquisas em Neurociência e Reabilitação (NPNR), um dos grupos de pesquisa colaboradores do NeuroMat. Ao longo de um dia no INDC, conheci as instalações do Instituto, incluindo ambulatório, laboratórios de neurologia, reabilitação e laboratórios de eletroencefalograma. Durante a visita, pude interagir com alguns pesquisadores que fazem parte do núcleo e pude visualizar especificidades da experimentação em neurobiologia por meio do acesso a equipamentos e ao ambiente de trabalho dos pesquisadores e ao ambulatório.

2.4.1 Entrevistas

No período de fevereiro de 2015 a outubro de 2017, foram realizadas 23 entrevistas com pessoas ligadas ao NeuroMat, entre pesquisadores docentes, pesquisadores colaboradores, pós-graduandos e técnicos, aqui referenciados como **Equipe NeuroMat**. Com auxílio de um roteiro de questões e tópicos (APÊNDICE – A), as entrevistas buscaram cobrir informações relativas à área do pesquisador, especialidade, a forma como o tema da Ciência Aberta é abordado dentro de sua área, e como o entrevistado se relaciona com o tema na sua prática cotidiana.

As entrevistas foram registradas por meio de um gravador e/ou por meio de notas escritas⁵⁴. A interação com algumas pessoas ocorreu em mais de um momento e em diferentes contextos, inclusive por meio de e-mails e chats em rede social.

No quadro 3, a seguir, estão as entrevistas realizadas com os pesquisadores docentes, alunos de mestrado e doutorado e pesquisadores de pós-doutorado⁵⁵. O quadro destaca a área de atuação dos pesquisadores e indica a forma de registro e a data de realização das entrevistas, incluindo as entrevistas que ocorreram mais de uma vez. Os nomes dos pesquisadores foram suprimidos e serão representados ao longo da pesquisa com um código (A+número de identificação).

54 O roteiro de questões funcionou apenas como uma bússola para as entrevistas. Procurei não ficar preso a um modelo de entrevistas dirigido para não engessar ou mesmo influenciar os informantes em suas respostas. Dentro dos limites impostos pelas preocupações da pesquisa, busquei dar certa liberdade para os entrevistados fazerem suas formulações.

55 Utilizo a categoria “**modeladores**” para me referir ao grupo de pesquisadores que trabalha diretamente com a aplicação de ferramentas de modelagem. Neste grupo estão incluídos os matemáticos, estatísticos, probabilistas, físicos e cientistas da computação. Em alguns momentos será importante demarcar a área de especialização do pesquisador, por essa razão, opto por chamá-lo, por exemplo, de “matemático”, “físico”, “cientista da computação”, etc.

Quadro 3: Pesquisadores Entrevistados.

Pesquisador	Perfil	Forma de registro e Datas
A1	<ul style="list-style-type: none"> Modelador Área: Modelagem estocástica do cérebro 	Gravada/notas 16/04/2015 e 18/10/2017
A2	<ul style="list-style-type: none"> Neurocientista Área: Neurofisiologia; Plasticidade sensoriomotora em humanos. 	Gravada 14/10/2015
A3	<ul style="list-style-type: none"> Modelador Área: Modelagem computacional; Neurociência computacional. 	Gravada/Notas 05/05/2015 e 12/10/2016 (Skype)
A4	<ul style="list-style-type: none"> Neurocientista Área: Neurociências da Visão. 	Notas 18/10/2017
A5	<ul style="list-style-type: none"> Neurocientista Área: Neurofisiologia da memória. 	Notas 06/05/2015
A6	<ul style="list-style-type: none"> Modelador Área: Sistemas de software 	Gravada 18/09/2017
A7	<ul style="list-style-type: none"> Modelador Área: Estatística, Análise de dados funcionais 	Notas 06/05/2015
A8	<ul style="list-style-type: none"> Neurocientista Área: Plasticidade cortical em pacientes com lesão do plexo braquial. 	Notas 18/05/2017
A9	<ul style="list-style-type: none"> Modelador Área: modelagem e gerenciamento de banco de dados. 	Gravada/Notas 05/05/2015 (Skype)
A10	<ul style="list-style-type: none"> Neurocientista Área: Neurofisiologia; fisioterapia; processos Perceptuais e Motores. 	Gravada 14/10/2015
A11	<ul style="list-style-type: none"> Modelador Área: Estatística, Probabilidade e estatística 	Notas 08/10/2015
A12	<ul style="list-style-type: none"> Modelador Área: Classificação de disparos neurais 	Notas 18/10/2017
A13	<ul style="list-style-type: none"> Modelador Área: Metodologia e técnicas da computação com especialidade em banco de dados. 	Gravada/ Notas 28/02/2015 06/05/2015 e 28/11/2015
A14	<ul style="list-style-type: none"> Modelador Área: Probabilidade e Estatística 	Notas 08/10/2015 e 29/08/2016
A15	<ul style="list-style-type: none"> Modelador Área: Probabilidade e Estatística Aplicadas. 	Notas 18/10/2017
A16	<ul style="list-style-type: none"> Modelador Neurociência Computacional; Processamento de sinais biológicos. 	Notas 29/08/2016
A17	<ul style="list-style-type: none"> Modelador Área: Processos estocásticos; grafos aleatórios. 	Notas 18/10/2017

Fonte: Elaborado pelo autor com base no Currículo Lattes.

A “**equipe técnica**” do NeuroMat ocupa um papel fundamental na produção e implementação da Ciência Aberta, o que justifica a inclusão no quadro de atores entrevistados. Foram incluídos nessa categoria os profissionais que atuam nas áreas de desenvolvimento de sistemas, profissionais da área de jornalismo, ciências sociais e linguística. No quadro 4, estão os entrevistados que fazem parte da equipe técnica. Esses entrevistados também tiveram os nomes suprimidos e serão identificados pelo código B+Número de identificação. Ao longo do texto esses informantes serão referenciados como “técnicos”.

Quadro 4: Equipe Técnica do NeuroMat.

Equipe técnica	Perfil	Registro da Entrevista e Data
B1	Ciências Sociais; Comunicação;	Notas 28/02/2015 29/08/2016 13/07/2017
B2	Analista de sistemas; Desenvolvedor <i>Web</i> .	Notas 28/02/2015
B3	Analista de sistemas; Gerente de transferência de tecnologia	Notas 28/02/2015
B4	Linguista; Assistente de disseminação científica.	Notas 16/04/2015 06/05/2015 07/05/2016
B5	Jornalista; bolsista de divulgação científica.	Gravação /Notas 19/09/2016 29/04/2017
B6	Jornalista; bolsista de divulgação científica.	29/08/2016

Fonte: Elaborado pelo autor.

Ademais das informações obtidas por meio das entrevistas, compõem o trabalho de campo o registro de declarações e comentários registrados em apresentações, discussões em reuniões e eventos, e mesmo conversas de corredor, em almoços e cafés.

De posse desses instrumentos e inserindo-me em diversos espaços por onde se conecta o NeuroMat, busquei descrever os significados, discursos e formas de representações com o intuito de dar forma ao objeto. Busquei registrar os pontos que me chamavam atenção, incluindo aqueles que não possuíam uma relação evidente como o objeto. Esse processo de registro foi essencial para entender a dinâmica interna, as formas de trabalho dos pesquisadores e seus instrumentos. No momento de produção do texto etnográfico, algumas percepções só vieram à tona no momento da sistematização e tentativa de construção desse

percurso por meio da confrontação com a literatura e um exercício de interpretação e escrita dos dados.

As análises que se seguem nos próximos capítulos são o produto da organização de anotações feitas nas atividades que acompanhei, transcrição das entrevistas, além de informações presentes em documentos e em outros materiais que coletei permanentemente a partir de buscas online. As reflexões desenvolvidas serão embasadas com o referencial teórico explicitado acima e com elementos empíricos do trabalho de campo.

Considerações Finais

Ao olhar para a Ciência Aberta como prática, parto de um questionamento no sentido de compreender como essa abordagem é significada e configurada nos espaços de produção da ciência. Nesse aspecto, o ferramental teórico e metodológico dos Estudos Sociais da Ciência e da Tecnologia possibilita explorar com cuidado as redes que atravessam o espaço da ciência. À medida que direcionamos o olhar para a prática e para a forma como os objetos são manipulados nessa prática, abandonamos a preocupação com significados fixos.

Baseando-me na perspectiva teórica e analítica descrita neste capítulo, o objetivo do trabalho de campo foi compreender a Ciência Aberta como algo interpretado e produzido nas e pelas redes em que os cientistas estão inseridos. Para tanto, parto de uma chave analítica atenta às mais diferentes formas que a Ciência Aberta pode assumir localmente; aos processos de negociação mais complexos e diversos; aos aspectos tecnológicos e as práticas culturais e ao papel dos atores envolvidos.

Por meio de uma perspectiva analítica da infraestrutura enquanto uma relação, procuro pensar as conexões de diferentes elementos associados que constituem a Ciência Aberta e como ela emerge como parte das práticas de um grupo de pesquisa, como redefine papéis, é incorporada e apropriada dentro e fora dessa comunidade, etc. Os subsídios que embasam respostas a alguns dilemas e questões acima colocados podem ser obtidos por meio de uma imersão no campo, proporcionada pela abordagem etnográfica. No capítulo seguinte abordo os resultados da pesquisa de campo desenvolvida.

3 A CIÊNCIA ABERTA NO CENTRO DE PESQUISA, INOVAÇÃO E DIFUSÃO EM NEUROMATEMÁTICA

Este capítulo trata da trajetória da pesquisa de campo no Centro de Pesquisa, Inovação e Difusão em Neuromatemática (NeuroMat) como forma de caracterizar o campo e identificar as atividades realizadas pelo centro. Busco aqui apresentar os elementos e características principais apresentadas pela abordagem de Ciência Aberta empregada no contexto das atividades do centro.

Ao longo do capítulo, descrevo a área de pesquisa da neuromatemática, destacando o projeto científico do centro, as principais abordagens e métodos empregados na área da modelagem neuronal. São apresentadas também a estrutura, a equipe e instituições participantes e eixos de atuação. Em seguida, o capítulo adentra os detalhes da pesquisa de campo, recortando as origens do envolvimento do NeuroMat com Ciência Aberta e as configurações que apresenta nesse espaço a partir do detalhamento das atividades desenvolvidas.

3.1 Adentrando o campo

A minha aproximação com o NeuroMat ocorreu em um momento do doutorado no qual eu buscava um local para desenvolver a pesquisa de campo. A primeira visita ao NeuroMat ocorreu no dia 16 de março de 2015, após algumas trocas de mensagens na página do Facebook do centro⁵⁶. Esse primeiro contato presencial foi especialmente marcante para os desenvolvimentos posteriores do trabalho de campo e para a reorientação da pesquisa, como explicitados na introdução desta tese.

56 Cf. F5

Naquele dia, meu anfitrião, um dos técnicos da equipe que naquele momento parecia cumprir o papel de assessor de imprensa, nada falou de Ciência Aberta sem antes me levar para fazer um *tour* pelo espaço para conhecer as instalações. Enquanto ele apresentava as salas e as instalações, explicava com detalhes a proposta científica do Centro. Entre detalhes sobre o que vinha sendo feito em cada um dos Eixos de atuação do NeuroMat, mencionava detalhes tais como o número de publicações feitas pelos pesquisadores até aquele momento, quem estava em congresso no exterior, os recursos financeiros disponíveis, os planos do centro para o futuro, entre outros detalhes próprios de ambientes acadêmicos como aquele.

Após conhecer o espaço, a conversa finalmente voltou-se para o tema que me levava até ali. Começou quando mencionei que entre as coisas que havia lido sobre o NeuroMat na viagem de ônibus de Campinas até São Paulo, havia sido um artigo que encontrei na Internet, de autoria de dois pesquisadores ligados ao centro, no qual os autores faziam uma reflexão sobre compartilhamento de dados na neurociência. Utilizei alguns pontos colocados pelo artigo para orientar a nossa conversa. Mencionei os pontos colocados pelos autores acerca do compartilhamento de dados na neurociência e a forma como os autores apontavam a existência de uma série de resistências entre os pesquisadores dessa área a esse tema. O artigo mencionava ainda o objetivo do NeuroMat de construir um banco de “dados neurofisiológicos” de acesso público.

A partir dos meus comentários, o meu anfitrião esclareceu que aquele artigo foi publicado com o propósito de marcar uma posição do centro sobre a Ciência Aberta na comunidade brasileira de neurociências. Segundo ele, aquele texto em especial poderia inclusive ser lido como uma espécie de “*manifesto do NeuroMat*”⁵⁷ pela Ciência Aberta.

Ao questioná-lo sobre a necessidade de haver um posicionamento do NeuroMat para entender qual a razão de se demarcar uma posição, ele disse que a preocupação com o compartilhamento aberto de dados é algo que esteve presente no projeto do NeuroMat desde a concepção do centro como parte dos objetivos de criar um centro de modelagem neuronal, do qual faria parte a criação de um banco de dados na área de neurociência para facilitar o trabalho conjunto de pesquisadores de diferentes áreas, como neurociência e matemática.

57 Ao longo da tese, trechos de fala e expressões atribuídas aos informantes serão diferenciadas com itálico e entre aspas. Trechos de fala longos serão recuados, obedecendo ao padrão de citações bibliográficas, indicando também quem é o informante. Os trechos que se referem a expressões minhas, geralmente indicando ênfase ou problematização, serão destacadas somente entre aspas.

Assim, publicar um manifesto direcionado à comunidade brasileira de neurociência pretendia chamar atenção para a necessidade de compartilhamento e para a construção de “*banco de dados de qualidade*” nessa área. A equipe formada por cientistas de diferentes áreas e profissionais com diferentes formações estava se envolvendo em várias atividades relacionadas à Ciência Aberta como parte do projeto científico do NeuroMat – especialmente ferramentas computacionais.

Logo em seguida nos dirigimos até uma novíssima sala de reuniões onde conversei longamente com alguns membros da equipe. Nesse momento foram feitas demonstrações de algumas ferramentas que vinham sendo desenvolvidas como parte de um “*esforço do NeuroMat pela Ciência Aberta*”.

Já naqueles momentos, as pessoas com quem interagi (técnicos desenvolvedores e pesquisadores) fizeram questão de deixar claro que as concepções de Ciência Aberta ali existentes iam além dos temas que frequentemente aparecem vinculados a essa agenda de discussões, tais como acesso aberto a publicações científicas, liberdade, democratização do conhecimento. Falava-se muito em softwares, código aberto, protocolos, metadados, licenças, frequentemente ligados a aspectos muito práticos das pesquisas ali desenvolvidas.

Além disso, outros elementos recorrentes na fala desses interlocutores sugeriam que havia “*entraves*” no centro em relação à adoção de uma abordagem de Ciência Aberta. Segundo eles, tais entraves podiam ser identificados na relação com a existência de “*diferentes culturas*” para se referir a uma resistência dos pesquisadores em compartilhar ou “*abrir o que quer que seja*”.

Os interlocutores demonstravam que ocorriam controvérsias internas ou, pelo menos, diferenças acerca da compreensão sobre o compartilhamento dos resultados utilizando uma abordagem de Ciência Aberta ou da participação dos pesquisadores nessa prática. A Ciência Aberta entre as normas de funcionamento do centro, isto é, como diretriz que obrigava os pesquisadores a abrirem os dados, fazia emergir perspectivas disciplinares conflitantes sobre o papel ou da necessidade de trazer essa questão “*tabu*” no contexto do projeto científico do centro.

Naquele momento, as diferenças internas se manifestavam de forma implícita na maneira como os interlocutores pareciam “entregar” os colegas, apontando nomes, separando-os por áreas, quem era declaradamente contrário ao compartilhamento aberto de dados e quem era a favor. Essas informações acabaram me conduzindo pelo caminho no sentido de tentar

entender os contornos dessas culturas e como elas se manifestavam em relação à Ciência Aberta. Foi a partir daí que comecei a interagir com várias pessoas ligadas ao NeuroMat com o intuito de compreender as dinâmicas particulares existentes e como o tema se inseria nas atividades ali desenvolvidas.

A primeira constatação que o campo ofereceu foi que, dificilmente, uma etnografia que propusesse compreender as práticas e as dinâmicas deste centro ficaria restrita àquele prédio de cor branca do Instituto de Matemática e Estatística. Desde o início, ficou evidente que o centro tinha uma dinâmica multidisciplinar e organizacional bastante particular, configurando-se como uma rede na qual se conectam diferentes culturas científicas, pessoas, equipamentos, computadores, softwares, dados.

De início não tive muitas dificuldades para me aproximar dos pesquisadores, especialmente daqueles que ficavam na sede, no campus Butantã. Desde o começo houve um relativo interesse da maioria dos membros da equipe em me receber e em conceder entrevistas. Com o passar do tempo e à medida que as visitas se tornaram frequentes, passei a me sentir mais à vontade para transitar pelos espaços. As pessoas pareciam cada vez mais acostumadas a me ver em reuniões, eventos e interagindo com a equipe. Inclusive ganhei um “passe livre” a partir da segunda visita para acessar as instalações, o que foi simbolizado com afixação da minha foto em um painel no qual eram identificados de todos os membros da equipe e colaboradores do NeuroMat, localizado na recepção do prédio.

A exceção foi com os neurocientistas ligados à prática clínica e neurocientistas do ICe/UFRN. O acesso a esses pesquisadores se mostrou bastante difícil, já que não retornavam os contatos via e-mail e dificilmente participavam das atividades presenciais em que estive. Destaco também a limitação de acessar a linguagem especializada dos vários campos presentes no NeuroMat, o que me deixou algumas vezes reticente sobre como abordá-los e como direcionar minhas questões. A “tradução” dessa linguagem de forma a deixá-la compreensível – de maneira a garantir a fidelidade às propriedades e especificidades do(s) domínios(s) ali presentes – foi um grande desafio enfrentando ao longo da pesquisa de campo e principalmente no momento da produção do texto etnográfico.

Um dos principais desafios no início foi me integrar em um ambiente essencialmente de ciências exatas, que embora tenha sido, de maneira geral, bastante receptivo, às vezes se mostrava muito intimidador para um etnógrafo iniciante como eu. Penso que essa dificuldade se deu tanto em razão de uma idealização que construí daquele

território como sendo um ambiente caracterizado pelo rigor e pela objetividade, o que poderia se traduzir em uma recusa ou estranhamento de um pesquisador “de humanas” entre eles, como também da dificuldade dos próprios indivíduos em compreender o que eu fazia ali.

Essa dificuldade, em especial, pôde ser percebida nos momentos em que fui definido por alguns membros da equipe como um “paparazzi” ou, simplesmente, como “jornalista” que estava ali para revelar os bastidores do centro de pesquisa. Ou mesmo quando alguns sujeitos se dirigiam a mim interpelando sobre a seriedade da minha pesquisa, como ocorreu em duas ocasiões; ou tentavam encerrar qualquer questão – cujas respostas me pareciam complexas – com respostas taxativas ou que fugiam do assunto.

Um desses momentos marcantes – que de certo modo, representou um choque com as práticas culturais desses pesquisadores – ocorreu quando, durante a primeira reunião que acompanhei como observador, após me apresentar como doutorando e explicar brevemente a pesquisa, mencionando que eu tinha interesse em investigar o que o NeuroMat compreendia como Ciência Aberta, um dos presentes faz o comentário, em tom jocoso, no qual colocava Ciência Aberta como sinônimo de ser aberto ao estranho, a quem não é da “área”. Ele disse: *“Aqui é muito aberto, a começar pela reunião... veja, você até está participando.”*

O comentário do pesquisador mencionado acima, bem como outros semelhantes surgidos nos momentos iniciais do trabalho de campo, chamou a minha atenção para o fato de que, muitas vezes, esses sujeitos assumiam que eu estava ali para (ou que minha pesquisa se resumia a) saber se eles eram “contra ou a favor” da Ciência Aberta e que, caso eu percebesse que eram contra, iria “denunciá-los” ou criticá-los por estarem fazendo algo “errado”.

A postura do pesquisador mencionado acima, assim como outras afirmações que surgiram ao longo das primeiras interações, sugeriam que havia um risco de reduzir a abordagem a essas categorias binárias (ser contra ou a favor). Colocava-se aí um desafio: como fazer um salto para pensar em questões mais complexas e amplas sobre Ciência Aberta na prática?

Em uma outra situação, aquela ideia de ser aberto ao diferente foi colocado novamente por outro pesquisador, porém de uma outra maneira. Enquanto retornava do almoço com outros membros da equipe, conversei com um pesquisador estrangeiro que acabara de ingressar no centro como bolsista de pós-doutorado. Quando comentei com ele sobre a proposta de investigar a Ciência Aberta no NeuroMat, ele comentou *“Aqui é muito*

aberto, eu sei bem disso...”. Naquele momento, ele fazia referência à sua múltipla condição de ser de fora: era estrangeiro, um físico em meio a muitos matemáticos e, até então, tinha pouca experiência no domínio da modelagem em neurociência. Ainda assim, os demais integrantes foram muito abertos em recebê-lo e integrá-lo ao grupo.

Foi a partir dessas interações que resolvi assumir também uma postura “mais aberta” em relação ao campo, inclusive deixando em suspenso certas concepções e conceitos preconcebidos sobre o objeto. O intuito era entender a circulação de significados, buscando nos discursos e nos vestígios materiais rastros e contornos da Ciência Aberta entre aqueles pesquisadores. Foi por isso também que o início da pesquisa de campo foi orientada por algumas questões empíricas iniciais: O que os pesquisadores daquele local estavam chamando de Ciência Aberta? E afinal, o que é a neuromatemática e o que isso tem a ver com Ciência Aberta?

A seguir, busco situar o leitor na proposta científica do NeuroMat antes de adentrar os detalhes da pesquisa de campo – embora o tópico seguinte seja também resultado dessa imersão no campo, que requereu um aprofundamento deste autor em temas, conceitos e instrumentos de pesquisa da área.

3.2 Neuromatemática: a nova ciência do cérebro

O NeuroMat é um centro de pesquisa avançada e multidisciplinar sediado no Instituto de Matemática e Estatística da Universidade de São Paulo (IME/USP). O centro de pesquisa faz parte do programa CEPIDs – Centros de Ensino, Pesquisa, Inovação e Difusão – da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). O CEPID NeuroMat é resultado do segundo edital do programa lançado em 2011 e divulgado em 2013, no qual foram aprovados ao final 17 centros no estado de São Paulo⁵⁸.

58 De acordo com Izique (2013), no edital de 2011 foram aprovados 17 centros no estado de São Paulo. Os centros estão distribuídos nos seguintes temas de pesquisa: alimentos e nutrição; vidros e cerâmica; materiais funcionais; neurociência e neurotecnologia; doenças inflamatórias; biodiversidade e descoberta de novas drogas; toxinas, resposta imune e sinalização celular; neuromatemática; ciências matemáticas aplicadas à indústria; obesidade e doenças associadas; terapia celular; estudos metropolitanos; genoma humano e células-tronco; engenharia computacional; processos oxidantes e antioxidantes em biomedicina; estudos da violência; óptica, biofotônica e física atômica e molecular.

O programa CEPID⁵⁹ consiste em uma linha de financiamento da FAPESP que apoia a formação de centros de pesquisa básica e aplicada no estado de São Paulo em temas científicos considerados estratégicos ou de grande importância para o desenvolvimento científico e tecnológico do país.

O programa faz parte de uma estratégia da FAPESP de financiamento a longo prazo e objetiva o desenvolvimento de novos paradigmas institucionais para a organização da pesquisa científica⁶⁰. Para tanto, volta-se para o fomento de equipes de pesquisa multidisciplinares e de alto nível dedicadas à produção de conhecimento científico com aplicações de impacto comercial e social, incluindo atividades de extensão e popularização da ciência.

O projeto do NeuroMat aprovado pela FAPESP afirma que o objetivo NeuroMat é integrar a modelagem matemática e computacional com pesquisa básica e aplicada na fronteira da neurociência por meio de uma abordagem fundamentalmente interdisciplinar que associa pesquisadores de diferentes especialidades das ciências duras, ciências da computação e neurociência (NEUROMAT, 2013)⁶¹.

A proposta do CEPID NeuroMat é desenvolver ferramentas matemáticas que possam integrar observações e explicações dos fenômenos cerebrais com a finalidade de preencher uma lacuna teórica da neurociência, geralmente apontada como sendo rica em produção de dados, porém os paradigmas de análise dos dados experimentais não estão consolidados.

A neurociência é uma área formada por diferentes disciplinas, dedicadas a estudar o sistema nervoso e suas ligações com toda a fisiologia do organismo, incluindo a relação entre cérebro e comportamento (VENTURA, 2010; BEAR et al, 2017). A pesquisa experimental nas diferentes subdisciplinas da neurociência emprega uma série de técnicas para investigar processos neurofisiológicos, geralmente associados à compreensão das patologias do sistema nervoso, utilizando uma variedade de métodos experimentais como eletroencefalografia (EEG), ressonância magnética e experimentação animal.

59 CEPID. Home page. Disponível em: <<http://cepid.fapesp.br/home/>>. Acesso em: 20 mar. 2018.

60 O período de financiamento dos CEPIDS é de cerca de 11 anos. O financiamento total destinado aos 17 Centros aprovados no edital de 2011 está estimado em cerca de R\$ 1,4 bilhão, com R\$ 760 milhões da FAPESP e R\$ 640 milhões em salários pagos pelas instituições sedes aos pesquisadores e técnicos. Recursos adicionais precisam ser obtidos pelos centros junto à iniciativa privada e outras agências de financiamento à pesquisa (FAPESP, 2018).

61 Cf. F1

De acordo com Ribeiro (2014), um dos desafios da neurociência é sair do nível dos fatos puramente descritivos ou fenomenológicos do sistema nervoso para um nível descritivo e preditivo das interações neuronais. Nesse aspecto, a hipótese que orienta o projeto científico do NeuroMat é de que o cérebro trabalha com certos padrões e regularidades (ainda invisíveis) que podem ser apreendidos e modelados utilizando instrumentos matemáticos, estatísticos e computacionais aplicados aos dados neurobiológicos coletados na pesquisa experimental desenvolvida em diferentes especialidades da neurociência. Coloca-se então a necessidade de instrumentos matemáticos para tratar os dados que deem conta da grande quantidade e complexidade dos dados neuronais coletados nesses experimentos.

A Figura 2, a seguir, ilustra uma das questões que norteia as investigações do NeuroMat: “O cérebro é um estatístico?”. O que se busca compreender por meio dessa questão é se os fenômenos neurofisiológicos, como conexões neuronais, obedecem a determinados padrões, e se existe regularidade nesses padrões de maneira que possam ser representados matematicamente. Por meio dos instrumentos matemáticos aplicados aos dados experimentais, os pesquisadores abrigados nesse centro buscam, assim, desenvolver modelos neuronais precisos capazes de representar essas regularidades.

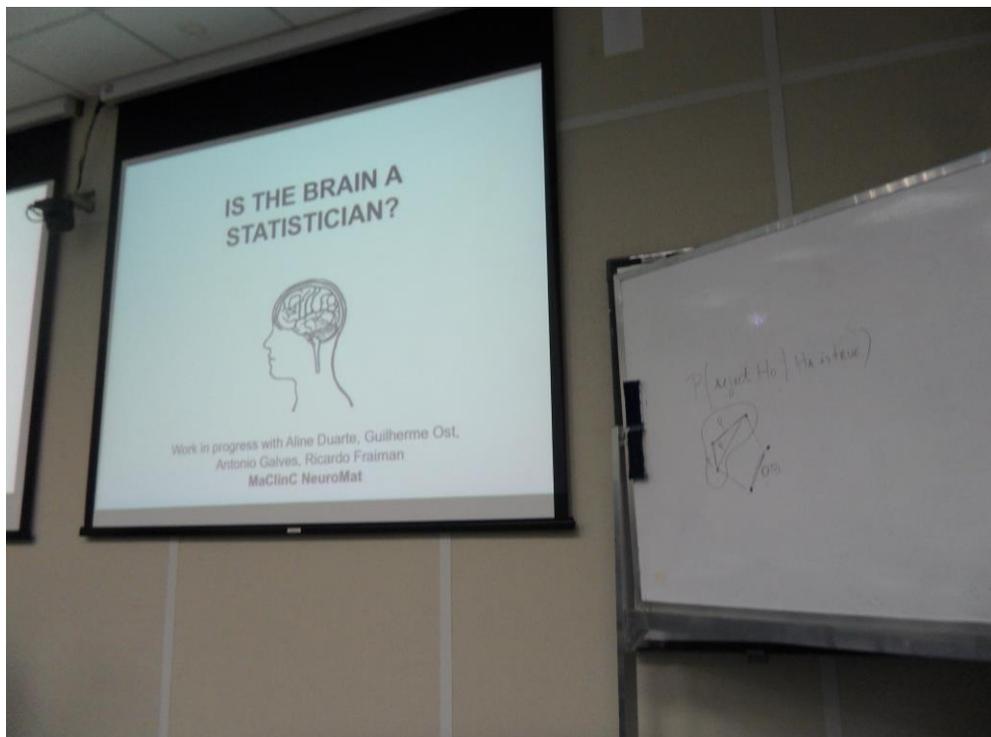


Figura 2: Apresentação da linha de pesquisa em neuromatemática durante workshop.
Fonte: Arquivo do autor.

Em matéria no portal da USP⁶² que trata do lançamento do projeto, a neuromatemática é definida como uma “nova ciência do cérebro” que pretende construir um quadro conceitual adequado à formulação dos problemas da neurociência por meio da aplicação de ferramentas de modelagem matemática aos dados experimentais produzidos no domínio da neurociência⁶³.

A abordagem matemática para os fenômenos cerebrais depende de uma nova classe de modelagem matemática que “junte” teoria da probabilidade, estatística, ciência da computação e neurociência para descrever e explicar com precisão os fatos experimentais, as múltiplas escalas e relações da atividade neural (redes neurais) buscando alcançar um nível explicativo para os fenômenos do sistema nervoso.

A construção dos modelos ocorre em conjunto com o desenvolvimento de novos métodos computacionais e aplicação de infraestruturas para armazenamento e análise de uma grande quantidade de dados experimentais, produzidos em experimentos da neurociência. A computação de alto desempenho é empregada em simulações de grande escala para testar as hipóteses sobre as regularidades de conectividade entre neurônios.

Na figura 3, técnicos trabalham na instalação dos equipamentos do *cluster* computacional de alta performance no NeuroMat, localizado em Ribeirão Preto, no qual são realizadas simulações do funcionamento do córtex cerebral⁶⁴.

62 Cf. F2

63 Existe uma disputa interna entre os pesquisadores do NeuroMat sobre o uso do termo “neuromatemática” como uma nova disciplina. Alguns apontam que o que o NeuroMat se insere no campo da “neurociência computacional” – ramo multidisciplinar da neurociência que combina computação, física, matemática e neurobiologia na modelagem dos fenômenos do sistema nervoso (SISNE, 2008). Não é objetivo da tese analisar o conteúdo da ciência produzida no NeuroMat tão pouco adentrar aspectos metodológicos e epistemológicos dessa área.

64 O córtex cerebral corresponde à camada cinzenta logo abaixo à superfície do cérebro, sendo uma região rica em neurônios (BEAR et al, 2017)



Figura 3: Técnicos instalam o supercomputador do NeuroMat.

Fonte: Página do NeuroMat no Facebook

A proposta científica do NeuroMat originou-se na linha de investigação desenvolvida pelo pesquisador e coordenador do projeto do centro de pesquisa, o matemático Antônio Galves, no campo da seleção estocástica de modelos, o que resultou na criação de um modelo matemático que permite simular neurônios e redes neurais, desenvolvido em colaboração com a pesquisadora do Université de Cergy-Pontoise, Eva Löcherbach. O modelo “*Galves-Loecherbach*” está na base de muitos projetos de pesquisa desenvolvidos no centro, o que é evidenciado pela menção recorrente nos artigos, apresentações e projetos de pesquisa⁶⁵.

Um dos fenômenos que o modelo *Galves-Loecherbach* permite descrever são os “disparos neuronais” ou “*spikes*”, que se referem ao potencial de ação para a transmissão de informações entre células do sistema nervoso (POUZAT, 2014). Os disparos neuronais podem ser visualizados por meio de registros de sinais elétricos captados nas imediações de neurônios individuais através de microelétrodos e são registrados como impulsos eletrofisiológicos.

65 O modelo “Galves-Locherbach” é um modelo com estocasticidade intrínseca para redes neurais, em que a probabilidade dos disparos dos neurônios depende da evolução total do sistema desde o último disparo (GALVES; LOECHERBACH, 2013).

Algumas das técnicas de modelagem desenvolvidas pelo NeuroMat compreendem a análise de dados eletrofisiológicos chamadas de “classificação dos disparos neuronais” ou *spike sorting* (QUIROGA, 2007). De acordo com informações veiculadas na revista *Mente e Cérebro*⁶⁶ sobre Workshop realizado pelo NeuroMat sobre a técnica de *spike sorting*, a classificação de disparos neuronais refere-se a um conjunto de técnicas que combina ferramentas matemáticas, estatística e processamento de sinais com a finalidade de identificar eventos – dentro de um conjunto de registros – em um espaço mais delimitado do que o conjunto de neurônios registrados por técnicas convencionais de registro de atividade neuronal. A técnica envolve a detecção de eventos, a “limpeza” de dados e a progressiva redução de amostras nesse conjunto de disparos de tal maneira que se aproxime de apenas um disparo neuronal. O processo de classificação de disparos envolve – a partir da discriminação do registro da atividade neuronal obtido – a identificação dos vários agrupamentos de neurônios. Os dados que passaram pelo processo de *spike sorting* permitem ao pesquisador seguir a análise de atividade neuronal e identificar formas de sincronismo entre neurônios e com isso obter modelos capazes de prever a probabilidade de um disparo do neurônio em dado momento⁶⁷.

Outra linha de atuação do NeuroMat insere-se na modelagem das redes neurais, que é feita por meio de uma abordagem probabilística aplicada à modelagem de dados das interações entre neurônios. Essa abordagem é chamada de “hipótese para distribuições aleatórias de grafos” ou “teoria dos grafos aleatórios” (*random graphs*) aplicada à compreensão das redes neurais ou da “conectividade do cérebro”⁶⁸.

A modelagem a partir do mapeamento de conexões neuronais é utilizada desde aplicações avançadas, como a interligação de modelos computacionais com artefatos mecânicos, possibilitando a reprodução de comportamentos por meio de pernas ou braços robóticos, a também ferramentas preditivas que permitem a compreensão de distúrbios cerebrais como epilepsia, enxaqueca, Alzheimer, Acidente Vascular Cerebral (AVC) e doença de Parkinson. Nas investigações que buscam a compreensão desses distúrbios, as ferramentas de modelagem matemática são utilizadas para “desemaranhar” as dinâmicas e variabilidade

66 Cf. F3

67 Gonzalo Ray et al (2015) explicam que a técnica de *spike sorting* é parte fundamental da eletrofisiologia moderna, sendo utilizada, por exemplo, na compreensão do comportamento e transmissão elétrica das células.

68 As aplicações práticas da teoria dos grafos ocorrem em todas as áreas de modelagem de redes complexas (HOFSTAD, 2016). Segundo Bullmore e Sporns (2009), os desenvolvimentos recentes na análise quantitativa de redes complexas são baseados em grande parte na teoria dos grafos.

das conexões neuronais possibilitando, eventualmente, formas de previsão da eficácia dos tratamentos para as doenças que afetam o sistema nervoso (LYTTON, 2008; RIBEIRO, 2014).

Um fenômeno que é possível compreender por meio da aplicação dessas ferramentas de modelagem é a chamada “plasticidade do cérebro” ou “neuroplasticidade”, isto é, o fenômeno que compreende a capacidade do sistema nervoso adaptar-se estrutural e funcionalmente quando submetido a um novo contexto (ZOLNERKEVIC, 2014).

O conceito básico por trás da plasticidade é de que os neurônios restauram suas funcionalidades, por meio de atalhos, quando ocorre algum rompimento das transmissões elétricas e neuronais⁶⁹. Diante de uma lesão e trauma, por exemplo, as células se reorganizam, completamente ou parcialmente, permitindo recuperar determinadas funções do organismo associadas a elas. As pesquisas realizadas no ramo da modelagem que buscam representar a reconstrução das redes neurais na plasticidade do cérebro podem levar ao desenvolvimento de modelos preditivos capazes de auxiliar na reabilitação de pacientes que sofrem de determinadas lesões do sistema nervoso (RIBEIRO, 2014).

O NeuroMat desenvolve estudos sobre plasticidade do cérebro em colaboração com o grupo de pesquisadores do Núcleo de Pesquisas em Neurociência e Reabilitação (NPNR) da Universidade Federal do Rio de Janeiro, localizado no Instituto de Neurologia Deolindo Couto (INDC/UFRJ). O NPNR desenvolve pesquisas na área de neurociência utilizando diferentes técnicas experimentais, principalmente na investigação da plasticidade do cérebro em contextos de lesão do sistema nervoso central e periférico em casos de lesão dos nervos associados ao controle motor humano. O Núcleo é composto por médicos, fisioterapeutas, psicólogos, estatísticos, probabilistas, cientistas da computação. Uma das principais especialidades do grupo é o estudo a plasticidade do cérebro em casos de lesões do plexo braquial – conjunto de nervos que liga a medula espinhal ao braços e que está associado à transmissão de comandos motores e sistema sensorial. Utilizando os dados experimentais coletados por esses pesquisadores em pacientes que sofreram rompimento dos nervos do plexo Braquial, os modeladores do NeuroMat produzem modelos que ajudam a compreender o comportamento das células cerebrais em casos de lesão desses nervos, auxiliando no desenvolvimento de tratamentos de reabilitação mais eficazes para pacientes que sofrem da lesão, assim como no desenvolvimento de modelos sobre plasticidade do cérebro.

69 Cf. F4

3.3 Estrutura do NeuroMat

O NeuroMat é um centro de pesquisa multidisciplinar, como perfil variado de pesquisadores e outros profissionais que compõe a equipe. Inclui matemáticos, probabilistas e estatísticos, cientistas da computação, físicos e diferentes especialidades da neurociência como biólogos, médicos, fisioterapeutas e psicólogos.

O centro é resultado de uma rede preexistente de pesquisadores interessados na modelagem neuronal, especialmente da modelagem da plasticidade do cérebro, organizada pelo coordenador do centro de pesquisa a partir de outro projeto financiado pela USP que deu origem ao Núcleo de Apoio à Pesquisa em Modelagem Estocástica e Complexidade (NUMEC)⁷⁰. De acordo com números divulgados do relatório de atividades da FAPESP 2016, o centro possui o total de oito pesquisadores principais; 20 pesquisadores associados; 14 colaboradores no exterior; cinco conselheiros internacionais; 22 instituições estrangeiras parceiras de sete países e oito instituições associadas brasileiras (FAPESP, 2017).

Seguindo a estrutura de organização dos CEPIDs, o centro está dividido entre pesquisadores coordenadores e pesquisadores associados do Brasil e do exterior. Existe ainda uma estrutura de profissionais que dá suporte às diferentes atividades realizadas, como uma equipe administrativa, formada por dois assistentes administrativos, um analista de sistemas e um gerente executivo; e uma equipe de suporte técnico formada por analistas de sistemas e desenvolvedores, jornalistas responsáveis pela assessoria de imprensa, linguista e bolsistas integrados ao projeto de difusão científica.

As instalações físicas ocupam um prédio inteiro de dois andares no IME/USP. O espaço é formado por uma sala de conferências; sala com estações de trabalho exclusivamente para a equipe do Eixo de Difusão Científica. Os alunos de mestrado e doutorado e os pesquisadores de pós-doutorado ocupam uma sala própria no segundo andar com estações de trabalho individuais. Há ainda duas salas em que estão alocadas a equipe de desenvolvedores e o suporte administrativo; uma área de convivência no primeiro andar; uma sala para realização de reuniões e outras salas para equipamentos de eletroencefalografia (EEG)⁷¹.

70 O NUMEC funciona nas mesmas instalações do NeuroMat.

71 O centro passou a ter a configuração descrita a partir de 2017, quando as instalações físicas foram ampliadas.

Os pesquisadores colaboradores são vinculados a diferentes instituições de pesquisa no Brasil e no exterior. A maior parte deles está vinculada a diferentes unidades acadêmicas da USP, especificamente a Faculdade de Medicina (FM/USP), Instituto de Biociências (IB/USP), Instituto de Física (IF/USP), Instituto de Matemática e Estatística (IME/USP), incluindo o Centro de Competências em Software Livre (CCSL). No campus da USP de Ribeirão Preto fica localizado o centro computacional de alta performance para a simulação do córtex cerebral e onde também o NeuroMat realizada anualmente a Escola Latino-Americana de Neurociência Computacional (LASCOM).

Também fazem parte da estrutura do centro o Instituto do Cérebro da Universidade Federal do Rio Grande do Norte (Ice/UFRN) e o Laboratório de Neurobiologia II do Instituto de Biofísica Carlos Chagas Filho da Universidade Federal do Rio de Janeiro (UFRJ). Há ainda colaboradores na Universidade Estadual de Campinas (UNICAMP), Universidade Federal de Ouro Preto (UFOP), Universidade Federal de Pernambuco (UFPE), Instituto de Matemática Pura e Aplicada (INPA) e Université de Cergy Pontoise. O centro integra também a *International Neuroinformatics Coordinating Facility* (INCF)⁷² por meio de um grupo de interesse especial sobre modelagem estocástica de sistemas neuronais (NEUROMAT, 2017).

3.4 Eixos de atuação do NeuroMat

Assim como os demais CEPIDS, o NeuroMat é organizado em três eixos de atuação: 1) Eixo de Pesquisa, descrito acima 2) Eixo de Transferência de Conhecimento e Tecnologia 3) Eixo de Educação e Difusão Científica. Os eixos de Transferência e Difusão desenvolvem atividades pautadas pelo arcabouço conceitual e experimental das pesquisas no campo da neurociência e matemática e cada eixo é coordenado por um pesquisador principal.

⁷² A INCF é uma rede acadêmica internacional de desenvolvimento de infraestrutura colaborativa e compartilhamento de recursos de dados e computação para a comunidade internacional de pesquisa em neuroinformática. Disponível em: <<https://www.incf.org/>>. Acesso em: 20 mar. 2018.

3.4.1 Eixo de Transferência de Conhecimento e Tecnologia

Este eixo compreende atividades de transferência de tecnologia e conhecimento nas áreas de reabilitação neurológica e diagnóstico relacionado à lesões, psicoses e doenças psiquiátricas e plasticidade do cérebro por meio de um programa voltado para hospitais e outras instituições de saúde.

Como parte das atividades de transferência de conhecimento foi criada uma rede de conhecimento sobre doença de Parkinson – Rede de Apoio NeuroMat a Amigos e Pessoas com Doença de Parkinson (Rede Amparo)⁷³. A Rede atua na organização de atividades voltadas para profissionais da saúde, familiares e portadores de doença de Parkinson e pesquisadores. A rede promove palestras, que são transmitidas por *streaming*, e oficinas sobre os desafios relacionados aos cuidados e melhoria da qualidade de vida de pacientes com Parkinson e desenvolvimento de políticas públicas.

Em 2017, também como parte das atividades de transferência de conhecimento e tecnologia, foi lançada a Iniciativa NeuroMat para a Lesão do Plexo Braquial (ABRAÇO). A iniciativa agrega um conjunto de ações em torno da pesquisa e transferência de conhecimento sobre lesão dos nervos que afeta a sensibilidade e a mobilidade dos do membro superior. A iniciativa foi criada pelo NeuroMat para acolher e orientar pacientes e familiares e oferecer treinamento e capacitação para profissionais de saúde, disponibilizar informações de pesquisa e recursos educacionais sobre a lesão, prognóstico e suas consequências funcionais por meio de uma plataforma online aberta, que pretende funcionar como um espaço de discussão de resultados de pesquisa sobre Plexo Braquial⁷⁴. A proposta também é que sejam disponibilizadas ferramentas computacionais orientadas ao aperfeiçoamento e ao avanço da pesquisa sobre o tema, dentro de uma perspectiva de modelagem da plasticidade do cérebro a partir de dados experimentais sobre essa lesão.

O desenvolvimento e disponibilização de ferramentas computacionais também são parte das estratégias de transferência de tecnologia adotada, integrando a missão mais ampla de oferecer suporte à pesquisa científica sobre o funcionamento do cérebro por meio da

⁷³ NeuroMat AMPARO. Home page. Disponível em: <<https://amparo.numec.prp.usp.br/>>. Acesso em: 20 mar. 2018.

⁷⁴ ABRAÇO. Home page. Disponível em: <<http://abraco.numec.prp.usp.br/>>. Acesso em: 20 mar. 2018.

modelagem dos dados neuronais⁷⁵. Dentre os principais produtos tecnológicos desenvolvidos estão as ferramentas de software e banco de dados experimentais de acesso aberto.

A plataforma do projeto ABRAÇO assim como as ferramentas computacionais desenvolvidas pelo centro estão inseridas também em um esforço mais amplo do NeuroMat em utilizar e disseminar os princípios de Ciência Aberta.

3.4.2 Eixo de Educação e Difusão Científica

O Eixo da Difusão Científica ocupa um espaço considerável na agenda do NeuroMat. É formado por uma equipe que inclui um coordenador principal, um supervisor de comunicação, um assistente de divulgação e bolsistas. A equipe de difusão produz materiais audiovisuais, desenvolve treinamentos voltados para professores e estudantes; organiza exposições em museus de ciência, conferências e cursos em escolas e participa de eventos de difusão e educação científica, além de manter um blog em que são registradas as atividades do grupo. A equipe também é responsável por fazer a manutenção do web site do centro, redes sociais e produzir uma *newsletter* mensal com o resumo das principais atividades.

Dentre essas iniciativas, a principal atividade do eixo é o projeto de “Wikidifusão” envolvendo as aplicações *wiki* da Fundação Wikimedia⁷⁶. O projeto inclui a produção de conteúdo para verbetes relacionados a temas da matemática e neurociência na Wikipédia⁷⁷; disponibilização de imagens e vídeos que explicam conceitos matemáticos na plataforma de conteúdo multimídia *Wikimedia Commons*⁷⁸ – que também são utilizados nos verbetes – e treinamentos para pesquisadores e para o público em geral sobre a utilização dessas aplicações.

75 As duas principais ferramentas desenvolvidas pelo NeuroMat são o software Neuroscience Experiments System para gerenciar e disponibilizar dados experimentais, que será descrito no capítulo 5, e Goalkeeper Game, um jogo de computador que coleta dados para estudo do aprendizado de processos determinísticos e probabilísticos. Disponível em: <<http://game.numec.prp.usp.br/>>. Acesso em: 20 mar. 2018.

76 A Fundação Wikimedia é uma organização sem fins lucrativos responsável por manter a enciclopédia eletrônica Wikipédia e outras aplicações *wiki* citadas no texto. A Fundação é uma das principais promotoras do “conhecimento aberto”, apoiando diferentes iniciativas do gênero.

77 A Wikipédia é uma enciclopédia aberta e editável construída em um sistema *wiki*. Atualmente é um dos principais sites acessados na Internet.

78 Cf. <<https://commons.wikimedia.org>>. Acesso em: 20 mar. 2018.

3.5 Configurando a Ciência Aberta no NeuroMat

A proposta de adoção de uma abordagem de Ciência Aberta como parte do projeto científico do NeuroMat, sobretudo obrigando os pesquisadores a compartilharem os dados em um repositório de acesso aberto, gerou uma série de controvérsias entre os pesquisadores associados ao centro. Alguns membros da equipe relataram diferentes conflitos internos relacionados às diferenças na forma de neurocientistas e modeladores abordarem o compartilhamento dos dados.

Um fato que ilustra esse conflito ocorreu durante a primeira reunião da equipe do projeto, em janeiro de 2014, logo em seguida à aprovação do projeto pela FAPESP. Nessa reunião, em que muitos integrantes da equipe se encontravam pela primeira vez, a coordenação do centro pretendia criar uma integração entre os membros e permitir que tivessem contato com a missão e visão multidisciplinar proposta, se inteirando das áreas de cada um.

A proposta do workshop mencionado era que os pesquisadores apresentassem suas linhas de pesquisa e a equipe do NeuroMat apresentasse os objetivos e a missão do centro. Durante a reunião, quando foi colocada aos integrantes do projeto a proposta de que o centro utilizaria princípios de Ciência Aberta para o compartilhamento de resultados experimentais por meio de um banco de dados, algumas reações contrárias vieram à tona imediatamente. Alguns informantes relataram que uma parte dos neurocientistas foi totalmente contrária naquele primeiro momento a abrir os dados. Os neurocientistas questionaram fortemente a pertinência de trazer aquela questão como pauta para o centro.

Na primeira reunião que realizei com a equipe envolvida com a iniciativa de Ciência Aberta, um dos presentes mencionou que houve um “*rebuliço*” quando surgiu o assunto durante o workshop. Para ele, um fato bastante inusitado durante essa atividade simbolizou como o compartilhamento aberto de dados seria um assunto controverso para o grupo: a convulsão de um rato.

Conforme ele descreveu, parte da apresentação de um dos pesquisadores colaboradores exibiu o vídeo de um experimento que consistia em induzir convulsões em ratos. Na sequência das apresentações no workshop, foi a vez de apresentar a proposta para cada um dos eixos de atuação do centro, incluindo a proposta de que o centro teria uma “*estratégia de Ciência Aberta*” em seu Eixo de Transferência de Conhecimento e Tecnologia, porém isso despertou naquele momento uma grande reação entre alguns dos principais colaboradores presentes na sala:

Em um primeiro momento a proposta não foi bem recebida... No primeiro evento foi bastante marcante porque nenhuma das apresentações, nem um vídeo de um ‘rato convulsionando’ (sic) gerou tamanha polêmica entre alguns pesquisadores como quando foi apresentada a proposta de que seríamos *Open Science* [Entrevista com B1].

Dada a repercussão negativa, temia-se, inclusive, que alguns dos colaboradores deixassem o projeto ou se recusassem a disponibilizar os dados, o que poderia comprometer o desenvolvimento do projeto científico do centro como um todo. Em outro momento, B1 destacou:

A “comoção” durante o workshop começou precisamente quando foi apresentado em um *slide* a proposta de transferência tecnológica (...) Foi um clima de guerra quando se falou de *open science*... Foi colocada uma questão de ordem prática: NeuroMat queria todos os dados de pesquisa disponibilizados. Naquele momento o coordenador foi contundente: ‘ou abre ou sai’. Ficamos com receio, achamos que iríamos perder a equipe [Entrevista com B1].

O workshop mencionado marcou algumas mudanças na forma como o assunto passou a ser tratado internamente, ajustando-se “*uma postura radical*” para uma postura voltada para o trabalho gradual de convencimento. O coordenador do centro foi uma figura central nesse momento, utilizando sua posição de líder do projeto para tentar convencer e envolver os demais. O coordenador é conhecido como um grande entusiasta de assuntos relacionados ao conhecimento aberto.

Alguns informantes afirmaram que o coordenador geralmente estava disposto a negociar e havia assumido uma postura mais conciliadora diante das reações iniciais, ainda que ele fosse um “*ferrenho*” defensor da Ciência Aberta como algo intrínseco ao funcionamento do NeuroMat.

Para não “*assustar*” os pesquisadores colaboradores, passou-se a tratar o assunto de uma maneira menos incisiva. A partir daquele momento houve um processo de negociação do coordenador com esses pesquisadores para tentar conscientizá-los da importância dos modeladores terem acesso aos dados produzidos pelos neurocientistas.

Embora existisse uma certa frustração com a receptividade de alguns pesquisadores, acompanhada de um certo temor de que o centro fosse prejudicado porque a proposta foi colocada como uma prioridade – porém alguns importantes colaboradores do projeto não pareciam estar dispostos a encarar facilmente – a ideia de tornar o NeuroMat comprometido com a Ciência Aberta não foi abandonada totalmente, havendo uma reconfiguração da proposta inicial para comportar os diferentes interesses que se manifestaram naquele momento.

Durante uma oficina sobre Ciência Aberta no I Workshop de Jovens Pesquisadores. Um dos pesquisadores principais, A1, dirigindo-se a um pós-doutorando da área médica que havia acabado de fazer uma extensa fala sobre as dificuldades relacionadas à abertura de dados na sua área, sugerindo que abrir os dados trazia dificuldades de tempo e cuidados, o coordenador pegou o microfone e, assumindo uma postura professoral, de forma a não intimidar o jovem pesquisador, disse que havia adotado um conselho de um colega: “*Se alguém quiser compartilhar seus dados, que compartilhe, isso é muito bom... mas se tem alguém que não quer, tudo bem, que seja assim. A questão é que no futuro vamos ter que compartilhar. É um futuro sem volta*”.

O modelador A1 afirmava em suas manifestações que seu entendimento era de que o acesso público e gratuito é um aspecto natural da ciência e, frequentemente, ele enfatizava, inclusive em apresentações públicas que fazia do NeuroMat, o “*fator investimento público*” como determinante para que os pesquisadores disponibilizassem seus dados. Em outro momento, ele destacou que a disponibilização de ferramentas de maneira que qualquer pesquisador pudesse utilizar colocaria o NeuroMat e os pesquisadores como “*participantes ativos*” no contexto de conhecimento aberto, Ciência Aberta e difusão científica, porém “*era difícil para alguns enxergarem o benefício disso*”.

Ele dizia não entender o porquê de alguns pesquisadores não assimilarem facilmente aquela ideia, que para ele era “*como uma regra*” do seu trabalho. Ele lamentava a falta de preocupação ou interesse nesses assuntos no meio acadêmico, algo que ele acreditava ocorrer muito menos por uma má vontade desses pesquisadores, mas por falta de

conhecimento dos colegas acerca da existência e possibilidades de ferramentas, como as licenças livres para a pesquisa. Também dizia não entender o porquê dos outros CEPIDs da FAPESP não demonstrarem preocupação em compartilhar seus dados e outros materiais como o NeuroMat vinha fazendo. Durante a primeira entrevista com A1, enquanto conversávamos, ele citava alguns projetos de Ciência Aberta que tinha conhecimento. Entre eles um projeto da área de linguística do qual ele participou. Para ele era um projeto que “*fazia um uso excepcional dessas coisas abertas*”. Comparativamente àquele projeto, ele apontava que havia uma infinidade de conteúdo produzidas pelos CEPIDs que poderia ser disponibilizados abertamente nos sites para uso de estudantes e professores, mas lamentava que esta prática não ocorria.

Em mais de um momento, ele sugeria que a pesquisa que eu vinha desenvolvendo poderia ajudá-lo a entender as razões que levavam os outros CEPIDs a não se preocuparem em fazer Ciência Aberta. Era possível observar, nesse sentido, que havia um grande interesse de sua parte em saber outras experiências de Ciência Aberta em outros centros de pesquisa. Em vários momentos, ele perguntou se eu já havia me deparado com alguma iniciativa de outro CEPIDs similar ao que o NeuroMat estava fazendo. Em alguns momentos sugeri que eu sondasse como outros CEPIDs estavam tratando do assunto.

Disponibilizar, abrir os dados é um tabu nos outros CEPIDs. Por que a X do CEPID X⁷⁹ não abre os dados? (...). “Eu acho que não é por ‘maldade’ (sic). Claro, existem alguns pesquisadores que têm má-fé ou interesses próprios, mas eu acho que no geral é mais falta de ‘cabeça’ (sic) para essas coisas. Você pergunta ‘por que não usar creative commons⁸⁰ na sua produção?’, [mas] a pessoa não compreende muito o que eu estou falando” [Entrevista com A1].

A percepção do caráter aberto e gratuito como característica quase natural da prática científica é algo que vi circular com frequência entre pesquisadores. Nas interações com a equipe, notei que tal aspecto era comum entre pesquisadores que possuíam um tempo de carreira acadêmica mais longo e que ocupavam a categoria de pesquisadores principais do projeto, como A1, mencionado acima.

79 Optou-se por ocultar a identidade de pessoas externas ao NeuroMat.

80 *Creative Commons* são licenças abertas que permitem a reprodução e compartilhamento de conteúdo com menos restrições.

Os pesquisadores seniores são justamente aqueles que mais sustentam discursivamente a estratégia de Ciência Aberta do NeuroMat em diferentes espaços e situações. Entre eles é facilmente identificável uma constante reafirmação do *ethos* normativo da ciência quando reiteram que a ciência deve ser aberta. Eles colocam o acesso aberto aos resultados como uma norma da prática científica e como obrigatório para aqueles que desenvolvem pesquisas com recursos públicos.

Falas como a de A3, a seguir, apareceram algumas vezes nas conversas e nas entrevistas com os pesquisadores seniores:

Na minha opinião a ciência tem que ser aberta. Nós que estamos na universidade pública recebendo dinheiro público deveríamos ser obrigados a tornar tudo o que a gente faz acessível, porque isso é uma característica natural da ciência (...)
[Entrevista com A3].

Nos discursos desses pesquisadores, o aspecto normativo da comunicação pública e aberta da ciência se mostrou bastante idealizado. Para convencer os outros colegas, parecia ser suficiente para eles o argumento de que os resultados da pesquisa financiada com recursos públicos deveriam ser abertos. Na prática, se mostrou algo difícil de implementar, demandando uma nova forma de abordar o assunto dentro do grupo.

Como parte das ações de caráter conciliador, os informantes mencionavam a existência de um trabalho de criação de uma cultura interna de Ciência Aberta por meio de um esforço de “convencimento”. A partir daquele momento, elementos que remetiam ao compartilhamento aberto passaram a estar presentes em diferentes formas e maneiras, o que incluiria atividades, infraestruturas e formas de organização perpassadas por elementos relacionados, no entendimento dos atores, como parte da Ciência Aberta.

Nesse aspecto, além de haver preocupações em tornar o assunto como parte da rotina, a equipe do NeuroMat passou a se envolver também com diferentes atividades relacionadas ao que alguns deles chamavam de “*conhecimento aberto*”, o que se refletiu desde a internalização de ferramentas abertas nas atividades, a maneira como o conteúdo do site do centro passou a ser disponibilizado, até a aproximação do NeuroMat com a Fundação Wikimedia, mantenedora da Wikipédia.

3.6 Licenças, recursos informacionais e profissionais da Ciência Aberta

Na segunda visita ao NeuroMat, reparei que a lousa da sala onde ficavam alguns membros da equipe técnica havia uma frase na forma de uma pergunta “*Como disponibilizar o conhecimento para todos?*”

Aquela frase ali, já meio apagada, era o resquício de uma reunião ocorrida dias antes, na qual a equipe havia discutido a estratégia de Ciência Aberta do centro, explicou, um tempo depois, um dos membros da equipe. A pergunta dava margem a várias respostas. Fiquei curioso em saber quais teriam sido as respostas das pessoas que participaram da reunião e se haviam chegado a alguma solução para aquele questionamento.

Ao colocar essa pergunta para um dos técnicos, ele disse que uma das “respostas” que chegaram era que eles deveriam “*espalhar*” os princípios de Ciência Aberta e que “*era necessário disponibilizar a produção interna o mais rápido possível*” com a utilização de licenças *Creative Commons* (CC). Seriam produzidos também vídeos tutoriais para explicar o funcionamento das licenças e um outro explicando a importância da Ciência Aberta. Ele aproveitou para mostrar um vídeo, ainda em processo de produção, com depoimentos dos principais pesquisadores e alguns membros da equipe técnica para fins de divulgação do centro. Enquanto falava do vídeo, o técnico relatava que o grupo tinha decidido que todos os vídeos e outros materiais produzidos seriam disponibilizados em CC como forma de assumir o comprometimento de todos ali com o compartilhamento aberto.

Não se pode deixar de mencionar que, no contexto do NeuroMat, atribui-se um grande valor à utilização de licenças *Creative Commons* para a disponibilização de alguns materiais em acesso aberto, como o conteúdo do site, relatórios de atividades, apresentações, tutoriais, artigos científicos, vídeos, imagens e códigos.

A importância atribuída ao licenciamento dos conteúdos em *Creative Commons* como parte da iniciativa de Ciência Aberta ficou evidente quando o coordenador apontou em diversos momentos de nossas conversas, com certo tom de orgulho, que o site do centro de pesquisa era o “*único dos CEPIDs que estava disponibilizado em Creative Commons*”⁸¹.

81 O site está disponibilizado com em licença *Copyleft Creative Commons* (CC BY 4.0), que permite copiar e redistribuir o material em qualquer meio e formato e permite adaptar e construir sobre o material para qualquer formato, mesmo que comercialmente. As licenças *Creative Commons* são instrumentos importantes no debate sobre abertura e “cultura livre” e estão bastante vinculadas à ideia de reutilização/remixagem de

O site⁸² possui um papel significativo na produção e no entendimento da Ciência Aberta nesse espaço⁸³. A atualização do site ocorre com bastante frequência com divulgação de eventos, publicação dos relatórios de atividades, relatos da participação dos pesquisadores em encontros acadêmicos, divulgação de publicações científicas e *links* para alguns dos projetos, apresentações, tutoriais e notícias. O site passou a registrar as atividades de forma detalhada, produzindo conteúdos extensos como *newsletter*, entrevistas e resumos dos eventos.

A produção científica do centro, majoritariamente composta por artigos de periódicos e *papers* em conferências, é divulgada de maneira sistemática no site e no Facebook, frequentemente acompanhados do *link*, *abstract*, um breve comentário sobre o conteúdo do artigo e breve perfil dos autores. Entretanto, B1 explicou que o foco das ações não era o acesso aberto aos artigos científicos, embora existisse interesse de alguns pesquisadores em criar um repositório em que todas as publicações estejam reunidas. Assim, os esforços se direcionavam para a disponibilização do que estava “à mão” ou disponibilizar conteúdos que não demandam negociações mais complexas, como *slides* utilizados em apresentações, mesmo porque muitos artigos publicados pelos pesquisadores já estão em acesso aberto.

A disponibilização dos *slides* utilizados em apresentações passo a ser tratada como uma parte importante da disponibilização de conteúdo, representando um processo de documentação das atividades. Durante os eventos internos, o coordenador muitas vezes assumia a função de pedir autorização aos palestrantes ou lembrá-los de disponibilizarem os *slides* para serem disponibilizados no site, segundo ele porque os arquivos das apresentações eram “*úteis para outros pesquisadores e professores*”.

O tema dos direitos autorais é uma das questões que frequentemente emergia das discussões sobre disponibilização e produção de conteúdo. O *copyright* é tratado pela equipe de difusão em vários momentos como um obstáculo ou um problema para a produção de materiais, especialmente quando envolve a reutilização de conteúdos produzidos por terceiros. Durante as discussões sobre a produção e disponibilização de materiais educacionais, muitas

conteúdo. Em busca realizada nos sites dos outros 16 CEPIDS, verificou-se que além do NeuroMat, havia mais dois sites estavam centros de pesquisa também licenciados com CC.

82 Cf. F6

83 Os sites mantidos pela equipe têm um papel importante nas atividades do centro, funcionando como uma das principais vias de transferência e difusão de informação. Foram criados alguns portais com a função de disponibilizar informações para diferentes públicos como o site da iniciativa ABRAÇO e AMPARO e um blog da equipe de difusão científica.

vezes os sujeitos recaíam no dilema dos direitos autores sobre determinados conteúdos. Cada vez que alguém propunha alguma atividade utilizando imagens ou vídeos produzidos por terceiros, outra pessoa chamava a atenção para a questão de violação dos direitos autorais. Isso pode ser ilustrado quando o coordenador relatou, durante uma reunião com a equipe de difusão, que uma das alunas de doutorado *“fez um sucesso estrondoso (sic) em uma apresentação para professores e estudantes sobre matemática e neurobiologia”* em uma atividade realizada em uma escola de ensino básico e, por conta disso, ele recomendava que os *slides* da aluna fossem disponibilizados no site de maneira que pudessem ser utilizados por outros interessados.

Naquele momento, sua empolgação deu lugar a uma expressão de frustração quando B1 informou que não era possível disponibilizar as apresentações, pois as *“imagens usadas pela aluna não estavam abertas”* e isso poderia gerar problemas com direitos autorais.

A preocupação com a autoria é algo que permeia também a produção de software, que é uma das principais atividades do centro e possui algumas dinâmicas e questões próprias. Uma delas diz respeito aos direitos autorais na produção do código. O desenvolvedor B2 trouxe algumas vezes esse aspecto. A apropriação do código por outros programadores é algo recorrente entre quem trabalha com desenvolvimento de software e banco de dados. *“Em alguns casos, o desenvolvedor copia o código de alguém e não se sabe onde ele pegou o código ou qual licença ele utilizou”*. Em outro momento B2 explicou que com essa prática: *“Pode-se incorrer contra os direitos autorais, em violação dos direitos autorais”*. Perguntei como ele lidava com essas questões no cotidiano do trabalho no NeuroMat e ele enfatizou que *“O NeuroMat tem a prática de não copiar o código alheio e por isso se passou a ter uma grande preocupação com os aspectos jurídicos e até se consultou um advogado”*.

B2 demonstrava bastante familiaridade com as ferramentas de licenciamento livre, de maneira que em suas formulações sobre seu trabalho de desenvolvimento, ele descrevia o processo de escolha das licenças, ressaltando a importância desses instrumentos para as atividades de desenvolvimento dos sistemas. Mas, como ele mesmo afirmou, sua preocupação com as licenças livres devia-se ao seu trabalho no NeuroMat: *“em anos trabalhando com TI nunca atentei muito para essa coisa de licenças nos sites. Passei a entender o significado das licenças livres, no sentido do simbolismo que têm, com a equipe do CEPID, com A13 e B1”*.

Em sua maioria, os profissionais que integram a equipe técnica assumem papéis essenciais na implementação e circulação dos instrumentos e conhecimentos sobre Ciência Aberta e estão familiarizados, ou diretamente envolvidos, com temas relacionados. Por essa razão, eles também são treinados dentro de uma abordagem que busca desenvolver habilidades necessárias para implementar as iniciativas de Ciência Aberta do NeuroMat de uma forma mais ampla.

Alguns dos principais membros da equipe técnica têm um histórico de atuação com projetos de conhecimento aberto anterior ao trabalho no centro. Esta “*expertise*” os torna atores fundamentais não somente para a forma como determinados temas são inseridos na agenda de atividades do centro como também os autoriza a treinar os outros membros da equipe técnica e pesquisadores acerca desses temas. Um desses profissionais é o cientista social B1, que possui um histórico como membro do grupo de editores da Wikipédia, e atua promovendo a ferramenta em diferentes espaços. Ele é responsável por operacionalizar uma série de atividades do gênero dentro do centro e, muitas vezes, assume a liderança na condução de projetos que envolvem difusão científica e projetos de extensão. Ele, inclusive, possui uma interlocução direta com os pesquisadores, e por meio disso ele envolve muitos deles nas atividades que envolvem o assunto.

Parte do trabalho de convencimento e também de operacionalização das atividades em torno da estratégia de Ciência Aberta é exercida pelos cientistas da computação e desenvolvedores. Nota-se que o papel assumido pelos indivíduos dessas áreas de “*espalhar os princípios de Ciência Aberta*” relaciona-se ao fato de que esse grupo está inteirado de alguma maneira acerca dos principais conceitos de Ciência Aberta, sobretudo na sua interface com o movimento de Software Livre e de Código Aberto.

Os cientistas da computação detêm conhecimentos e técnicas específicos da modelagem computacional e normalmente estão familiarizados com o movimento FLOSS, com ferramentas de código aberto, licenças abertas, entre outros. Por essa razão, aparecem também como responsáveis por operacionalizar as iniciativas de Ciência Aberta do centro. Os cientistas da computação atuam sobretudo na circulação de conceitos, disponibilização de ferramentas e oferecimento de oficinas e palestras, além de explorarem esses conceitos no *design* das infraestruturas de informação do NeuroMat.

Torna-se central nesse processo da operacionalização da iniciativa de Ciência Aberta a proximidade do NeuroMat com pesquisadores ligados ao Centro de Competência em Software Livre (CCSL)⁸⁴. Essa proximidade é também física, já que os prédios dos dois centros se encontram lado a lado. A colaboração entre os dois centros ocorre por meio do treinamento oferecido pelo CCSL para os desenvolvedores e no desenvolvimento das ferramentas computacionais. Alguns pesquisadores do CCSL também participavam dos workshops por meio de oficinas em que algumas ferramentas que auxiliem na disponibilização de conteúdo, como licenças e bases de dados de acesso aberto.

3.7 Abrindo a ciência na Wikipédia

Parte do envolvimento do NeuroMat com atividades de conhecimento aberto e desenvolvimento da cultura interna de Ciência Aberta foi a criação do projeto Wikidifusão, que compreende a utilização de ferramentas abertas e colaborativas da Fundação Wikimedia. O projeto busca compreender principalmente a produção de conteúdo sobre as áreas de atuação do NeuroMat para a Wikipédia.

As ações envolvendo a Wikipédia tiveram início em 2014 como parte do comprometimento do centro de pesquisa com a popularização da ciência e com a Ciência Aberta. Na ocasião, foram lançadas duas notas e um “chamado ao dever” do próprio coordenador com o propósito de convocar os pesquisadores a contribuírem com a representação da neuromatemática na Wikipédia. Como “lição de casa”, os membros e colaboradores do centro foram estimulados a melhorar conteúdos especializados para verbetes da Wikipédia.

84 O CCSL é ligado ao IME/USP. O centro realiza atividades voltadas para o desenvolvimento, uso e divulgação do software livre dentro e fora da USP. CCSL. **Home page**. Disponível em: <<http://ccsl.ime.usp.br/>>. Acesso em 20 mar. 2018.

Tabela 1: Contribuições do NeuroMat na Wikipédia.

Tipo de Edição	Quantidade
Adição de imagem	166
Criação de artigo	3
Criação de verbete	67
Melhoria de conteúdo	83
Melhoria de verbete	1
Total geral	320

Fonte: Repositório NeuroMat na Wikipédia.

A tabela 1 apresenta a quantidade de contribuições feitas pelo NeuroMat à Wikipédia no período de 2014 a 2017⁸⁵. A partir de 2015 as atividades ficaram mais sistemáticas e permanentes. Desde então, os membros da equipe passaram a ser envolvidos no melhoramento de verbetes já existentes na Wikipédia e na criação de novos. Foram estabelecidas algumas metas para as atividades, como a atualização dos artigos sobre neuromatemática (pesquisa, teoria, terminologia) em uma linguagem apropriada para o público em geral e baseadas em informações confiáveis.

Os pesquisadores foram engajados também na avaliação de artigos relacionados à matemática e neurociência, identificando problemas conceituais (NEUROMAT, 2014)⁸⁶.

Entre 2015 e 2017 foram realizadas seis maratonas de edição de verbetes abertas ao público⁸⁷, algumas especificamente para edição de tópicos sobre matemática e neurociência⁸⁸ e treinamentos em difusão científica, com foco nas aplicações da Fundação Wikimedia.

85 Cf. F21

86 Os aplicativos Wiki são ferramentas abertas, existe toda uma infraestrutura de software, acesso, editores, e toda uma comunidade em voltado, assim como uma Fundação portadora de recursos.

87 WIKIVERSIDADE. Pesquisa em Difusão Científica do CEPID NeuroMat. S/d. Disponível em: <https://pt.wikiversity.org/wiki/Pesquisa_em_Difus%C3%A3o_Cient%C3%Adfica_do_CEPID_NeuroMat>. Acesso em 20 mar. 2018.

88 A equipe de difusão passou a realizar oficinas de edição e treinamento sobre as ferramentas *wiki* para outros centros de pesquisa. Já foram realizadas oficinas de edição de conteúdo sobre anatomia veterinária e sociologia da violência.

As “editatonas” ou maratonas de edição são geralmente abertas ao público, integrando as atividades de divulgação e educação científica. Alguns verbetes produzidos pela equipe fazem parte do projeto “Wikipédia falada”, que consiste na audiodescrição de verbetes de probabilidade e estatística com o objetivo de facilitar o acesso às fórmulas de probabilidade estatística por pessoas com deficiência visual⁸⁹.

As maratonas envolvem atividades práticas de edição e melhoria de conteúdo, assim como treinamento sobre os conceitos e funcionalidades das ferramentas da Fundação Wikimedia. A Figura 4, abaixo, mostra o cartaz de divulgação de uma das maratonas de edição.

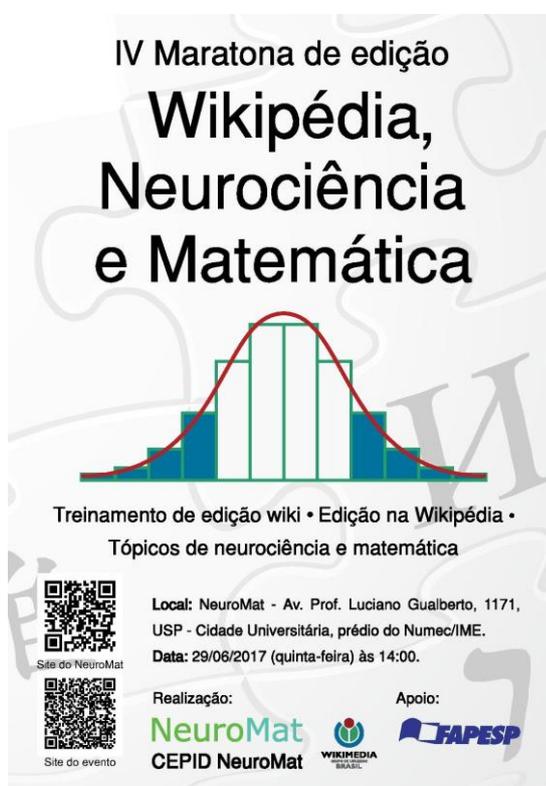


Figura 4: Cartaz da IV Maratona de edição.

Fonte: NeuroMat

A equipe envolvida com o projeto Wikidifusão faz a publicação de conteúdo multimídia (figuras, fotos, imagens animadas e vídeos) no repositório de mídias Wikimedia, a *Wikimedia Commons*. Destaco a edição do verbete “classificação de disparos neuronais”⁹⁰ que envolveu a produção de um vídeo explicativo que ilustra a versão em português e inglês do

89 TRAÇO DE CIÊNCIA. **CEPID NeuroMat cria nova versão da página Projeto Wikipédia Falada**, 7 fev. 2017. Disponível em: <<https://difusaoneuromat.wordpress.com/2017/02/07/cepid-neuromat-cria-nova-versao-da-pagina-projeto-wikipedia-falada/>>. Acesso em 20 mar. 2018.

verbete. As fotos de eventos e dos membros da equipe também são publicadas imediatamente na *Wikimedia Commons*.

O NeuroMat também estabeleceu uma parceria com o Museu de Medicina Veterinária da Universidade de São Paulo para carregar e indexar na *Wikimedia Commons* um conjunto de imagens produzidas pelo museu, dentro da abordagem Open GLAM⁹¹. Houve algumas experiências com a utilização da Wikipédia em sala de aula, em escolas de ensino médio e na disciplina de graduação “Neurociência computacional” oferecida pela USP de Ribeirão Preto. Durante o período da pesquisa de campo estava também em desenvolvimento um curso online de “Introdução ao jornalismo científico” na plataforma de cursos Wikiversidade.

3.8 Ferramentas abertas e a organização da neuromatemática

O envolvimento do NeuroMat com práticas de Ciência Aberta passou a ser feito de diferentes maneiras. Uma das formas significativas por meio da qual ocorre esse envolvimento passa a ser por meio da introdução da temática e assuntos correlatos em palestras e oficinas internas, o que está relacionado à importância atribuída, pelos pesquisadores seniores que apoiam essa prática, à formação dos pesquisadores mais jovens dentro dos princípios de compartilhamento aberto. As ações de formação voltam-se especialmente para mestrandos e doutorandos.

Aqui trazemos algumas evidências importantes da pesquisa de campo que dizem respeito à relação da estratégia da Ciência Aberta com a institucionalização e formação da disciplina científica da neuromatemática. As atividades de formação voltadas para os jovens pesquisadores revelam um aspecto fundamental do desenvolvimento dessa cultura interna de Ciência Aberta.

Algumas evidências importantes da pesquisa de campo sugerem uma associação entre as práticas de compartilhamento que se estabelecem nesse espaço com um aspecto de organização da disciplina científica da neuromatemática. Durante a pesquisa de campo, notei que esse aspecto se manifesta de duas maneiras principais.

91 GLAM é o acrônimo de Galleries, Libraries, Archives and Museums. Open GLAM refere-se a projetos que promovem o acesso livre e aberto à produção cultural e intelectual existente nessas instituições por meio da digitalização e disponibilização online.

A primeira diz respeito a uma associação da neuromatemática com práticas de abertura. A Ciência Aberta aparece em muitos momentos entrelaçada com o propósito mais amplo do centro de pesquisa em treinar pesquisadores e⁹² de *“formar uma geração de novos pesquisadores em neurociência com essa perspectiva do compartilhamento”*, explicou o pesquisador sênior A3.

A prática também expressa um interesse dos pesquisadores seniores que lideram o projeto de que a disciplina da neuromatemática venha a ser caracterizada ou organizada em torno de ferramentas abertas e compartilhamento aberto de dados, diferenciando-a da neurociência, em que as práticas de abertura são menos uniformes e onde os neurocientistas *“não entendem”* a importância de tais práticas.

Ao incorporar esse elemento como parte da formação dos jovens pesquisadores, A3 afirmou que isso *“coloca a ciência aberta como parte dessa área nascente que é a neuromatemática”*, e treina-se pesquisadores com uma *“percepção aberta”* às práticas de compartilhamento, demarcando-se a neuromatemática como uma disciplina relacionada à reutilização e a reprodutibilidade experimental, possibilitadas pelas práticas de Ciência Aberta.

O aspecto da organização disciplinar se manifesta com bastante evidência nas atividades que empregam as ferramentas *wiki*. As atividades de produção de conteúdo para Wikipédia passaram a acontecer como uma forma de *“demarcação online”* do campo da neuromatemática. Como detalhado anteriormente, os pesquisadores têm produzido uma grande quantidade de conteúdo para a Wikipédia. No processo de construção dos verbetes, são inseridas informações científicas baseadas nas áreas de especialização do centro e melhoria de verbetes de saúde relacionados à área de interesse dos pesquisadores (ALVES et al, 2016).

92 A preocupação com a formação dos pesquisadores dentro de novos métodos e abordagens para a neurociência é algo bastante presente entre os pesquisadores seniores, o que se manifesta na forma como os jovens pesquisadores são inseridos nas atividades. Parte dessa formação envolve preparar os Jovens Pesquisadores para dialogarem com uma comunidade científica internacional. Praticamente todos os workshops e palestras são ministradas em inglês. Nesses momentos, todos apresentam suas pesquisas e interagem em inglês, desde docentes aos alunos de iniciação científica. Em parte isso ocorre porque costuma haver não-falantes de português acompanhando as atividades e, ao mesmo tempo, é uma forma de treinar os pesquisadores mais jovens a se integrarem a um circuito de pesquisa onde a língua falada é o inglês. Essa prática de falar em inglês nas atividades em grupo é algo comum especialmente entre os matemáticos.

Alguns pesquisadores passaram a fazer também o trabalho de verificação de conteúdo antes da disponibilização. O pesquisador A2, por exemplo, contou como se dava seu grande envolvimento com essas atividades, o que inclui a revisão de verbetes de neurociência e apresentação dos mesmos em congressos.

Embora essas atividades estejam inseridas em atividades de difusão e divulgação científica, os usos das aplicações da Fundação Wikimedia no NeuroMat fazem parte de um processo bastante específico relacionado à concepção e delimitação da disciplina da neuromatemática – às vezes considerada ou contestada como uma nova disciplina ou área do conhecimento pelos próprios membros do grupo.

O primeiro relatório do projeto explica que o trabalho envolvendo a criação de conteúdo sobre neuromatemática na Wikipédia se justificaria tanto pelo alcance da enciclopédia global, potencializando a difusão de informações, e o entendimento de que a Wikipédia e seus usos desenvolvidas nesse espaço relaciona-se ao esforço mais amplo do NeuroMat em disseminar a Ciência Aberta funcionando como “uma estratégia genuinamente eficiente para esclarecer e aperfeiçoar conceitos recém-concebidos no âmbito de uma área científica recém-projetada” (NEUROMAT, 2014, p. 27).

No processo de produção de conteúdo para a Wikipédia, os pesquisadores, que possuem jargões e conhecimentos distintos, realizam um processo de identificação e seleção da terminologia da área (a fim de selecionar um vocabulário comum), e de definição dos conceitos-chave do campo, que envolve explicações matemáticas para fenômenos neurofisiológicos que consideram como parte do campo de estudos da disciplina. Por meio dessas práticas, a produção de conteúdo para a Wikipédia, por exemplo, não ocorre de maneira deliberada, mas insere-se em um conjunto de práticas, técnicas, sociais de organização de um centro de pesquisa e de uma disciplina científica novos (BOWKER, 2005). Nesse sentido, abrange processos de “manutenção da memória” da disciplina, sendo empregada na documentação do conhecimento produzido, classificação e delimitação do campo em espaço de acesso aberto e colaborativo e altamente visível na Internet. Esses processos, por sua vez, funcionam como uma prática documental agregativa que, com a função semelhante de artigos científicos, relatórios de pesquisa e tutoriais, instruem o grupo acerca das práticas e referências (BOWKER, 2005; GEIGER; RIBES, 2011).

Sem a intenção de fazer um aprofundamento nos usos específicos da Wikipédia e de outras ferramentas *wiki*, dada a complexidade e a extensão dessas atividades que fogem do objetivo central da tese, algumas evidências empíricas dos usos das ferramentas *wiki* direcionam o olhar para esse processo de organização disciplinar no qual se inserem também outras ferramentas de Ciência Aberta e também o treinamento dos pesquisadores.

Em 2015, participei de uma longa reunião com a equipe de difusão na qual foi discutido com os pesquisadores seniores e outros colaboradores do centro sobre a definição de estratégias e atividades relacionadas aos Eixos de educação⁹³ e difusão científica do projeto. Na ocasião, discutiu-se principalmente a organização do I Workshop de Jovens Pesquisadores, que seria realizado dali a menos de um mês.

A organização do workshop já trazia esse propósito de formar os pesquisadores com a perspectiva das ferramentas abertas. Foi o que enfatizaram dois professores antes do início da reunião quando perguntei o que motivava a realização de um workshop especificamente para jovens pesquisadores. Com essa pergunta eu buscava entender o porquê de todos parecerem estar tão atarefados e preocupados com o evento, afinal organizar eventos faz parte da rotina de quem trabalha na universidade. A2 e A3, os dois pesquisadores principais envolvidos com o workshop, explicaram que o NeuroMat tinha vários alunos de pós-graduação e pós-docs em São Paulo e Rio de Janeiro que não sabiam o que era o centro. O objetivo portanto era reunir esses jovens pesquisadores para que conhecessem o centro, apresentassem as pesquisas para serem comentadas por pesquisadores mais experientes e que pudessem “*vestir a camisa do centro*”.

Aquela reunião na qual eu participava tinha justamente o objetivo de discutir a inserção de determinados temas na programação do workshop: “*A gente quer educar os jovens pesquisadores com essa perspectiva do NeuroMat, que inclui a coisa de compartilhamento*”. Com esse objetivo, também faria parte da programação do workshop uma apresentação sobre Ciência Aberta, incluindo uma oficina de editoração de verbetes da Wikipédia⁹⁴.

93 Durante o debate sobre as estratégias de educação e difusão, foram frequentes as questões identitárias da área e de como as atividades voltadas para a formação de alunos e professores também contribuiriam para a divulgação do conhecimento da neuromatemática e expertises do NeuroMat em escolas por meio da realização de oficinas.

94 A convite de B1, um representante do Grupo de Usuários da Comunidade Wikimedia no Brasil participou da reunião para apresentar dos projetos da Fundação e fazer uma demonstração das funcionalidades da Wikipédia e possibilidades de utilização no meio acadêmico. O representante era o linguista B4, que posteriormente passou a fazer parte da equipe do NeuroMat como assistente da equipe de difusão. É

Durante a discussão sobre a organização da oficina de edição houve um debate sobre os conteúdos que os jovens pesquisadores editariam durante o workshop e que passariam, portanto a estar abertos na Wikipédia como uma perspectiva do NeuroMat para determinados assuntos. Ao longo da reunião, percebia-se em vários momentos uma preocupação com a efetividade da oficina tanto no aspecto da produção de conteúdos que de fato representassem o centro como com aspecto instrutivo da atividade no sentido de transmitir os princípios de Ciência Aberta.

A1 reforçou recorrentemente durante a reunião que ele queria uma atividade que *“fizesse com que os alunos se apropriassem da ideia central do projeto”*. Ele tinha dúvida sobre o que se ganharia em termos de produção de conteúdo sobre o tema inserindo uma atividade como aquela em uma oficina com jovens pesquisadores. Ao mesmo tempo que se dizia fascinado pelas funcionalidades e possibilidades de uso da Wikipédia apresentadas por B4, ele se colocou como alguém que tinha uma *“visão utilitarista”*, preocupada em fazer com que a atividade contribuísse para o entendimento dos jovens acerca de conceitos da área.

O termo “Lesão do Plexo Braquial” (LPB) foi colocado como possibilidade de edição. Mas imediatamente ocorreu um debate sobre a pertinência de abordá-lo na oficina de edição e como seria a abordagem da neuromatemática para o verbete. Uma das preocupações colocadas por A1 era o “enlace” entre matemática e neurobiologia. Ele tinha dúvidas se os pesquisadores que editariam os verbetes dariam conta da grande complexidade do conteúdo matemático durante a oficina, e se o termo LPB poderia ser considerado parte do vocabulário da área, uma vez que no estágio inicial que estavam nas pesquisas sobre LPB ainda não havia muitas referências produzidas.

Enquanto ouvia o colega falar, A3 abriu o navegador e digitou “Plexo Braquial”. A busca no Google retornou entre os primeiros resultados o verbete da Wikipédia. Segundo ele, o conteúdo do verbete da Wikipédia estava “muito pobre”, porém ele enxergava a possibilidade de utilizar na divulgação do conhecimento da neuromatemática. Nesse momento ele apresentou a sugestão de reunir os neurocientistas que estudavam a LPB e, posteriormente, se fosse necessário, solicitava-se aos pesquisadores que trabalhavam com modelagem matemática que verificassem a qualidade do conteúdo.

importante ressaltar que B1 também faz parte do Grupo de Usuários Wikimedia no Brasil, sendo um dos membros mais atuantes na edição de conteúdo e promotor da Wikipédia como ferramenta de ensino e divulgação científica. Dentro do NeuroMat, ele tem sido responsável por supervisionar o projeto de Wikidifusão e a equipe de bolsistas do projeto.

Para além de um aspecto académico, é interessante essa coisa da *wiki* porque: vai que algum famoso tem lesão do Plexo Braquial? Com certeza todo mundo vai querer saber o que é isso. Podemos achar que lesão do Plexo Braquial não é importante para os nossos interesses aqui, mas com as ferramentas que eles [Os Jovens Pesquisadores] vão aprender, eles podem voltar para suas casas e editarem o que eles quiserem. Os jovens podem falar de modelagem de qualquer coisa; isso já é uma contribuição, mas se a gente achar melhor, podemos induzir deixando claro que aqui o centro tem uma abordagem e temos que ter essa preocupação. No terceiro dia [do workshop] podemos editar os verbetes, mas o mais importante é a discussão sobre essas coisas, a coisa da Wikipédia, do compartilhamento. Podemos achar que lesão do Plexo Braquial não é central aqui, mas de algum modo está presente aqui no NeuroMat. Dentro de um verbete maior vai haver seções para inserir a abordagem matemática, a abordagem computacional, biológica. [Registro de comentário de A1]

Destaco um momento ocorrido posteriormente, já na oficina de edição, em que houve um grande impasse entre os pesquisadores seniores líderes do projeto em torno do verbete “neuromatemática”. Os três pesquisadores ficaram um bom tempo para chegar em um acordo em relação ao conteúdo do verbete. A falta de consenso estava em definir os objetos e objetivos da neuromatemática. Após algumas horas de discussão, escrita e reescrita, ao final da atividade, o verbete ficou com um pouco mais de uma linha escrita. De acordo com A3, o principal ponto de conflito naquele momento era: *“em que medida a neuromatemática não se iguala ao que se convencionou chamar de ‘neurociência computacional’ ou ‘neuroinformática’”*.

Um aspecto importante desses processos é que, nesse espaço aberto e colaborativo, os pesquisadores podem empregar conhecimento especializado e autoridade científica para produzir “conteúdo de qualidade” e assim ocupar esse espaço favorecendo a circulação do conhecimento produzido pelo centro. Destacam-se, por exemplo, representações visuais de conceitos e instrumentos matemáticos utilizados na modelagem neuronal. Algumas dessas representações inclui o modelo matemático “Galves-Loecherbach”, criado por pesquisadores do centro, a “lei de Hick”, “lei dos grandes números” e “classificação de disparos neurais”⁹⁵.

95 Cf. F21

A Figura 5, a seguir, é um exemplo de representação gráfica animada criada pela equipe para representar o conceito “lei dos grandes números”⁹⁶, disponibilizada no Wikimedia Commons.

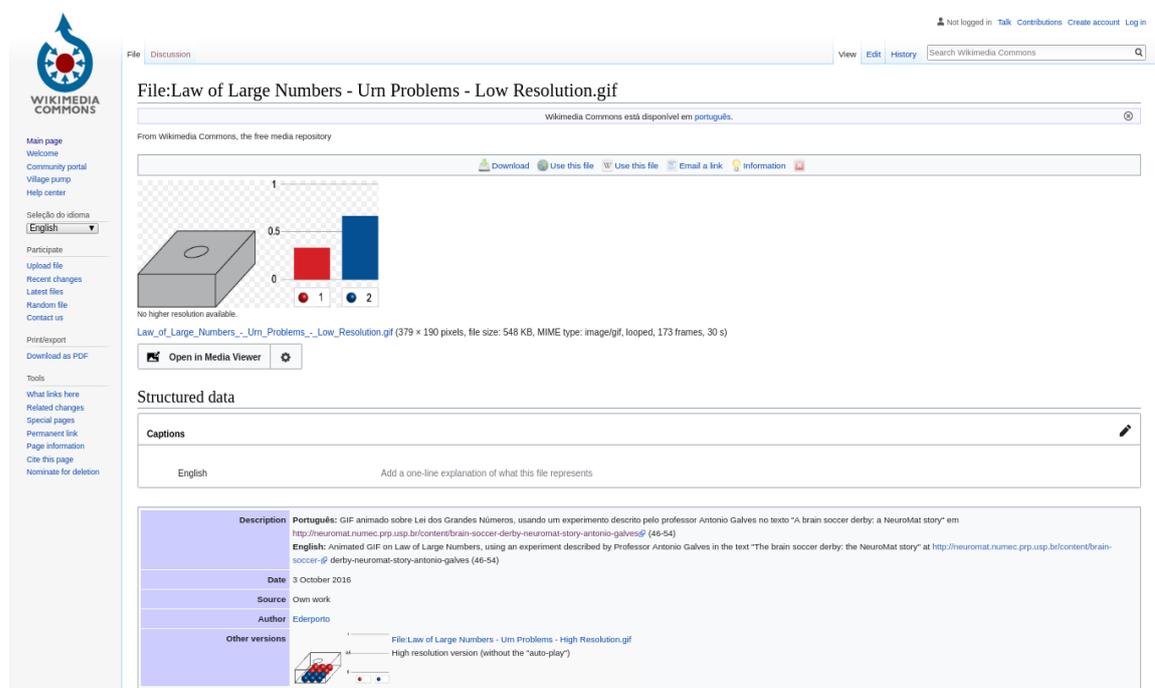


Figura 5: Animação ilustrativa do conceito matemático "Lei dos Grandes números".

Fonte: Wikimedia Commons

Inseridas nesses processos de produção de conteúdo científico, as *wiki* respondem a uma necessidade concreta do CEPID em relação à produção de indicadores de difusão científica. Dada a força tarefa que foi estabelecida para a produção de conteúdos de autoria do centro, passou-se a obter uma série de indicadores de atividades na Wikipédia.

96 Cf. A lei dos grandes números é um conceito da teoria da probabilidade que descreve o resultado da realização de um evento realizado repetidas vezes. Cf. <https://pt.wikipedia.org/wiki/Lei_dos_grandes_n%C3%Bameros>.

As atividades com as ferramentas *wiki* ganharam grande centralidade nas atividades de difusão a partir do momento que o projeto recebeu uma avaliação positiva na avaliação periódica da FAPESP, em parte porque os números do projeto (acessos, edição e inserção de conteúdo, oficinas) em termos de alcance de pessoas foram satisfatórios. Além disso, alguns verbetes ocuparam posição de destaque na comunidade de editores da Wikipédia por serem considerados “bem escritos” e “de qualidade”. Esses fatos e números possibilitaram ao centro tornar-se, e colocar-se, como o maior produtor de conteúdo matemático para Wikipédia lusófona⁹⁷.

Considerações Finais

O capítulo apresentou o Centro de Pesquisa, Inovação e Difusão em Neuromatemática. O NeuroMat configura-se como um espaço multidisciplinar que desenvolve pesquisa avançada em modelagem de dados neuronais e realiza atividades de difusão científica e transferência de conhecimento e tecnologia. O centro insere-se no campo da pesquisa sobre o funcionamento do cérebro, atuando em modelagem de dados. O capítulo detalhou aspectos dessa estratégia, destacando os fatores que a originaram, o papel de diferentes ferramentas abertas e dos profissionais envolvidos. Desenvolvem-se aí um conjunto de atividades próprias associadas a uma estratégia de Ciência Aberta.

No capítulo a seguir, apresento os resultados do trabalho de campo que retratam a aproximação com as práticas dos pesquisadores do centro. A discussão é alimentada com a caracterização detalhada das práticas de modeladores e neurocientistas. Esse entendimento é fundamental para compreender como são produzidos e utilizados os dados neuronais e quais são suas propriedades e fronteiras, ajudando assim a analisar como os diferentes grupos se relacionam com a Ciência Aberta nesse espaço.

97 Cf. F12

4 PRÁTICAS DE COMPARTILHAMENTO DE DADOS: FRONTEIRAS DISCIPLINARES E CULTURAS EPISTÊMICAS

O capítulo anterior apresenta discussão sobre as configurações do campo e elementos que caracterizam a Ciência Aberta no NeuroMat. Neste capítulo apresento elementos sobre as práticas de pesquisa dos neurocientistas e modeladores.

Ao longo do capítulo, busco descrever o trabalho etnográfico realizado entre neurocientistas teóricos e experimentalistas, enfocando nos processos de produção e compartilhamento de dados desses pesquisadores. Em seguida, relata a aproximação com os modeladores, descrevendo os processos de aplicação de ferramentas matemáticas na produção de modelos como forma de ilustrar o trabalho desses sujeitos e as ferramentas que empregam na modelagem de dados.

4.1 Neurocientistas e práticas de compartilhamento

Knorr-Cetina (1999a) compreende que a maneira como as comunidades científicas “sabem” está fortemente ligada à natureza dos instrumentos utilizados; ao significado e importância dos dados de pesquisa; à maneira como as comunidades se estruturam; aos meios de comunicação que os cientistas empregam. Em outros termos, está ligado às culturas epistêmicas que as comunidades científicas constituem ao longo dessas relações. Nessa perspectiva das culturas epistêmicas, os modos como os cientistas arquivam e compartilham seus dados podem variar consideravelmente e refletir questões, preferências e premissas mais amplas.

A autora aponta que a colaboração científica em larga escala em torno de instrumentos de laboratório complexos, onde os instrumentos e os resultados pertencem a todos, se dá em campos disciplinares com histórico e cultura de compartilhamento de dados

bem estabelecidos⁹⁸. Outros autores apontam que disciplinas que contam com práticas mais dispersas e artesanais tendem a possuir tradições marcadas por práticas, normas e convenções específicas para equipes e laboratórios, e até mesmo para estilos pessoais de trabalho (BORGMAN 2015; MILLERAND, 2011).

Em se tratando do compartilhamento de dados de pesquisa, Millerand (2011) e Bowker (2000) referem-se a uma dimensão política e ética implicada em tal processo. Descrevendo especificamente a representação de dados de pesquisa em repositórios públicos, os autores apontam para a ação de diferentes tradições ou diferentes culturas na constituição da gestão e compartilhamento de dados em repositórios científicos. Tais culturas atuam como padrões e convenções que governam as representações dos pesquisadores sobre sua pesquisa, suas práticas e resultados do seu trabalho.

Millerand (2011) chama atenção para a instrumentação dos processos de preparação e compartilhamento de dados em repositórios como algo que permite identificar as fronteiras disciplinares, evidenciando a grande disparidade de práticas no que se refere aos dados e, portanto, à preponderância das culturas epistêmicas.

O desenvolvimento da infraestrutura é acompanhado por um processo de instrumentação das atividades de publicação e compartilhamento de dados que permite observar a existência de culturas epistêmicas altamente diferenciadas dentro do coletivo de pesquisadores, e que intervêm nos relatos dos pesquisadores a respeito dos seus dados de pesquisa. Em termos mais gerais, as novas práticas associadas ao uso da infraestrutura levantam a questão da possibilidade de uma reorientação do olhar dos cientistas para os seus dados de pesquisa (2011, p. 12).

Essa diferenciação fica ainda mais exposta em projetos multidisciplinares – como é o caso do NeuroMat – nos quais ocorrem muitas vezes confrontações entre culturas epistêmicas diferentes.

A interação com os atores durante a pesquisa de campo no NeuroMat mostrou que a forma como os pesquisadores se relacionam com a abertura dos resultados é fortemente dependente da natureza da pesquisa e das relações diferenciadas que os pesquisadores estabelecem com os seus dados, que por sua vez são elementos intrinsecamente ligados às tradições epistêmicas em que os indivíduos estão inseridos. Essa complexa relação pode ser observada explicitamente entre os pesquisadores neurocientistas.

98 Este é o caso da física de altas energias, estudada por Knorr-Cetina (1999a), área em que as equipes de pesquisa podem reunir dezenas de pesquisadores em torno de instrumentos gigantescos (aceleradores de partículas) e onde os dados “pertencem” a todos.

A constante menção aos dados experimentais foi um ponto que me despertou a atenção desde o início do trabalho de campo. Percebendo essa agência dos dados, busquei aproximação com os neurocientistas para compreender suas práticas, sobretudo a produção, os processos e as atividades que remetem ao compartilhamento aberto de maneira mais ampla.

Por ser um campo de investigação muito amplo, a neurociência é marcada por uma grande variedade de métodos, abordagens e teorias. A pesquisa experimental é parte fundamental do processo de compreensão dos fenômenos neuronais e produz uma grande quantidade de dados neurofisiológicos. Um dos dilemas da neurociência é desenvolver instrumentos capazes de tratar a grande quantidade de dados experimentais produzidos nas mais diferentes subdisciplinas que compõem a área. De um lado aponta-se uma grande capacidade da área em coletar dados sobre o funcionamento do sistema nervoso e, do outro lado, aponta-se uma carência de instrumentos capazes de analisar e interpretar esse volume de resultados. Por essa razão, a neurociência é comumente descrita como uma área “rica em dados e pobre em teoria”, e isso precisa ser considerado no desenvolvimento desse campo.

Os dados produzidos no campo da neurociência vão muito além de dados sequenciais, numéricos e textuais, e englobam, por exemplo, imagens morfológicas e gravações fisiológicas, como ilustrou um pesquisador em um workshop com um *slide* no qual havia várias imagens que mostravam diferentes experimentos realizados representativos da pesquisa experimental em neurociência: um rato dentro de uma gaiola representando a experimentação animal; uma cabeça com uma touca com eletrodos representando a encefalografia; um cientista em frente a uma tela de computador em que eram exibidas imagens de ressonância magnética do cérebro. As pesquisas em neurociência também empregam diferentes equipamentos na realização dos experimentos eletrofisiológicos como eletroencefalograma e estimulação Magnética Transcraniana, de neuroimagens e comportamentais.

Como ficou evidente nos discursos sobre a abertura de dados experimentais, não apenas os neurocientistas do NeuroMat como a comunidade de neurociências, de maneira geral, é caracterizada por uma acentuada resistência ao compartilhamento aberto dos dados. Apesar de ser apontada uma relativa mudança em relação a uma cultura do compartilhamento na área, a aceitação entre os neurocientistas ainda é lenta e desigual. Em grande medida, essa lentidão está relacionada com uma certa complexidade atribuída à produção de dados (KON; VARGAS, 2014; BORGMAN, 2012).

A forma como os neurocientistas se relacionam com os dados, como apresentam resistências ao compartilhamento aberto, remete muitas vezes a essa complexidade, representada em muitos momentos pela natureza da pesquisa (se é experimental ou teórica), à origem dos dados (humano ou animal, por exemplo), o tipo de técnica experimental, equipamentos empregados, e mesmo ao ambiente onde a pesquisa é desenvolvida.

Em diversos casos, a pesquisa experimental em neurociência se insere no campo da pesquisa em saúde e, geralmente, envolve seres humanos, o que remete à afirmação de Millerand (2011) de que problemáticas em relação ao compartilhamento aberto de dados envolvendo questões de confidencialidade são mais comuns para dados de pesquisa em saúde e ciências sociais. Sobre isso, Santos et al (2017, p. 100) argumenta que:

O debate sobre abertura de dados ganha ainda mais relevância e complexidade quando se trata do campo da saúde, onde são desenvolvidas pesquisas de caráter sensível e também estratégico em muitos sentidos, incluindo a produção de conhecimento e tecnologias de importância para a saúde pública e de interesse e potencial mercadológico. A área de saúde também abrange um campo vasto de pesquisas e ações, o que implica reunir práticas de produção e comunicação científicas bem distintas entre si e pautadas por interesses muitas vezes contraditórios.

Desse modo, por ser um ramo fundamental da produção de conhecimento em saúde, os dados clínicos e da pesquisa experimental em neurociência estão na maioria das vezes restritos e “isolados” de acesso por requisitos legais, éticos, de segurança e confidencialidade (NOOR et al, 2015).

Kon e Vargas (2014) apontam a neurociência como um campo em que o compartilhamento aberto não se estrutura como uma prática trivial. De modo geral, tanto a coleta quanto o arquivamento dos dados são feitos de modo pouco estruturado. Existe ainda uma grande variabilidade dos dados coletados nas pesquisas. Os bancos de dados nessa área precisam incluir não apenas os dados brutos, mas informações que envolvem a forma e comportamento dos neurônios individuais, fornecidos por meio de medidas de funcionamento do cérebro e por medidas comportamentais, o que complexifica o trabalho necessário à disponibilização.

As práticas dos neurocientistas dos NeuroMat refletem muitas dessas diferenças. Os modeladores e outros profissionais técnicos que estiveram envolvidos com a estratégia de Ciência Aberta apontaram em vários momentos que os neurocientistas costumam apresentar uma resistência maior à disponibilização de dados e foram apontados em muitos momentos como barreiras à implementação da estratégia de Ciência Aberta.

De acordo com os entrevistados, a resistência à abertura dos resultados, especialmente entre os neurocientistas, têm relação com o perfil profissional ou a formação acadêmica desses pesquisadores. Os próprios pesquisadores dividem os colegas neurocientistas em dois grupos no momento que relataram como era a relação interna com os neurocientistas em relação à disponibilização de dados: os “neurocientistas experimentalistas” abarcariam os pesquisadores que trabalham com coleta de dados mais diretamente, aqueles que lidam com a pesquisa observacional. Do outro lado estariam os “neurocientistas teóricos”, que são descritos como pesquisadores que possuem uma formação mais de pesquisa básica e teórica e empregam instrumentos estatísticos e computacionais para fazerem análises de dados experimentais.

Essa diferenciação está bastante relacionada às especializações desses grupos dentro da grande área da neurociência, apresentando formações e métodos científicos diferentes, o que se reflete na maneira como interpretam a Ciência Aberta e, especificamente, como se relacionam com o compartilhamento de dados. Por consequência, entende-se que um dos fatores que contribuem para a resistência dos neurocientistas à disponibilização de dados tem a ver com os tipos de experimentos realizados pelos pesquisadores desses dois perfis.

Busquei, tanto quanto possível, uma aproximação com o trabalho dos pesquisadores neurocientistas para compreender as culturas epistêmicas subjacentes. A seguir apresento uma descrição das práticas desses grupos a partir de questões e temas específicas que emergem dessa interação.

4.1.1 Neurocientistas teóricos

Os neurocientistas teóricos costumam expor preocupações que passam geralmente por questões relativas à utilização ou interpretação dos dados originais. Na neurociência, conforme caracteriza Vargas e Kon (2014), as reações contrárias ao compartilhamento aberto de dados estão no argumento de que os dados brutos possuem uma complexidade para serem

interpretados por outros pesquisadores fora do domínio disciplinar e contexto de sua produção. Essa complexidade gera preocupações e críticas acerca da má utilização ou má interpretação dos dados originais que se distancie dos resultados originais.

Para compreender como ocorre a experimentação e a produção dos entre esses pesquisadores, visitei o Instituto de Neurologia Deolindo Couto (INDC), localizado na Universidade Federal do Rio de Janeiro, onde está localizado o Núcleo de Pesquisa em Neurociências e Reabilitação – um dos colaboradores do NeuroMat. Esse grupo é liderado por neurocientistas de tradição teórica.

As pesquisas desenvolvidas pelos neurocientistas teóricos referem-se principalmente à pesquisa sobre plasticidade do cérebro envolvendo a Lesão do Plexo Braquial (LPB). O plexo braquial é um grupo de nervos que se estendem – como um feixe – da medula espinhal pelo membro superior e está associado aos movimentos e a sensibilidade dessa região do corpo. Os modeladores utilizam os dados produzidos nessas pesquisas para o desenvolvimento de modelos da neuroplasticidade⁹⁹ no contexto da LPB. O desenvolvimento do primeiro módulo de criação do software para gestão e compartilhamento de dados experimentais teve como base as pesquisas desenvolvidas nesse laboratório e foi centrado especificamente nos dados neurofisiológicos da Lesão do Plexo Braquial (LPB)¹⁰⁰.

No INDC era perceptível a sobreposição entre o ambiente hospitalar e o ambiente acadêmico. Essa mistura era marcada pelos letreiros indicando as salas de grupos de pesquisa e pelos pacientes que aguardavam atendimento médico ou para realizar exames de eletroencefalografia, por exemplo. Durante a interação com os pesquisadores ligados NPNR, eles chamaram atenção para o fato de que a maior parte das pessoas que participam dos estudos realizados pelo grupo são pacientes do hospital no qual o laboratório está localizado fisicamente e, geralmente, são pacientes que passaram por intervenção cirúrgica e estão em acompanhamento médico. Por essas razões, esse ambiente traz uma série de especificidades para os dados coletados e para a forma como esses dados são manejados e tornados públicos. O que foi colocado por esse grupo foi que havia aspectos mais amplos da “*ética em pesquisa*” fortemente presentes na forma como eles pensavam o desenvolvimento da pesquisa. O contexto de um hospital de neurologia e a participação de pacientes nas pesquisas traziam implicações para o compartilhamento dos dados.

99 Neuroplasticidade ou plasticidade do cérebro refere-se à capacidade das células se adaptarem ou se reconstituírem, por exemplo, quando ocorre uma lesão cerebral.

¹⁰⁰ O software será abordado no próximo capítulo.

Durante a minha visita ao INDC/UFRJ, o neurocientista A2, que trabalha em colaboração com médicos, proveu uma perspectiva detalhada de como os colegas geralmente abordavam o compartilhamento de dados. Os dados produzidos nas pesquisas são muitas vezes descritos como “sensíveis”, pois podem envolver informações pessoais que atingem a privacidade do paciente. Por essa razão, os neurocientistas ligados à prática clínica geralmente são contrários à disponibilização aberta dos resultados alegando que frequentemente se relacionam com aspectos da ética clínica e preocupações legais mais amplas que permeiam a pesquisa em saúde

A maneira dos pesquisadores relatarem suas pesquisas sobre a Lesão do Plexo Braquial (LPB) revela muitos desses aspectos considerados sensíveis. Eles utilizam uma série de imagens nas apresentações que ilustram o nervo e as partes do corpo que são afetadas pela lesão desses nervos. Além desses recursos visuais, eles apontam para regiões do próprio corpo com o intuito de deixar clara a localização dos nervos no corpo humano. Na Figura 6, a seguir, as linhas coloridas indicam as regiões do braço humano por onde se estendem os nervos do plexo braquial.

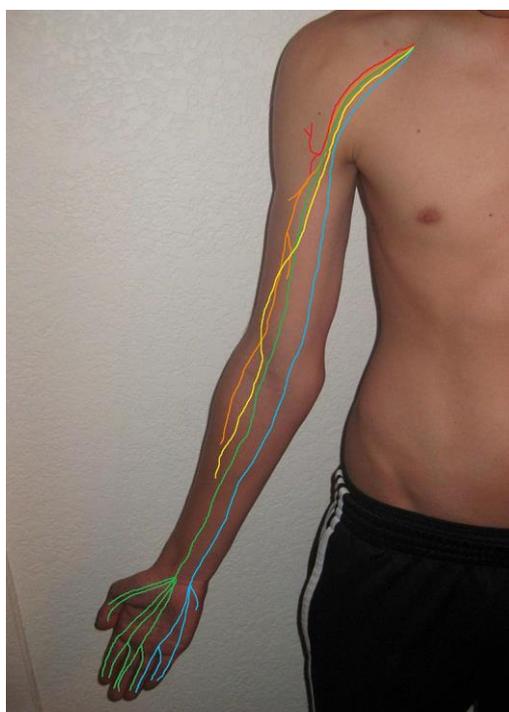


Figura 6: Nervos do plexo braquial no braço humano. Fonte: Wikimedia Commons

O neurocientista A8, que na época da entrevista estava no mestrado, fala de sua pesquisa sobre avaliação sensorial em pacientes que sofreram lesão do plexo braquial e que foram submetidas a reparo cirúrgico. A8 explica que a avaliação sensorial é fundamental para modelar a percepção motora. No contexto do projeto NeuroMat, ele contribui com informações sobre como essas sensações táteis são mapeadas ou compreendidas pelo cérebro, utilizando a coleta de dados a partir de imagem por ressonância magnética funcional (fMRI) e eletroencefalografia (EEG). Por essa razão, a coleta de dados envolve diferentes aspectos fisiológicos e neurológicos relacionados à lesão desse conjunto de nervos.

Percebia-se pela explicação do neurocientista que o plexo braquial possui uma relação anatômica bastante intrincada com outras estruturas do corpo humano, incluindo ossos, articulações, músculos do pescoço e do ombro. É essa relação anatômica que o torna uma estrutura bastante suscetível a diferentes tipos de lesões que podem afetar as inervações da pele que recobrem o músculo dos ombros e o antebraço lateral, polegar e dedo indicador, a inervação do dedo médio e dos quarto e quinto dedos e a pele do antebraço medial, causando principalmente alterações sensoriais e motoras nessas regiões¹⁰¹.

A Figura 7, abaixo, aparece em muitas apresentações dos integrantes do laboratório do INDC. Nela percebe-se as diferentes partes afetadas pela LBP.

101 Os pesquisadores do NPNR ressaltam com frequência em suas apresentações a informação de que os acidentes de motocicletas estão entre as principais causas de lesão do Plexo Braquial. Considerando os altos índices de acidentes de trânsito causado por motocicletas no Brasil, este tipo de lesão vem sendo tratado como um problema de saúde pública, que afeta principalmente homens envolvidos em acidentes automobilísticos – que representa um terço das mortes no trânsito no período analisado. Os pesquisadores mencionam em suas apresentações os dados do Mapa da Violência sobre acidentes de trânsito e motocicletas produzido pela FLACSO, que evidencia o aumento expressivo de 15% nas mortes de motociclistas no Brasil no período de 1996-2011 (WAISELFISZ, 2013).

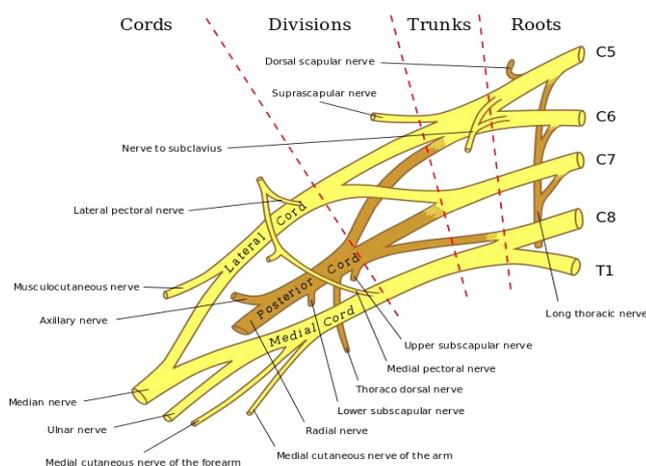


Figura 7: Ilustração anatômica do Plexo Braquial.
 Fonte: Wikipédia

Segundo Lima (2015), as lesões podem ser causadas tanto por traumatismos de baixo impacto como estiramentos do braço e pescoço, que inclui lesões obstétricas (causadas durante o parto) como por acidentes de alto impacto (acidentes de automóvel). A lesão pode ocorrer por avulsão, que é quando os nervos são arrancados de sua raiz na coluna, sendo essa o tipo de lesão mais grave. Outra forma da lesão se dá por ruptura do nervo sem, no entanto, arrancá-lo da raiz. Há ainda o caso em que o acúmulo de tecido nervoso no local da lesão dificulta a passagem de sinais elétricos do membro superior ou também pelo estiramento do nervo em que o nervo é danificado sem que haja o seu rompimento. As lesões podem ser reparadas por meio de técnica cirúrgica de ligação dos nervos rompidos ou também por meio de implantação de enxertos nervosos. Há casos em que o paciente é submetido a cirurgia para religar os nervos que permanecem funcionando, permitindo a recuperação da função de alguns nervos associados aos movimentos e à sensibilidade de partes importantes como controle do ombro e flexão do cotovelo¹⁰².

Por causa da complexidade desses processos, essa dimensão cirúrgica aparece como um aspecto importante para o debate sobre compartilhamento de dados. Para os informantes, a cirurgia realizada nos pacientes que sofrem da lesão é um elemento que insere outros condicionantes para a abertura ou não dos dados. Sobretudo porque, no meio clínico, existem médicos que receiam que a disponibilização desses dados referentes a procedimentos médicos possa trazer implicações para sua carreira em casos envolvendo procedimentos

¹⁰² Este detalhamento acerca da LPB foi fornecido pelos informantes A8 e A10. As informações foram complementadas por materiais recuperados no site da iniciativa ABRAÇO, onde é possível encontrar artigos científicos, vídeos e outras fontes sobre LPB.

cirúrgicos malsucedidos – que são informações inseridas em prontuários médicos – e que podem associar erro médico ao nome do profissional.

Parte do processo de desenvolvimento da estratégia de Ciência Aberta do NeuroMat envolveu o desenvolvimento de instrumentos que garantissem a anonimização dos dados experimentais provenientes de seres humanos, o que demandou um extenso processo de discussão e negociação com médicos e outros profissionais do INDC/UFRJ para a construção dos formulários para a coleta de metadados referentes aos experimentos realizados com os pacientes. Ao falarem desse processo, evidencia-se uma preocupação desses neurocientistas com aspectos éticos e morais implicados na abertura desses dados.

A privacidade e a restrição de acesso a informações sensíveis têm um peso muito grande nos argumentos desse grupo e, muitas vezes, deixa os outros pesquisadores e técnicos responsáveis pelo desenvolvimento das ferramentas sem argumentos. No limite, eles reconheciam que as questões levantadas pelos clínicos deviam ser consideradas mesmo que a equipe responsável soubesse que não se tratava de disponibilizar publicamente todo e qualquer dado e que existe um tratamento antes de serem tornados públicos, mas não bastava apenas eles saberem disso. Como a equipe não tinha instrumentos sólidos e confiáveis para apresentar para esses neurocientistas, a privacidade passou a ser uma demanda válida que a equipe técnica teve que incluir na ferramenta que foi desenvolvida.

Nesse sentido, a dimensão ética se manifesta como uma característica da cultura epistêmica desse grupo envolvido com a prática clínica, diferenciando-o dos experimentalistas.

4.1.2 Neurocientistas experimentalistas

Se no caso dos neurocientistas teóricos a própria natureza dos dados pode facilitar ou, pelo contrário, limitar seu compartilhamento aberto, o envolvimento ou uma implicação pessoal do pesquisador na sua produção parece ser um elemento ainda mais determinante para limitar sua abertura (MILLERAND, 2011).

Os experimentalistas são apontados como os principais focos de resistência ao compartilhamento de dados. O trabalho destes pesquisadores implicado na coleta dos dados é muitas vezes descrito como um trabalho “artesanal” de coleta. Esse trabalho quase artesanal de coleta dos dados – que geralmente envolve ir a campo; trabalho substancial de preparação

para o processo de coleta de dados como preparar cobaias, manter cadernos de pesquisa, observar o experimento, tabular dados – se traduz em um sentimento de posse muito forte do pesquisador sobre os instrumentos da pesquisa e sobre os resultados obtidos.

Os neurocientistas experimentalistas que, na maior parte das vezes aparecem como pesquisadores do ICe/UFRN, trazem elementos associadas a uma necessidade de secretismo dos resultados ou de retenção de determinadas informações como algo esperado no trabalho da experimentação. A necessidade de secretismo emerge muito associada às disputadas por prioridade e reconhecimento que são características do meio acadêmico. Nesse contexto, os esforços despendidos com a experimentação e o tempo que esse trabalho leva para ser desenvolvido aparece como elemento influenciador na forma como os dados são geridos, em outras palavras, como são tornados públicos ou reservados de acordo com a necessidade (HILGARTNER, 1994).

Associado a isso, como a pesquisa desenvolvida por esse grupo é baseada na experimentação comportamental, há um aspecto da natureza do trabalho desses pesquisadores que reforça uma resistência em tornar públicos conjuntos de dados. O tempo e o trabalho implicados nesse tipo de pesquisa está entre os fatores que reforçam o secretismo. Nesses casos, geralmente por motivações ligadas à carreira ou ao reconhecimento desse trabalho, é possível que o neurocientista experimental tenha uma forte resistência a disponibilizar os dados que possam gerar um número de artigos de periódico durante um certo período de tempo.

A seguir, descrevo alguns pontos apresentados por um dos neurocientistas do Instituto do Cérebro de Natal (A4¹⁰³). O nome desse pesquisador é mencionado em diversos momentos por outros informantes, sendo identificado como alguém “*contra a Ciência Aberta*”. Logo no início de nossa conversa, coloquei a afirmação de que ele era visto como uma das principais vozes dissonantes em relação à proposta do centro de pesquisa em utilizar uma abordagem de Ciência Aberta para compartilhamento de dados. Quando mencionei que gostaria que ele relatasse sua opinião sobre o assunto, imediatamente o pesquisador se interpôs como se tentasse se justificar: “*veja, na minha área é um pouco diferente*”. Em seguida, ele passou a se referir às especificidades do tipo de experimento que ele desenvolve em sua área e também apontar para a forma como os dados são produzidos como justificativa para que sua forma de abordar o compartilhamento aberto de dados seja “*diferente*”, isto é,

103 Tentei entrevistar esse pesquisador em vários momentos. Após algumas trocas de e-mail, finalmente, o encontrei no Workshop *Random Structures in the Brain*.

contrária à abertura, como outros informantes haviam relatado. Pedi, então, que ele descrevesse a natureza dos experimentos e como isso se relacionava com a forma como ele disponibilizava os dados.

Naquele momento, A4 ressaltou diversas vezes as propriedades e especificidades dos experimentos realizados por ele e pelo seu grupo de pesquisa como forma de sustentar sua posição em relação à abertura dos dados. Ele fazia referência a um trabalho geralmente longo e “árduo” de produção dos experimentos e de coleta dos dados. Ele mesmo se descrevia como um “experimentalista” e, como um experimentalista, afirmou que não podia simplesmente disponibilizar o conjunto de dados coletados nos experimentos e que tomavam tanto tempo do seu trabalho.

O laboratório do neurocientista entrevistado aplica métodos para estudar as interações neuronais durante a percepção visual, utilizando principalmente a experimentação em primatas. A pesquisa desenvolvida tem por objetivo a compreensão do processamento cortical e visual que possibilite, por exemplo, o desenvolvimento de próteses mais eficazes para deficientes visuais. Ele em específico, investiga “mecanismos de percepção visual a partir do estudo do comportamento de primatas.” A equipe do laboratório de A4 coleta uma série de dados durante determinados períodos de tempo.

Em diversos momentos, ele fez referência ao processo de experimentação com primatas e ao próprio manejo dos animais como um motivo para não abrir os dados. O esforço individual de trabalho para realizar os experimentos e para compor o conjunto de dados experimentais que, em seguida, se tornarão artigos publicados em periódicos, também foi um aspecto que foi bastante enfatizado por ele. Enquanto tecia para mim os detalhes do processo experimental, A4 fazia inclusive demonstrações gestuais e expressões faciais que simulavam cansaço e esse esforço quase braçal implicado nesse processo.

Para ele “*não existe*” a divulgação de partes dos dados de um experimento que ainda não foi concluído para esse tipo de pesquisa.

A narrativa desse neurocientista evidenciava, sobretudo, uma preocupação com a carreira que passa necessariamente por uma retenção dos dados experimentais por um determinado tempo. Ele afirma que a disponibilização dos dados fora desse período de retenção dos resultados é uma forma de “*suicídio acadêmico*”, pois impossibilita que o pesquisador seja reconhecido por meio da publicação de artigos originais:

Eu, como experimentalista, não posso simplesmente pegar e publicar meus dados assim de bandeja (sic). Se eu pego e disponibilizo meus dados de experimento com o macaco, como eu vou publicar artigo? Se eu faço isso eu mato minha carreira. [Entrevista com A4].

Dessa forma, o pesquisador acaba também colocando a abertura dos dados como algo incompatível com a ascensão na carreira e o reconhecimento. Sugeriu assim que os experimentalistas operam com a ideia de que existe uma “vantagem” em termos da aquisição de recompensa ao se manter algum tipo de restrição ou segredo em torno do seu trabalho. Esse “segredo” inclui não apenas os dados, mas os métodos e os equipamentos, assim como o investimento na produção de uma determinada quantidade de artigos a partir dos dados coletados, em oposição a disponibilizar os resultados fora de uma lógica que não seja aquela que recompense a originalidade e esforço despendido pelo pesquisador. Em casos como esses, as práticas de publicação dos cientistas tendem a se concentrar em poucas formas de comunicação pública dos resultados. Ele mesmo comentou que a “*pulverização de publicações não é considerada estratégica*” do ponto de vista do reconhecimento do seu trabalho. Nesse aspecto ele disse que prefere “*focar*” nos periódicos de maior relevância e não divulgar os dados para os “*concorrentes*” antes de publicar.

Um fator adicional que reforça essa predominância da publicação de artigos científicos relaciona-se ao fato de que o instituto de pesquisa do pesquisador acima entrevistado dialoga relativamente com uma comunidade internacional de neurocientistas. Nesse contexto, supõe-se que as disputas de prioridade – que passam pela originalidade, exclusividade e inovação dos resultados com mais chances de serem aceitos pelos periódicos de impacto e, portanto, de terem mais visibilidade – devam ser muito mais intensas.

O neurocientista A2, que frequentemente se manifestava em defesa da Ciência Aberta, também apontou a relação entre a disponibilização dos dados e o reconhecimento acadêmico ao fazer referência a um outro colega neurocientista experimentalista que faz parte do projeto e que tem uma posição contrária. Ele reconhecia como reivindicação legítima a escolha do colega de priorizar a publicação de artigos científicos e manter os dados restritos:

Eu compreendo meu colega X. Quando você começa a fazer um experimento, você já tem hipóteses, perguntas. Você não parte do nada. Você já tem uma noção da originalidade daquilo que você quer fazer. A escolha por não disponibilizar os dados é compreensível. Quando você tem um estudo inovador em mãos, achados inéditos, você quer veicular aquilo numa boa revista [Entrevista com A2].

Invariavelmente, a discussão envolvendo o compartilhamento aberto recai no papel condicionante das avaliações e do sistema de reconhecimento acadêmico no sucesso ou fracasso da adesão dos pesquisadores à Ciência Aberta. A falta de mecanismo para a recompensa e de quadro institucional para as práticas de compartilhamento reforçam a relutância dos pesquisadores a nível individual a aderirem a essas práticas (MILLERAND, 2011). Esse aspecto foi apontado explicitamente por alguns pesquisadores neurocientistas mais jovens.

Os pós-graduandos, por exemplo, trazem nesses momentos questões próprias relacionadas ao reconhecimento do seu trabalho. Durante o I Workshop de Jovens pesquisadores, alunos de doutorado e recém-doutores trouxeram em momentos da apresentação alguns pontos de vista sobre o efeito da disponibilização de seus dados, sejam eles brutos ou tratados, sobre suas carreiras acadêmicas. Para alguns desses jovens pesquisadores, a inversão do processo, isto é, a disponibilização dos dados antes do artigo, é também tratada como um “suicídio acadêmico”. Aqui faço menção à observação de um pesquisador, bolsista de pós-doutorado, de que o sistema de avaliação acadêmica ainda é *“estruturado a partir da quantidade de artigos publicados pelo pesquisador e não pela disponibilização de dados”*. Essa centralidade do artigo científico se traduz em resistência dos pesquisadores a se utilizarem de outros formatos e estratégias de comunicação dos resultados de pesquisa – especialmente se essas outras formas e estratégias não forem avaliados da maneira adequada ou “justa” pelas avaliações acadêmicas e servirem ao propósito do pesquisador de subir na carreira e galgar postos em instituições de C&T.

Um outro aspecto dessa utilização dos dados como recurso para obter crédito está envolvida em um processo de abertura seletiva dos dados, no qual abrir e fechar aparecem como complementares. Em minhas observações, alguns pesquisadores, incluindo neurocientistas que têm uma postura geralmente reticente ao compartilhamento, utilizam os próprios dados em práticas de barganha para ter acesso a uma determinada informação ou ferramenta de pesquisa de outro pesquisador que possa ser útil tanto para a sua própria pesquisa como para atrair algum tipo de benefício para seu grupo.

Assim, não chega a ser contraditório o mesmo pesquisador, que em um momento foi categórico ao se colocar contra a abertura de dados, reiterando o que outros entrevistados já haviam dito de que ele era “contra a Ciência Aberta”, mencionasse a existência de uma rede de cooperação entre as pessoas de sua comunidade em torno do compartilhamento de dados e recursos de pesquisa. Quando apontei para ele que aquilo parecia contraditório, ele pontuou:

Veja, não é que não se compartilha de jeito nenhum. Se um colega que trabalha em coisas parecidas pede algum material, algo que você fez... porque esta é uma área muito pequena, as pessoas conhecem umas às outras (...) então se alguém pede alguma coisa que você fez, você pode passar (sic) porque você confia. Tem essa rede entre as pessoas, uma troca... [Entrevista com A4].

A existência dessa rede de compartilhamento entre pares “confiáveis” fica explícita quando o pesquisador descreveu uma situação em que precisou acionar os pares de sua rede. Isso já havia acontecido, segundo ele, quando pretendia enviar um artigo para um periódico que exigia a utilização de um software de análise específico cuja licença de uso era paga, mas ele não tinha a licença do software. Naquela ocasião, ele recorreu a uma prática comum da comunidade de procurar suas redes informais. No seu caso, a saída foi enviar o conjunto de dados a um colega de uma universidade no Japão que tinha acesso às ferramentas para que ele “rodasse” a análise e pudesse viabilizar a obtenção dos resultados demandados.

A partir desse exemplo, observa-se uma prática de troca dentro da comunidade na qual se imbricam a disputa pela prioridade – evidenciada pelo segredo em torno dos dados experimentais – e a revelação espontânea de uma cooperação sustentada em uma espécie de confiança mútua em uma rede personalizada. Em sua fala, A4 contrastava essa rede de confiança personalizada com a existência de uma desconfiança em partilhar os resultados de pesquisa com “*colegas desconhecidos*” até mesmo em casos em que existe a obrigação de compartilhar.

Ainda trazendo exemplos relacionados à publicação de artigos, o pesquisador fez referência à política dos periódicos científicos de cada vez mais de requererem que o autor disponibilizem os dados de pesquisa quando submetem os artigos. Com uma expressão que sugeria indignação, ele manifestou a existência de uma desconfiança em relação às diretrizes adotadas pelos periódicos acadêmicos. Em sua percepção, existe um risco implicado para o pesquisador ao disponibilizar os dados para editores e pareceristas, que para ele são pessoas “desconhecidas”.

As revistas pedem os nossos dados com frequência, mas isso aí é um problema (...) porque há casos em que você tem uma ideia original, achados originais, mas aí os avaliadores não aceitam (...) mas aí pouco tempo depois você vê algo muito parecido com sua ideia na mesma revista. Os pareceristas têm um papel nisso também [Entrevista com A4].

A seguir descrevo as práticas dos modeladores, buscando compreender como os pontos relacionados ao compartilhamento aberto aparecem entre esses pesquisadores.

4.2 Modeladores e práticas de compartilhamento

Matemáticos, cientistas da computação e físicos, aqui reunidos na categoria “modeladores”, possuem um conjunto de práticas estabelecidas em torno do uso de ferramentas abertas. Muitos modeladores aparecem frequentemente associados a um discurso sobre Ciência Aberta, existindo entre eles uma reiteração recorrente do caráter aberto como uma propriedade intrínseca da ciência.

Algo perceptível entre os modeladores são seus padrões de comunicação científica, em grande medida centrados em publicações de Acesso Aberto. Durante as apresentações das pesquisas percebe-se que muitos referenciam artigos publicados em periódicos de acesso gratuito na Internet. Boa parte dos artigos que tive acesso por meio de notificações do Google e listagem de publicações anexadas a relatórios de atividades, são artigos publicados em periódicos ou repositórios abertos.

Existe um padrão, sobretudo entre os matemáticos, de utilização do repositório de pré-publicações abertas arXiv¹⁰⁴. Delfanti (2016) ressalta que o arXiv não funciona como um periódico, mas sim como uma forma de validação pela comunidade de referência das áreas que utilizam esse repositório. Os matemáticos orientam-se muitas vezes por essa centralidade que o arXiv tem para a área enquanto espaço oficial da comunicação da comunidade matemática. Assim, mesmo que publiquem posteriormente a versão final do artigo disponibilizado no arXiv em um periódico (de acesso pago ou não), os matemáticos costumam publicar primeiramente nesse repositório de revisão aberta em primeiro lugar.

104 O arXiv é um repositório de pré-publicações ou “*pre-prints*”. É muito utilizado na física e na matemática, sendo, inclusive o meio de divulgação preferencial de algumas subdisciplinas. O arXiv utiliza um modelo de revisão ampla e aberta, no qual os autores submetem rascunho de um artigo para avaliação de toda a comunidade antes de enviá-lo para um periódico.

Outros periódicos que aparecem com frequência nas referências em artigos e *slides* de modeladores estão periódicos da PLoS e o Frontiers¹⁰⁵, ambas editoras de periódicos de acesso aberto da área de ciências da vida e ciências biológicas¹⁰⁶. A17 justificou a predominância desses periódicos porque “*Publicam muita coisa de biologia quantitativa*”.

A olharmos para as práticas individuais do grupo de modeladores, os matemáticos possuem padrões mais homogêneos e, ao mesmo tempo, certa dinamicidade e rapidez em termos de tempo na forma como desenvolvem o trabalho e como se utilizam dos canais de comunicação de comunicação científica, se comparados aos neurocientistas.

Durante as atividades, os matemáticos passam a maior parte do tempo escrevendo e interagindo em torno das demonstrações numéricas. Por essa razão, é comum ver equações escritas nos muitos rascunhos espalhados em mesas ou afixados nos murais e nas várias lousas existentes em praticamente todas as salas do centro de pesquisa.

A lousa é um instrumento integrado ao trabalho dos matemáticos. Durante as apresentações das pesquisas, a maior parte deles utiliza a lousa associada à apresentação de *slides*, uma vez que não basta apenas falar sobre o teorema, frequentemente é preciso demonstrá-lo. Era também bastante recorrente nos intervalos das reuniões e eventos observar dois matemáticos iniciando uma conversa no café e, logo em seguida, vê-los diante de uma lousa: enquanto um escreve uma longa demonstração em uma lousa, outro observa atentamente ou toma notas.

Em uma reunião realizada em junho de 2017 com a equipe de difusão, comentei, com um pouco de estranhamento, que havia percebido que alguns alunos que conheci em 2015 quando ainda estavam no mestrado, já estavam terminando ou já haviam terminado o doutorado. Comentei ainda que a quantidade de fotos dos integrantes da equipe, afixadas em um totem na recepção, aumentava cada vez que eu retornava ao NeuroMat. As caras que eu via pelas instalações mudavam muito rapidamente. Sobre isso, B1 proveu algumas informações que permitem compreender um pouco das características do grupo de modeladores matemáticos. Ele explicou que a carreira dos matemáticos costuma ser “*meteórica*”. Como não existem muitas pessoas se dedicando àquele tipo de pesquisa e área, tudo acaba sendo muito rápido. O pesquisador encontra o teorema e sua pesquisa terminou.

105 FRONTIERS. **Home page**. Disponível em: <<https://www.frontiersin.org/>> Acesso em 20 mar. 2018.

106 Os neurocientistas pertencentes ao grupo possuem um padrão de publicação com relação ao acesso aberto menos evidente. Apesar de mencionarem com frequência que os critérios de seleção dos locais de publicação são muito mais difusos e depende de cada caso, os critérios que costumam pesar geralmente é a visibilidade do periódico dentro da área de especialidade e a originalidade do estudo.

As teses são relativamente curtas em comparação às teses das ciências humanas. Os matemáticos não costumam ficar sem bolsa ou emprego por muito tempo já que tudo é muito rápido.

Inclusive, dois alunos que foram fazer doutorado sanduíche na França já tinham retornando e prestariam um concurso para docente no IME logo em seguida. Outro aluno, dentre aqueles que finalizaram o doutorado, havia passado em um concurso para docente em uma universidade do Nordeste. Outro pós-doc havia finalizado a bolsa e conseguido uma outra bolsa de pós-doutorado nos Estados Unidos. O projeto também havia recebido muitos bolsistas de pós-doutorado estrangeiros e, dali a pouco, chegariam outros mais, já que ainda havia bolsas disponíveis.

Durante essa reunião, também compartilhei a impressão de que os modeladores demonstravam uma maior adesão aos meios de publicação em acesso aberto, e portanto, pareciam mais abertos à Ciência Aberta. B1 chamou atenção para que eu me atentasse também às “*relações de poder*” na determinação da relação dos pesquisadores com a Ciência Aberta. Ele então apontou a existência de uma forte hierarquização na matemática entre pesquisadores júniores e pesquisadores seniores.

Na matemática, o orientador possui uma grande centralidade nas escolhas do jovem pesquisador. Alguns informantes apontaram a existência de uma grande hierarquização na relação entre os pesquisadores dessa área. Acerca disso, B1 disse que, devido à sua convivência com os matemáticos, vinha percebendo que os “*o pesquisador júnior não tem poder de escolha. Não é uma discussão. Não se discute muito o que o orientador pede para fazer*”. Por outro lado, na neurociência, comparativamente à matemática, o pesquisador júnior parecia ter muito mais “*margem de manobra*” em suas escolhas de como e onde publicara. Comparativamente aos matemáticos, o pesquisador neurocientista exerce muitos papéis e muitas vezes faz o trabalho que ninguém quer fazer, como preparar as cobaias, acompanhar o experimento, fazer as anotações, tabular os dados, acompanhar o experimento em horários “*ingratos*”.

Essas diferenças nas relações estabelecidas ao longo da cadeia entre os pesquisadores de acordo com os níveis que ocupam se manifestam na forma como matemáticos e neurocientistas se relacionam com os dados. Na ciência experimental, a manipulação de dados é uma responsabilidade geralmente delegada aos técnicos ou estudantes. De acordo com Roth e Bowen (2001), esse trabalho de “*pôr a mão*” nos dados

constitui uma etapa fundamental do treinamento, visto que possibilita contato do jovem pesquisador com processos tangíveis e, muitas vezes, restritivos da pesquisa. Em outros casos, a tarefa técnica, laboriosa e repetitiva demanda trabalho técnico com um certo *know-how*, o que torna o apoio do pessoal técnico especializado fundamental. Nesse sentido, o jovem neurocientista costuma ter muito mais poder de negociação sobre onde e como ele quer divulgar os resultados do seu trabalho. Esse aluno também costuma ser o primeiro autor, e isso é algo que costuma ser negociado com o orientador.

A comunidade de neurociência computacional se assenta na construção de recursos compartilhados, sobretudo para o compartilhamento do código dos modelos e modelos de validação. Gardner et al (2008) argumentam que o surgimento da disciplina neurociência computacional ou neuroinformática está vinculado às infraestruturas de compartilhamento de dados. Embora os bancos de dados individuais, principalmente para atlas cerebrais, já existissem no final da década de 1980, a neurociência computacional surge oficialmente como disciplina junto com os grandes projetos governamentais de pesquisa sobre o cérebro, com o financiamento das agências americanas e europeias que passaram a aceitar projetos de pesquisa que combinassem a pesquisa em neurociência com a informática.

Nesse sentido, durante entrevista com A3, ele relatou que o compartilhamento de dados não é uma rotina da neurociência, mas o ramo da neurociência computacional ou neuroinformática – disciplina que ele considera ser bastante jovem – possui uma tradição em desenvolver e disponibilizar ferramentas de pesquisa abertas, como softwares e bancos de dados/repositórios de acesso aberto. O pesquisador ressaltou que *“existe uma grande preocupação da área em acessar as informações padronizadas, assim como a necessidade de definição de terminologias comuns, necessidade de união de esforços”*.

Adicionalmente, os softwares fazem parte do ferramental básico da neurociência computacional e muitas das práticas desses pesquisadores envolvem o desenvolvimento e aplicação dessas ferramentas. Muitas das discussões sobre a abertura nesse campo também giram em torno da disponibilização dessas ferramentas em código aberto.

De acordo com o entrevistado, *“disponibilizar essas ferramentas para a neurociência computacional não é muito uma questão, geralmente as pessoas compartilham”*. Muitas das ferramentas existentes na neurociência computacional foram desenvolvidas pela comunidade ao longo do tempo e são de uso relativamente generalizado entre os pesquisadores desse campo.

As ferramentas aparecem em diversos momentos ao longo das apresentações e no dia a dia das pesquisas. Algumas delas foram o *OpenWorm*, que foi objeto de debate em uma reunião do NeuroMat¹⁰⁷; a “*Scholarpedia*”, que é uma enciclopédia de artigos acadêmicos curtos, escritos por especialistas e revisados por pares; o repositório *Neuroscience Information Framework*¹⁰⁸, que agrega dezenas de outras bases de pesquisa em neurociências; e CARMEN¹⁰⁹, projeto de análise de modelagem em neurociências; o NeuroMorpho.org¹¹⁰, repositório de imagens de neurônios reconstruídos digitalmente.

Muitos dos projetos originaram bancos de dados abertos e softwares de pesquisa de código aberto – abrangendo subcampos de biologia molecular através de morfologia e fisiologia celular, a circuitos neuronais, atlas cerebrais, exames cerebrais, clínicos, neurologia e comportamento, principalmente com a funcionalidade de visualização dos modelos. A maior parte das ferramentas mencionadas são disponibilizadas em código aberto, conforme os próprios pesquisadores mencionam e como é destacado nos próprios sítios *web*.

A utilização de bases de dados públicas de acesso aberto também está fortemente presente no trabalho de modelagem neuronal. As análises estatísticas e modelos fazem grande uso de repositórios de dados de pesquisa públicos. Isso pode ser observado durante as apresentações e nos artigos científicos produzidos pelo grupo, nos quais são frequentemente citadas as bases em que foram coletados os dados utilizados ou em que foram depositados os conjuntos de dados.

A medida que fui tentando compreender o processo de modelagem foi possível observar não apenas a forma como os modeladores acessam e processam os dados experimentais, como também averiguar, na prática, que esse trabalho, ao menos no contexto analisado, envolve constantemente a mobilização de diferentes recursos de pesquisa abertos.

Em posse dos modeladores, os dados adquirem diferentes propriedades, formas e significados. Como os dados são transformados em números em tabela, são portados nos computadores e levados para as reuniões, assim, é possível observar as interações em torno da elaboração das representações gráficas dos modelos. Um exemplo é a técnica de

107 OPENWORM. Home page. Disponível em: <<http://openworm.org/>>. Acesso em 20 mar. 2018.

108 NEUROSCIENCE INFORMATION FRAMEWORK. Home page. Disponível em <<https://neuinfo.org/>>. Acesso em 20 mar. 2018.

109 CODE ANALYSIS, REPOSITORY & MODELLIN FOR E-NEUROCIENCE. Home page. Disponível em: <<http://www.carmen.org.uk/>>. Acesso em 20 mar. 2018.

110 NEUROMORPHO.ORG. Home page. Disponível em: <<http://neuromorpho.org/byRandom.jsp>>. Acesso em 20 mar. 2018.

“classificação de disparos neurais” (*spike sorting*), utilizada para representar a ação de transmissão elétrica dos neurônios.

O modelador A12, especializado na técnica de classificação dos disparos neuronais, me explicou como tratava os dados na produção dos modelos de disparos neurais¹¹¹.

As classificações convencionais de disparos elétricos geralmente são dependentes da medição de uma variedade de neurônios simultaneamente, o que gera riscos de erro na interpretação e compreensão. Segundo ele, o estudo de células isoladamente oferece uma visão muito limitada do quadro inteiro. Nesse aspecto, a compreensão do cérebro depende em grande medida da capacidade de registrar simultaneamente grandes populações de células. Fica mais difícil, por exemplo, saber se o registro coletado é efetivamente o registro daquilo que se esperava coletar e não de ruídos ou ecos da amplificação de atividades neuronais já registradas ou irrelevantes, sendo mais difícil também discriminar a atividade de cada neurônio na massa de dados (Figura 8).

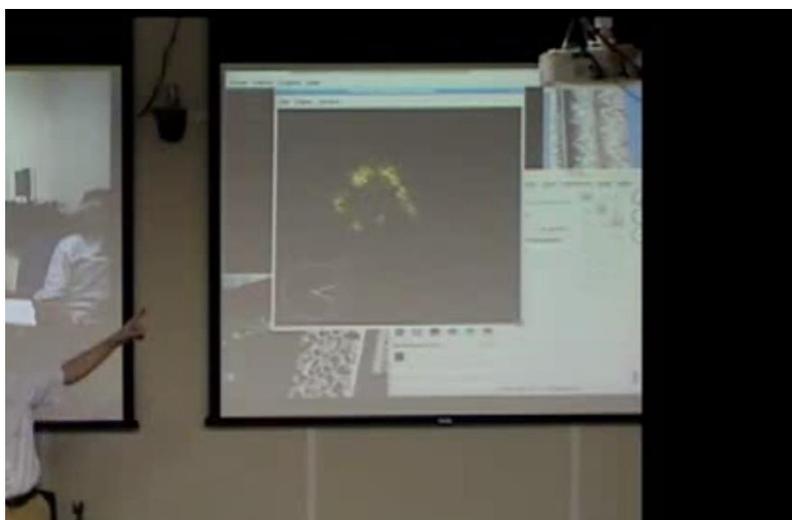


Figura 8: Modelador mostra as atividades elétricas de um grupo de neurônios. Fonte: IPTV USP

¹¹¹ Para compor a descrição do processo de classificação de disparos neurais me vali de informações providas tanto na entrevista como em materiais produzidos pelo NeuroMat, como o material do curso de disparos neurais oferecido 2014 (F19) e vídeo explicativo produzido pelo centro para o verbete “Disparos Neurais” na Wikipédia (F17).

Nesse momento, o pesquisador mostra dados captados por um eletrodo: um eletrodo mostra amplitude mais alta, representada por picos na curva, e outra de amplitude mais fraca, representada por declives na curva. Ele explicou que a visualização daqueles dados era feita por meio de um software de código aberto que possibilita uma representação de conjuntos de dados com mais de duas dimensões. Em seguida o pesquisador mostra uma animação em 3D que representa completamente o conjunto de dados. A imagem é formada por uma espécie de nuvem de pontos coloridos, na qual cada um desses pontos coloridos é associado aos neurônios.

A12 disse que costumava utilizar uma ferramenta de software *open source* que permite a exploração visual de dados de alta dimensão e a observação em diferentes ângulos, além da manipulação do modelo, como alguém que observa uma pedra preciosa por diferentes ângulos. O software utilizado chamava-se Ggopi, um software baseado em linguagem de programação R ¹¹², utilizado principalmente por cientistas de dados e estatísticos.

Após obter aquele resultado, a tarefa passa a ser a descoberta da quantidade de nuvens e o centro de cada nuvem, e para cada disparo ou cada ponto, o centro mais próximo. Após observar a rotação do conjunto de dados na figura em 3D (o pesquisador interage com uma imagem em 3D na tela), ele afirma que identificava algumas “nuvens” presentes. Em seguida, aplica-se um algoritmo de mineração de dados para fazer o agrupamento em *clusters* em milissegundos (*k-means*¹¹³). A figura 9 mostra os *clusters* em 3D. Cada cor representa a atividade de um neurônio.

112 R é uma linguagem de programação utilizada para cálculos estatísticos e elaboração de gráficos.

113 O *k-means* é um método de mineração de dados que produz agrupamentos de dados de maneira automática de acordo com o grau de semelhança (clustering).

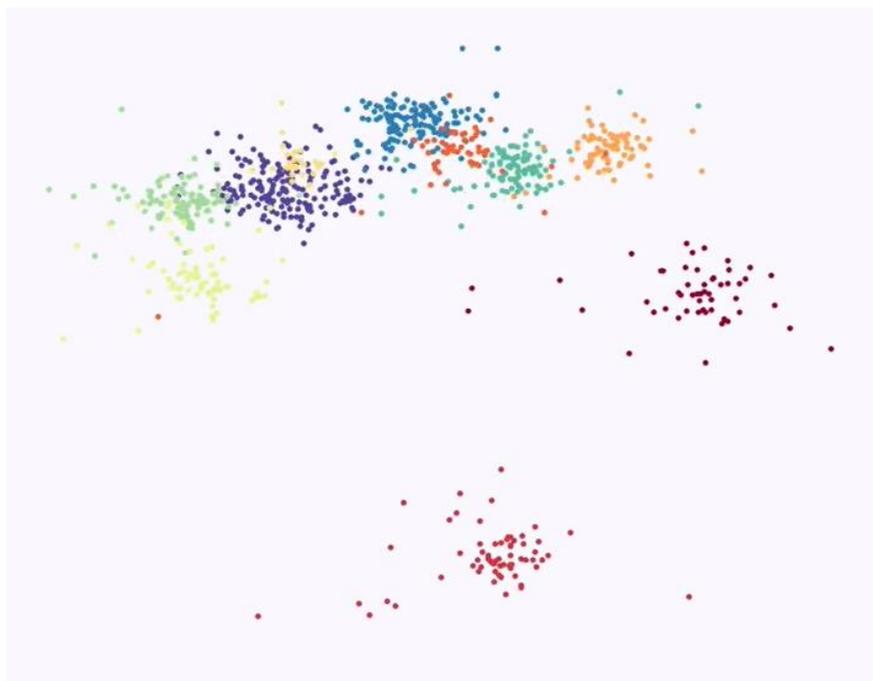


Figura 7: Visualização em 3D dos disparos neurais. Fonte: NeuroMat.

Fonte: Pouzat (2014)

O pesquisador volta para a exibição em 3D e, com a ajuda de outro software, vai colorindo todos os pontos utilizando cores correspondentes ao seu conjunto de dados enquanto avalia a qualidade da visualização do modelo. Ele então chama atenção para pontos com a mesma coloração permanecem juntos. Em alguns casos, a classificação de disparos termina nesta fase de produção do modelo, mas o problema é que para produzir aquele modelo se utilizou somente dados de amplitude do neurônio, o que traz algumas limitações interpretativas. Os dados de amplitude não funcionam para todos os conjuntos de dados, segundo ele *“porque pode acontecer da atividade de um neurônio interferir ou gerar ruídos na atividade de outro neurônio”*.

Em um dos cursos oferecidos no NeuroMat¹¹⁴, ele exemplificou essa situação mostrando na tela do computador o funcionamento de três neurônios no qual um deles gera atividade nas regiões dos outros dois, o que significa que se forem utilizados somente os dados de amplitude não seria possível identificar de qual neurônio se origina a maioria dos disparos. Um contexto como este descrito pelo modelador demandava mais informações do que somente os dados de amplitude. Como os neurônios disparam com tempos bastante

114 Cf. F19

específicos, idealmente, o modelador deveria incluir essas informações pois é esse nível de detalhamento que torna os modelos confiáveis. Entretanto, a inclusão dessas informações requer o desenvolvimento de modelos estocásticos de descarga neural precisos e que levem a algoritmos eficientes.

Os dados coletados em experimentos são traduzidos em inscrições na forma de algarismos em planilha, que após passar pelo processo de manipulação por meio da aplicação de ferramentas matemáticas e computacionais são transformados em representações objetivas – gráficos de barras, gráficos com linhas com picos ou declives, grafos, modelos tridimensionais, “*scatters*” com pontos coloridos ou diagramas com cores – em torno das quais se estabelecem discussões, inferências e interpretações. Durante os workshops, é possível ver os modeladores aplicando ou discutindo formas de olhar “matematicamente” para o conjunto de dados. As reuniões e eventos científicos parecem extensões do laboratório. Durante as apresentações, muitos permanecem executando tarefas como preparar ou fazer ajustes em *slides*. Percebia-os com frequência manipulando planilhas e softwares especializados para analisar o conjunto de dados e gerar diferentes tipos de visualizações. Estavam constantemente consultando bases bibliográficas, periódicos e baixando artigos, digitando linhas de código ou mesmo rascunhando demonstrações matemáticas em folhas.

Nos momentos de interação entre neurocientistas e modeladores é possível identificar as fronteiras entre essas culturas. Notava-se que os dados eram frequentemente objeto de barganha para os modeladores também nesses contextos. Isso pode ser observado em um momento durante um workshop, quando observei um dos modeladores mostrar interesse na técnica de análise de dados apresentada por um pesquisador estrangeiro, cuja abordagem poderia ser complementar para a análise dos dados coletados a partir de um dos softwares de coleta de dados desenvolvidos pelo centro. O pesquisador disse que colocaria à disposição do colega “*todos os nossos dados*”, propondo também que eles intercambiassem alunos de pós-graduação, inclusive afirmando que o seu doutorando “*levaria*” os dados para o laboratório do colega nos EUA. Os modeladores também estão a todo momento expondo aos colegas as representações dos modelos, inclusive solicitando ajuda na compreensão dessas representações feitas a partir da aplicação de ferramentas de análise aos dados.

Durante o Workshop “*Random Structures in the Brain*”, A15 apresentou resultados da pesquisa que vinha desenvolvendo sobre descrição matemática da evolução temporal de sistemas de neurônios. Enquanto examinava com expressão de dúvida a representação gráfica de uma cabeça com regiões com cores diferentes, representando reações

dos neurônios em pontos específicos, ela comentou que não conseguia inferir ou sugerir uma interpretação para aquelas visualizações. Ela passa um bom tempo antes da apresentação fazendo ajustes e checagem no conjunto de dados. Ao longo desse processo, a pesquisadora alternava entre uma planilha na qual os dados apareciam tabulados e o MatLab¹¹⁵. Ela também mostrava o gráfico a alguns colegas e, em determinado momento, fez o seguinte comentário: *“olha aqui o meu gráfico, não ficou bonito? Agora só falta achar um significado para ele...”*. A certa altura do workshop, a neurocientista que participou da coleta dos dados de EEG utilizados por A15 sentou ao seu lado e ambas discutiram sobre interpretações possíveis de serem feitas a partir daquelas representações que a modeladora produziu, no intuito de ajudar no entendimento e interpretação das representações.

Em momentos como esses, os modeladores demonstram certo “desapego” em relação aos produtos do seu trabalho, quando expõem para várias pessoas, e não “escondem”, o trabalho, revelando de certa maneira um aspecto colaborativo da pesquisa, em muitas vezes dependente da certificação dos colegas. Essa maneira desapegada se reflete também na forma como normalmente tentam convencer os colegas neurocientistas a disponibilizarem os dados.

115 MatLab é uma aplicação utilizada para cálculos numéricos e visualização de gráficos, frequentemente utilizada por matemáticos e estatísticos.

Quando fiz esse comentário de que “*os modeladores demonstram maior desapego aos dados e são mais abertos ao compartilhamento aberto dos resultados e ferramentas*”, o modelador A12, que é matemático atuando em classificação de disparos neurais reagiu com a afirmação de que é simples de compreender o porquê dos matemáticos normalmente defenderem o compartilhamento de dados:

Os matemáticos defendem o compartilhamento de dados porque, afinal, nós não produzimos dados, os dados não são seus. O modelador está preocupado com a aplicação do teorema, com as ferramentas, em gerar modelos confiáveis. Mas ele tem que se preocupar com os protocolos experimentais, qualidade dos dados que tem em mãos, na integridade do conjunto de dados. [Entrevista com A12].

Nas ciências exatas, o trabalho de preparação dos resultados da pesquisa, que é um processo anterior à sua disponibilização, geralmente não exige um esforço idêntico ou levantará as mesmas questões em comparação à preparação dos resultados nas ciências da saúde e nas ciências sociais (MILLERAND, 2011). Nesse sentido, a abertura dos resultados na matemática não aparece como um grande problema, visto que, muitas vezes, o que está em jogo para os pesquisadores dessas áreas é a aplicação do conhecimento matemático para resolver um problema.

Mas mesmo essas características não são absolutas. É possível observar modulações desse comportamento entre os modeladores que trabalham diretamente com o desenvolvimento de ferramentas de software.

Outro aspecto que as ferramentas abertas permitem observar diz respeito ao seu significado enquanto produto da atividade científica. É interessante olharmos para as características particulares desse tipo de produção acadêmica, especialmente do software de código aberto. Wiley e Michaels (2004) observam que as ferramentas de software tornaram-se essenciais para a biologia moderna em razão da centralidade da análise de dados em disciplinas como bioinformática e biologia molecular e cada vez mais fazem parte de outras disciplinas em diferentes campos.

As peculiaridades do desenvolvimento do software científico traz questões em torno da atribuição e do status do desenvolvimento de produtos de código aberto como contribuição intelectual que se tornam ainda mais importantes no contexto de políticas de Ciência Aberta. Algumas das preocupações que permeiam contextos acadêmicos nos quais os pesquisadores desenvolvem ferramentas e serviços para a comunidade científica, como softwares de pesquisa de código aberto, são perpassadas por questões sobre a atribuição de

crédito e autoria e mesmo por preocupações relativas às aplicações que serão feitas do software.

Levin e Leonelli (2016) afirmam que, para alguns pesquisadores, o tempo e esforço exigido pelo desenvolvimento de software é visto como um imperativo para compartilhamento do código e de toda a documentação do software, assim como a colaboração que se estabelece em torno da disponibilização aberta do código fonte é considerada como um dos aspectos mais importantes e valiosos de seu desenvolvimento. Por outro lado, alguns pesquisadores tendem a tratar o trabalho de desenvolvimento de ferramentas e serviços como uma forma inferior ou menos valiosa do trabalho acadêmico (LEVEN; LEONELLI, 2016).

Quando o desenvolvimento de software está atrelado a uma pesquisa, a decisão por reter o código-fonte relaciona-se a fatores diversos. Uma das práticas existentes é liberar o código fonte quando o projeto de pesquisa estiver concluído ou quando as publicações associadas já estiverem sido feitas, evitando-se revelar aos “concorrentes” o que o pesquisador está trabalhando; ou disponibilizar quando o laboratório não tiver mais recursos para a manutenção do software. Muitas das iniciativas de Ciência Aberta pretendem inverter essa lógica colocando a abertura dos processos, o que inclui as ferramentas, como algo que deve ocorrer em tempo real, no momento em que os resultados da pesquisa são produzidos. Segundo as autoras, experiências negativas ou frustrações de pesquisadores que criam ou disponibilizam ferramentas para a sua comunidade com questões de propriedade intelectual ou recompensa se traduziram em comportamentos reticentes quanto à disponibilização em código aberto.

Vejamos com essa questão foi apontada por um dos pesquisadores do NeuroMat. O pesquisador A9, que é da área de Ciência da Computação e vinculado ao CCSL, se coloca como parte de um “ramo” de Ciência Aberta que faz software livre. O pesquisador oferece uma perspectiva de como os colegas encaram o compartilhamento das ferramentas de software. Segundo ele:

Não sei afirmar como é em outras áreas, o que posso falar é mais da área de desenvolvimento de software. Aqui no IME existem três tipos de pesquisadores. Tem o colega que desenvolve software individualmente ou junto com um orientando e não disponibiliza de jeito nenhum... Ele tem mais essa lógica de que ‘foi eu que fiz, eu que desenvolvi e não quero que ninguém se beneficie’. Tem aquele outro que pensa ‘não vou disponibilizar porque meu código não está bom’. Esse no máximo ele coloca um *link* no artigo para um repositório, uma nota de rodapé, ele não acha que é algo que mereça ser disponibilizado. Tem outros aqui do IME que não disponibilizam porque têm preguiça (sic). No máximo eles disponibilizam em um arquivo *zip* e colocam o *link* do site pessoal no artigo, sem nenhuma documentação. [Entrevista com A9].

Ele disse ainda que é comum ver a frustração do pesquisador que sente que sua contribuição intelectual e sua mão de obra não receberam o crédito adequado – como inclusão como coautor de artigos resultantes de pesquisa que utilizaram o software que ele desenvolveu para analisar dados e gerar resultados.

Situações como essas relacionadas à falta de reconhecimento adequado para as contribuições no desenvolvimento de software se traduzem, por exemplo, na disponibilização das ferramentas em formato de “arquivo binário” (LEVIN et al, 2016), que ao contrário do código aberto, é um tipo de arquivo de programa disponibilizado publicamente apenas para utilização, o que significa que os usuários do software podem utilizar os recursos de análise de dados, porém não são autorizados a acessar o código fonte para fazer modificações. Segundo Leonelli, a disponibilização em formato binário pode ser uma forma de expor o trabalho, mas também é um modo de assegurar que não será revelado aos concorrentes no que o pesquisador está trabalhando.

Outro aspecto observado entre os modeladores diz respeito à colaboração desses pesquisadores com os neurocientistas por meio do compartilhamento dos resultados de pesquisa como instrumento necessário à produção de modelos, o que justifica, em diversos momentos, uma certa postura apologética dos modeladores em relação ao compartilhamento dos dados.

Algo reiterado por outros modeladores foi a associação do compartilhamento de dados com uma noção de qualidade e reprodutibilidade experimental. Essa preocupação surgiu em alguns eventos quando foram apresentadas as dificuldades relacionadas ao acesso a bases de dados consistentes e padronizadas em neurociências. Isso ocorre porque, para entender os fenômenos do sistema nervoso, a ciência do cérebro precisa analisar vários dados de experimentos e disciplinas em conjunto e que são coletados muitas vezes usando técnicas diferentes, e por essa razão os modeladores associam fortemente a Ciência Aberta com

padrões de compartilhamento. Nesse aspecto, o compartilhamento de protocolos de pesquisa apareceu como uma questão colocada frequentemente nas interações com modeladores.

A interação dos modeladores com os neurocientistas, por meio do compartilhamento dos resultados de pesquisa, é parte fundamental na produção de modelos, o que justifica, em vários momentos, uma postura frequentemente apologética dos modeladores em relação ao compartilhamento dos dados a partir de uma abordagem de Ciência Aberta. Como os modeladores demonstram, a produção de modelos neuronais é um trabalho essencialmente multidisciplinar que necessita de colaboração. Desse modo, o desenvolvimento de ferramentas computacionais que viabilizem essa colaboração aparece como um elemento de interesse dos modeladores e central para as práticas estabelecidas no espaço do NeuroMat.

Considerações Finais

Em torno das dinâmicas que se estabelecem em torno das infraestruturas de Ciência Aberta é possível identificar características e propriedades específicas de cada comunidade epistêmica envolvida com o seu desenvolvimento e utilização, possibilitando, assim, observar os contornos das fronteiras dessas áreas e como se comportam em relação ao compartilhamento aberto dos resultados de pesquisa.

No decorrer do capítulo, busquei evidenciar que as formas como os pesquisadores se apropriam de ferramentas abertas e principalmente como se relacionam com a disponibilização dos resultados do seu trabalho são relativamente diversificadas e relacionadas a aspectos contextuais diversos. A interação com modeladores e neurocientistas permitiu elaborar um panorama acerca das diferenças entre/através dos domínios científicos e os limites entre culturas epistêmicas em interação no espaço do NeuroMat.

Como parte do estudo etnográfico, o próximo capítulo foca nos aspectos da infraestrutura de compartilhamento desenvolvida pela equipe do projeto, especialmente nas dinâmicas e interações em torno do software de gestão e compartilhamento de dados experimentais.

5 DINÂMICAS EM TORNO DE UM SOFTWARE DE GERENCIAMENTO E COMPARTILHAMENTO DE DADOS

Trago para este capítulo análises sustentadas na etnografia do desenvolvimento e implementação de uma ferramenta de código aberto para gestão e compartilhamento de dados experimentais, o software *Neuroscience Experiments System* (NES). A ferramenta é central na estratégia de Ciência Aberta do centro de pesquisa, inserindo-se em uma discussão mais ampla sobre abertura de dados e construção de uma infraestrutura de uso comum entre modeladores e neurocientistas. Ao focar a análise na ferramenta, o objetivo do capítulo é discutir diferentes dinâmicas e interações que ocorrem em torno e através dessa ferramenta relacionada à estratégia de Ciência Aberta, identificando o trabalho envolvido, os papéis e significados atribuídos a esse desenvolvimento.

Busco aqui descrever os processos envolvidos no desenvolvimento do software, destacando especificamente o papel dos padrões e os significados desses processos. Em seguida, discute os dilemas que permeiam o software de código aberto no contexto das atividades do eixo de transferência de tecnologia.

5.1 Neuroscience Experiments System

Entre os objetivos do NeuroMat está o desenvolvimento de ferramentas computacionais para a coleta, gestão e disponibilização de dados experimentais dos laboratórios colaboradores do centro. Os softwares e banco de dados para armazenamento de dados experimentais são parte da estratégia de Ciência Aberta já mencionada. Quando iniciei a pesquisa de campo em 2015, havia poucos meses que a equipe havia dado início à primeira fase do projeto de construção de um banco de dados neurofisiológicos de acesso aberto. O objetivo do projeto era disponibilizar publicamente em acesso aberto dados, metadados e processos de análise dos dados.

Naquele período, foi realizada a implementação de um protótipo do projeto, em modo experimental, no Núcleo de Pesquisa em Neurociências e Reabilitação do Instituto de Neurologia Deolindo Couto (INDC), e posteriormente seria implementado nos outros laboratórios colaboradores. O protótipo tratava-se de um banco de dados para registro, armazenamento e disponibilização dos dados produzidos nas pesquisas sobre plasticidade do cérebro em caso de lesão e reconstrução do Plexo Braquial, que vinham sendo desenvolvidas pelos pesquisadores do NPNR. O banco de dados incluiria também histórico médico, registro das lesões e evolução clínica dos pacientes. Como parte desse projeto foi desenvolvido um software para gestão e compartilhamento de dados experimentais, chamado de Neuroscience Experiments System (NES)¹¹⁶. No repositório GitHub, o NES é assim conceituado¹¹⁷:

NES (Neuroscience Experiments System) é uma ferramenta de código aberto que visa auxiliar os laboratórios de pesquisa de neurociência em procedimentos de rotina para coleta de dados. O NES suporta a reprodutibilidade de experimentos, permite a comparação de dados em todos os estudos e mantém a proveniência dos dados. Além disso, promove formatos padronizados para experiências e relatórios de análises.

No material de divulgação, o NES é descrito como uma ferramenta que auxilia os pesquisadores no processo de coleta de dados e metadados durante o desenvolvimento dos experimentos e que possibilita a obtenção de um registro integrado de diferentes tipologias de dados – clínicos, eletrofisiológicos, imagem e comportamentais produzidos por um laboratório, grupo ou projeto de pesquisa – por meio de uma plataforma *web* para registro de dados e metadados de cada etapa de um experimento em neurociência. Toda a atividade de registro no NES é enviada para um repositório central mantido pelo NeuroMat para ser disponibilizado publicamente¹¹⁸.

A versão 0.1 do software foi desenvolvida com base nos experimentos realizados por pesquisadores do Laboratório de Neurociência e Reabilitação do INDC/UFRJ. A implementação ocorreu sob a coordenação das pesquisadoras A2 e A13 a partir de novembro de 2014, em estágio experimental, visando ao gerenciamento e compartilhamento dos dados neurofisiológicos coletados por esse grupo. Em 2015, foi lançada a versão 0.2.1, que em relação à versão anterior (0.1), teve duas contribuições principais. Uma delas incluiu a

116 No período em que realizei a pesquisa de campo, além do NES, estava em planejamento um outro módulo para gerenciar dados coletados em exames de neuroimagem (como imagens de ressonância magnética) e um módulo para gerenciar dados obtidos em experimentação animal. Um dos desenvolvedores relatou que esses módulos seriam implementados de acordo com as demandas apresentadas pelos laboratórios colaboradores.

117 Cf. F23

118 Cf. F24

implementação de um recurso para vincular os dados experimentais registrados por um projeto de pesquisa. A outra contribuição incluiu a representação de todas as definições das condições experimentais a que os indivíduos seriam submetidos, isto é, os protocolos experimentais. Entre as funcionalidades da versão, foi implementado o registro de pacientes e gerenciamento de experimentos a partir de questionários eletrônicos. A figura 10, a seguir, mostra um dos responsáveis pela ferramenta apresentando a estrutura do repositório durante evento sobre Ciência Aberta.

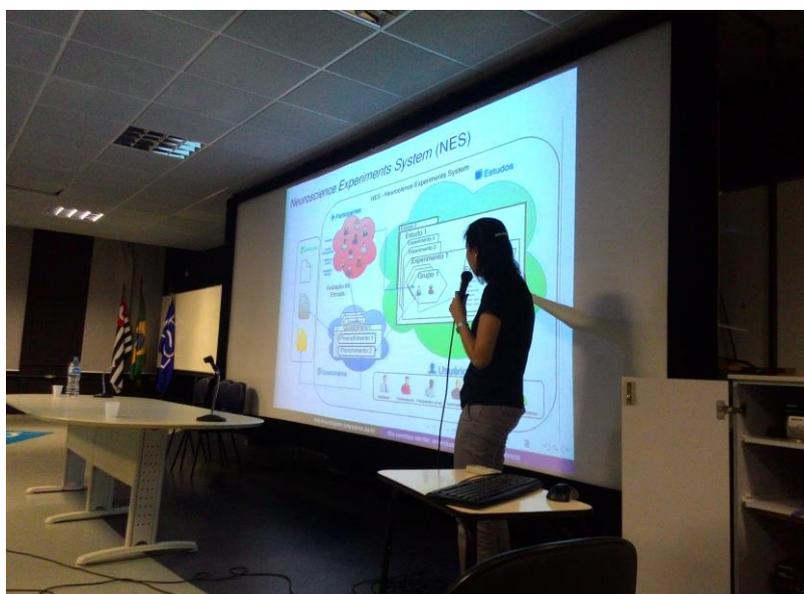


Figura 8: Apresentação do banco de dados durante o II Encontro de Ciência Aberta. Fonte: Arquivo do autor.

Na primeira reunião com alguns membros do grupo de trabalho (A13 e B2) a centralidade do NES era perceptível. Naquele momento, os informantes pareciam muito empolgados com os resultados do projeto, em particular com o software que havia sido desenvolvido pela equipe. O software era uma das principais etapas de criação do banco de dados. Inclusive havia causado certa repercussão positiva na área. A maior representação daquela empolgação ocorreu quando o coordenador do centro entrou na sala no meio de nossa conversa e pediu para que A13 o acompanhasse até uma outra sala. Ele queria que ele apresentasse a ferramenta para um pesquisador de Cuba que estava visitando o local naquele mesmo dia.

O coordenador se mostrava bastante empolgado. Explicou que o colega cubano estava *“impressionando”* e *“superinteressado”* na proposta da ferramenta. Havia ainda a possibilidade de se estabelecer uma cooperação entre os centros de pesquisa dos dois países, pois o colega tinha *“uma quantidade enorme de dados coletados [...] até falou de uma possibilidade da ferramenta ser usada e em hospitais por lá”*.

A13 se retirou da reunião para acompanhar o coordenador. Enquanto ele não voltava, o desenvolvedor ligou o monitor que estava ao lado da mesa. Se referindo ao software, ele disse *“vou te mostrar a cara dele”*. Ele abriu a aplicação e navegou pela área de gestão dos dados. Mostrou uma tela com planilhas contendo formulários onde eram solicitadas descrições diversas. Tais formulários, segundo o desenvolvedor, serviam para registrar os dados e os metadados de um experimento. Ele passou a relatar uma quantidade de detalhes técnicos acerca da elaboração dos questionários que, em um primeiro momento, pareciam pouco informativos e frios, mas não pude deixar de achar complexo aquilo tudo.

Tal complexidade parecia residir principalmente no fato de que foram empregadas vários outros recursos para chegar à ferramenta, como a seleção de modelos de banco de dados e escolhas de padrões para registro de dados experimentais em neurociência e uma série de outros padrões para a elaboração do questionário para coletar os dados¹¹⁹. Além de todo esse trabalho de programação feito no escritório, o desenvolvedor relatou ainda que ele havia se deslocado até o Rio de Janeiro para reuniões com alguns pesquisadores do NPNR que utilizariam a ferramenta. O objetivo era discutir com esses pesquisadores possíveis formas de

119 Alguma das descrições que devem ser registradas referem-se à avaliação clínica do paciente, como exame físico, avaliação de fisioterapia, ficha sobre procedimentos cirúrgico, grau de lesão, entre outras informações.

tratar os dados sensíveis de pacientes, já que aquilo havia sido uma demanda daquele grupo de pesquisadores.

Naquela interação com o desenvolvedor, percebi que parecia haver muito trabalho envolvido em torno das ferramentas, além disso se atribuía muitos papéis ao software.

O NES é antes de tudo uma ferramenta de trabalho. O pesquisador A13 afirmou que o software nasceu para facilitar a cooperação entre pesquisadores a partir do intercâmbio de dados, metadados e descrição dos protocolos experimentais. Objetivava também à integração entre equipes do ambiente hospitalar e laboratório a partir de um ambiente comum para relatar e armazenar os dados do paciente.

É atribuído ao software a missão de tornar os princípios de compartilhamento de dados difundidos e assimiláveis entre seus usuários. A função da ferramenta nesse aspecto é essencialmente padronizar, desde a forma como os dados são registrados até a forma de coleta desses dados. Ao ser inserido em uma dinâmica de trabalho, o NES assume a função de incentivar o pesquisador responsável por produzir os dados a registrar de uma forma também padronizada.

Alguns informantes associavam ainda à construção, manutenção e curadoria de bancos de dados um papel ligado ao avanço efetivo na compreensão do funcionamento do cérebro e para o tratamento de doenças. Como é analisado adiante, essa preocupação com as ferramentas e seus processos subjacentes, como a padronização, é recorrente em diferentes manifestações, sendo postos como parte dos “desafios computacionais” da modelagem neuronal como área que necessita de instrumentos de registro, armazenamento e acesso aos dados experimentais visando à reutilização dos mesmos.

5.2 Padrões e infraestruturas de compartilhamento

Como pode ser observado até aqui, falar de Ciência Aberta no NeuroMat significa considerar o compartilhamento de dados e ferramentas, mas principalmente “padrões”. A associação da abertura de dados a partir de padrões específicos se relaciona com o aspecto fundamental da reutilização dos resultados de pesquisa e informações associadas a eles.

A13 menciona com frequência a noção de “*agregação de valor aos dados*”. A forma de se agregar valor se dá por meio de metadados associados aos dados. Tais metadados trarão uma descrição do que é, da origem e do próprio significado de cada dado, imprimindo maior facilidade de recuperação do conjunto nos bancos de dados. O processo de curadoria, por sua vez, agrega qualidade aos experimentos, como reprodutibilidade e meta-análises. Durante o II Encontro de Ciência Aberta, A13 descreveu:

No cenário atual os processos se dão por meio de análise computacional dos dados experimentais e os dados abertos têm um papel importante nisso, porque permitem uma ciência com mais qualidade e reprodutível mas existem vários desafios, como a falta de formatos padronizados para a representação dos dados. A representação digital dos dados de pesquisa necessita de padrões que possibilitem que sejam armazenados de forma adequada, eficiente e segura e que seja de fácil manuseio pelos pesquisadores. É necessário também metadados sobre sua origem e a estrutura para que os dados abertos sejam compreendidos e possam ser reutilizados. Essa curadoria dos dados demanda que existam repositórios de dados e software livre, inclusive para manutenção dos dados a longo prazo¹²⁰.

Essa noção de agregação de valor através dos metadados e protocolos foi ilustrada de uma outra maneira por um engenheiro da computação que acompanhava o *Workshop Random Graph in the Brain*. Comentei que me faltava clareza sobre o papel dos padrões de metadados no processo de compartilhamento de que os modeladores tanto falavam. Ele começou falando que os metadados estão intrinsecamente ligados a um processo de reutilização dos dados experimentais. Ele, como um cientista de dados, ou seja, alguém que analisa e extrai conhecimento dos dados, tem a necessidade de ter acesso à especificação de como a coleta de dados experimentais foi desenvolvida. Durante a nossa conversa, o engenheiro da computação reforçou em muitos momentos essa ideia de que um futuro ideal para a neurociência seria sua constituição cada vez mais como uma ciência de dados. Esse pesquisador relatava que os metadados associado aos dados e à padronização poderiam fornecer “inteligência” e capacidade para a reutilização tanto dos códigos das ferramentas computacionais empregadas quanto dos dados obtidos nas pesquisas experimentais.

Esta ideia de reutilização está muito associada a um esforço para não duplicar o trabalho – a que os modeladores se referem como “*Dont repeat yourself*” (DRY). Esses modeladores acreditam que a padronização evitaria alguns esforços desnecessários ou que acabam levando muito tempo. Evitaria, por exemplo, que neurocientistas e outros cientistas da vida tenham que desenvolver habilidades computacionais (como aprender a programar)

120 Registro de apresentação no II Congresso de Ciência Aberta.

para analisar os próprios dados; assim como seria fundamental para aqueles que possuem habilidades computacionais, mas não têm domínio do paradigma experimental ou não sabem de onde extrair dados de qualidade. Nesse futuro ideal, os neurocientistas não necessitariam desenvolver habilidades de programação avançadas, uma vez que a padronização possibilitaria a eles se tornarem responsáveis por produzir as hipóteses, coletar, organizar e disponibilizar os dados, enquanto os modeladores ou cientistas de dados estariam responsáveis pelo desenvolvimento e aplicação das ferramentas estatísticas para extrair resultados relevantes desses dados.

Outra possibilidade de aplicação dos metadados pode ser o desenvolvimento de ferramentas que permitam obter, marcar (atribuir *tags*) e relacionar os dados com os metadados de maneira automática. Essas características são apontadas como aspectos que qualificam certos bancos de dados como de “qualidade”. Esses tipos de funcionalidades dependem da existência de padrões de metadados, como modelos de registro de dados que sejam consenso e compartilhados pela comunidade de neurocientistas e modeladores.

Star e Bowker (1999, p. 13) definem padrões como “conjunto de regras que foram acordadas para a produção de objetos (textuais ou materiais)”¹²¹. Do ponto de vista do desenvolvimento de sistemas técnicos, os padrões funcionam como “*gateways*” ou arranjos sociotécnicos que permitem a interoperabilidade com outros sistemas possibilitando a utilização como um único sistema integrado (STAR; RUHLER, 1996; RIBES, 2014). Assim como as normas sociais e padrões culturais orientam o comportamento dentro de uma comunidade (STAR, 1999), o padrão implica a “diluição de diferenças” e a “adoção de uma base comum” (VICENTIN, 2016, p. 110), atuando na redução das lacunas entre culturas epistêmicas (LEONELLI, 2016)¹²². Nos sistemas de compartilhamento de informação, a base

121 De acordo com Vermeulen et al (2013, p. 168), a padronização, em diversos sentidos, é um dos principais fatores que motivam o aumento dos níveis globais de colaboração nas ciências da vida, sobretudo nas últimas décadas. Como um motivador da colaboração científica, a padronização associa-se à busca de “compressão” ou “aceleração” do tempo das pesquisas por meio de instrumentação avançada e à internacionalização da ciência. Nesse contexto, os padrões comuns cumprem o papel de criar dados compatíveis a partir de diferentes fontes e locais. A codificação universal do DNA, por exemplo, usando as letras A, G, C e T, associada à negociação de métodos comuns de pesquisa para decifrar o código foram fundamentais para o sucesso do Projeto Genoma Humano.

122 Baker et al (2016) sugere que o significado dos padrões pode ser entendido por meio da ideia de interoperabilidade, que aparece como um processo que objetiva a montagem e o compartilhamento de dados, permitindo a reutilização por uma variedade de indivíduos, através de disciplinas divergentes, territórios e/ou por longos períodos de tempo. A interoperabilidade de dados, embora muitas vezes vista como uma questão técnica em termos de escolhas padrão, desdobra-se em uma preocupação mais complexa diante de comunidades, organizações e questões sociotécnicas, uma vez que o trabalho de alcançar a interoperabilidade é sempre um trabalho de reorganização, construção de comunidade e tomada de decisão e escolhas técnicas.

comum é representada por padrões que visam à interoperabilidade ou livre circulação de informações em diferentes ambientes, como vocabulários controlados, ontologias e padrões de metadados (BAKER et al, 2006).

De acordo com Edwards et al (2011), os padrões de metadados são instrumentos essenciais na redução da fricção entre as diferentes disciplinas, sendo colocados como objeto de preocupação de desenvolvedores e pesquisadores envolvidos com o desenvolvimento de sistemas compartilhados. Os autores utilizam o conceito de “fricção de dados” para se referir aos os pontos em que os dados se movem entre pessoas, laboratórios, computadores, disciplinas e formatos, produzindo esforços e tempo adicionais, energia e atenção humana. A fricção de dados produz o que autores chamam de “fricção de ciências”, ou seja, os atritos e dificuldades produzidos quando duas ou mais disciplinas científicas que trabalham em problemas comuns tentam interoperar.

No contexto matemático e computacional aqui retratado, dada a complexidade e o grande volume de dados para serem analisados por pequenas comunidades, o desenvolvimento de modelos neuronais significa, em muitos momentos, um trabalho orientado à formação de “terreno comum” entre áreas e pessoas envolvidas em sua produção. Tal terreno comum é representado por infraestruturas de compartilhamento como repositórios, padrões de classificação e suas ontologias subjacentes e protocolos que possibilitem essas disciplinas “conversarem” entre si (EDWARDS, 2017; EDWARDS et al, 2009).

Entretanto, os informantes apontaram que muitos bancos de dados neurofisiológicos, empregados largamente nas pesquisas de modelagem e outras técnicas analíticas, apresentam uma série de problemas. Alguns desses problemas foram recorrentes nas interações com o grupo de trabalho envolvido com o NES.

As pessoas envolvidas relataram suas experiências com o desenvolvimento de bancos de dados de experimentos de neurociência, explicitando como modeladores compreendem o papel dos padrões de dados. Nas interações com alguns desenvolvedores (como B2 e B3), à medida que contavam de suas próprias tarefas diárias, ficou evidente que o trabalho de desenvolvimento do banco de dados é orientado por fatores considerados por eles como “a realidade” dessas ferramentas no campo da neurociência. Entre esses fatores estariam as “limitações técnicas” à construção das bases de dados de acesso aberto em neurociência, sobretudo em relação à “documentação” e ao “conhecimento computacional” insuficientes dos neurocientistas para a manutenção dos bancos de dados.

Nas descrições feitas por A13, ele afirmava que não se tratava de simplesmente abrir os dados, mas que havia toda uma preocupação envolvendo o registro de metadados e descrição de protocolos experimentais. Havia naquele momento uma expectativa de que o software incentivasse, por exemplo, os pesquisadores a registrarem os protocolos utilizados nos experimentos, prática pouco difundida na neurociência. O compartilhamento dos protocolos experimentais é parte da reprodutibilidade na neurociência. Porém *“Não existe consistência nos protocolos experimentais. Cada pesquisador faz de um jeito [...]”*. Ele via a importância do detalhamento de parâmetros de configuração dos equipamentos utilizados pelo pesquisador durante a pesquisa para a reutilização dos dados da seguinte forma:

Ele [o pesquisador], antes de inserir os dados vai descrever o protocolo experimental. Tem que falar (sic) se colocou o experimento ‘assim ou assado’ (sic), se usou luva e que tipo de luva foi [usado], se colocou o eletrodo em tal posição... essas coisas. [Entrevista com A13]

Em relação à quantidade, variedade de dados experimentais e técnicas de coleta dos dados, estas são apontadas como um problema para a qualidade do armazenamento. O mesmo informante enfatizou que:

A quantidade e variação dos dados produzidos na pesquisa experimental em neurociência demandam ferramentas de banco de dados que comportem essas características e que permitam também a reutilização dos dados e a produção de resultados analíticos relevantes [Entrevista com A13].

Falava-se ainda com frequência de uma *“baixa qualidade”* das ferramentas existentes que estaria associada às funcionalidades técnicas das bases, ainda que na maior parte das vezes remetesse à ausência de algum tipo de padronização: *(Nota de conversa com A9) “é muito complicado acessar os dados [nessa área] por uma falta de padrões que permitam a convergência de dados e de ferramentas”*. Em outro momento, A13 disse que *“o desenvolvimento de descritores adequados do funcionamento neuronal é um dos desafios no desenvolvimento de bases de dados [em neurociência]. Os parâmetros para descrever as lesões neurológicas em bases de dados não são consolidados”*, indicando as dificuldades técnicas existentes no desenvolvimento dessas ferramentas.

5.3 Ferramentas abertas e as redes de validação

O desenvolvimento do NES levou em conta padrões para compartilhamento e construção de bases de dados neurofisiológicos estabelecidos internacionalmente¹²³. Parte do trabalho dos desenvolvedores e cientistas da computação envolvidos com o projeto foi levantar e analisar iniciativas internacionais para padrões de bases de dados neuronais com o objetivo de garantir uma integração com um circuito global de pesquisa. A outra faceta dessa preocupação com “*seguir um padrão*” manifesta-se nas relações científicas e institucionais estabelecidas como projetos internacionais com propósitos semelhantes ao que vinha sendo desenvolvido internamente. Essa aproximação busca geralmente possibilitar o contato da equipe com “*boas práticas em compartilhamento de dados*”. Observou-se também um interesse em alcançar uma “aceitação” da comunidade e reconhecimento pelas contribuições do centro por meio das ferramentas desenvolvidas, possibilitando a integração da ferramenta a uma infraestrutura de informação global.

Desse modo, esses padrões utilizados parecem assumir um papel que vai além da sua dimensão “técnica” que visa o funcionamento das infraestruturas e, como tal, é um trabalho restrito aos técnicos. O trabalho por trás do desenvolvimento do NES sugere que o processo de padronização, associado à abertura de dados guardam questões disciplinares/científicas mais amplas que perpassam as áreas envolvidas com a produção de modelos e com a própria disciplina da neuromatemática.

Isso ficou bastante evidente durante o workshop “*High-Performance Computing, Stochastic Modeling and Databases in Neuroscience*”, um dos eventos que acompanhei durante a pesquisa de campo e no qual alguns desses assuntos foram debatidos¹²⁴. Este workshop foi o primeiro a discutir em toda a programação o compartilhamento de dados e ferramentas de código aberto na modelagem neuronal. Por isso foi bastante importante nas atividades do centro. Durante sua realização, ocorreu o lançamento do supercomputador do NeuroMat e foi realizada uma série de interlocuções com representantes de projetos de ferramentas de simulação de redes neuronais em grande escala. Buscou-se com isso divulgar o NES e inseri-lo em uma rede de ferramentas utilizadas na área de modelagem.

123 Assim que foi lançado houve uma repercussão positiva na comunidade científica, o que muitas vezes parecia ser visto pela equipe como um demonstrativo de sucesso do software. A equipe relatava interesse de algumas em utilizar a ferramenta.

124 O workshop foi realizado para estreitar relações com os principais centros de pesquisa internacionais sobre cérebro. Foi realizada também uma discussão sobre participação do NeuroMat na INFC e a colaboração do NeuroMat com o grupo de trabalho sobre padrões para compartilhamento de dados eletrofisiológicos. O evento teve a participação de pesquisadores estrangeiros dos principais empreendimentos científicos de pesquisa sobre o cérebro, incluindo BRAIN, financiado pelo governo dos EUA, e o Projeto Human Brain, financiado pela Comissão Europeia, Allen Institute, também dos EUA, Riken Brain Science Institute, do Japão.

Para a pesquisa de campo, o workshop forneceu elementos para uma compreensão dos aspectos operacionais do compartilhamento, possibilitada pelo detalhamento de diferentes ferramentas computacionais durante as exposições. Em muitos momentos esses tópicos relacionados à integração se manifestaram de diferentes maneiras, desde os aspectos tidos pelos sujeitos como “estritamente técnicos” a uma discussão sobre cooperação acadêmica em seu sentido mais tradicional, que acaba recaindo nos desafios computacionais para o avanço da colaboração científica em neurociência. Menciono especificamente a discussão realizada em duas mesas redondas em que se debateu desafios e as novas configurações da colaboração científica na área de neurociência computacional.

Em uma das mesas do evento, representantes do NeuroMat e pesquisadores estrangeiros discutiram a pesquisa em neurociência no contexto de uma realidade de *big science* na neurociência¹²⁵, na qual se falou do papel do compartilhamento de recursos tecnológicos e dos dados como lastros da pesquisa em neurociência na atualidade. O que percebi durante essa mesa foi que para esses pesquisadores existe uma tendência da pesquisa sobre as bases funcionais do cérebro em se organizar a partir dos grandes projetos de longo prazo que envolvem muitos recursos, técnicas, equipes trabalhando em conjunto e equipamentos computacionais de alto desempenho¹²⁶.

Nesse cenário, diante de uma realidade que se volta para a *e-science*, os pequenos laboratórios ou os pequenos grupos de pesquisa têm pouca capacidade de mobilizar os recursos necessários para tratar grandes quantidades de dados. Consequentemente, os pequenos laboratórios passam a ter pouca ou nenhuma chance de darem contribuições significativas à pesquisa sobre o cérebro, à medida que a produção de conhecimento nesse campo vem se dando cada vez mais pela capacidade dos laboratórios modelarem grandes quantidades de dados. Para o pesquisador, a resposta está na consolidação de um ambiente aberto para convergirem os desenvolvimentos em neurociências.

Nesse aspecto, a criação de repositórios de uso comum possibilitará aos pequenos laboratórios participarem, disponibilizando dados e reutilizando ferramentas e outros recursos de pesquisa produzidos por laboratórios com capacidade produtiva mais elevada. Esses pontos

125 Mesa “‘Big science’: the case for neuroscience. What are the goals? What are the research questions? What can neuroscience benefit from the big science approach”.

126 Os marcos para essa configuração da área de neurociência são os projetos Brain Initiative, lançado pelo governo dos EUA em 2013 e Human Brain Project da União Europeia, também lançado em 2013. Ambos os projetos fazem parte de políticas estratégicas que visam incentivar a pesquisa sobre o cérebro humano.

foram reforçados em outra mesa¹²⁷ em que se debateu formas de viabilizar o compartilhamento e o desenvolvimento colaborativo na modelagem. Mostrou-se como consenso dessa discussão que o “*pilar*” da colaboração científica e uma “*grande prioridade*” para a neurociência deve ser o compartilhamento aberto de dados. De acordo com um dos palestrantes, existe uma mudança ocorrendo na área em relação à percepção sobre a disponibilização aberta dos dados.

Tais pontos se manifestaram no uso recorrente, pelos modeladores, das palavras “*convergência*” e “*integração*” em diferentes momentos e contextos. O uso dessas expressões aparece tanto associado a esse novo paradigma para a neurociência, como na forma de um “princípio orientador” do desenvolvimento de ferramentas, isto é, o desenvolvimento das ferramentas computacionais busca convergir ou integrar as ferramentas a outras infraestruturas técnicas. Nesse contexto, o compartilhamento, os dados e o código aberto remetem direta ou indiretamente a essa ideia de cooperação ou de “interoperabilidade global” em que sistemas técnicos devam convergir¹²⁸. Essa dificuldade relatada como um dos principais desafios da modelagem neuronal é associada às poucas normas e requisitos de compartilhamento existentes na comunidade de neurocientistas e modeladores. Um dos modeladores mencionou, inclusive, a necessidade de um “*código da pesquisa*” compartilhado na neurociência, em uma referência ao código-fonte do software e às normas científicas de publicação dos resultados. Assim como o abrir o código-fonte do software possibilita o aperfeiçoamento da ferramenta, da mesma forma, abrir o “código fonte da pesquisa”, possibilita que as ideias e os conhecimentos por trás dos modelos possam ser testados, reproduzidos, reaproveitados e aperfeiçoados por outros pesquisadores.

Um momento que pareceu sintetizar todos esses pontos ocorreu durante a discussão da modelagem de integração de dados anatômicos e fisiológicos em modelos computacionais unificados, ocorrida em um dos dias do evento como parte do Workshop mencionado. Esse tipo de modelagem demanda o desenvolvimento de um modelo em que diferentes resultados experimentais sejam integrados a uma simulação computacional. Por isso a modelagem de integração foi descrita como um desafio colocado para a biologia computacional, trazendo uma complexidade para a realização de simulações das interações biológicas em múltiplas escalas e múltiplos algoritmos.

127 Mesa do Evento: “*Open Source Brain: enabling sharing & collaborative development of models in computational neuroscience*”.

128 Cf. Gleeson et al (2010).

A modelagem integrativa ocorre em um cenário de computação intensiva, visto que a realidade que busca representar envolve um grande volume de dados e pessoas dedicadas ao gerenciamento e manutenção de arquiteturas, bases de dados, ferramentas, artefatos de informação digital e códigos. No modo como foi colocada, a integração de modelos parece representar uma evolução na modelagem de organismos vivos, em parte porque opõe-se a um cenário de fragmentação – ou, como mencionou um dos palestrantes, um cenário em que ainda “persistem silos”, que parecem impedir o desenvolvimento de modelos do sistema nervoso mais precisos e realísticos¹²⁹. Para ilustrar as possibilidades representadas pela integração de esforços foi feita a demonstração de um modelo computacional chamado de *OpenWorm*. O modelo faz parte de um projeto de código aberto que simula, por meio de uma animação computadorizada, os comportamentos moleculares, celulares e mecânicos do verme *C. Elegans*¹³⁰. O modelo foi desenvolvido por um grupo de pesquisadores norte-americanos que trabalha com a modelagem integrativa.

A demonstração da ferramenta foi feita pelo pesquisador da *University College London*, o neurocientista computacional Padraig Gleeson. Após uma sessão em que foram apresentados diferentes projetos de simulação, a apresentação de Gleeson chamou atenção em um primeiro momento porque, para um leigo, a animação não impressionava. Era uma animação em 3D simplória de algo que lembra uma larva.

129 Para um detalhamento do desenvolvimento do OpenWorm Cf. Szigeti et al. (2014).

130 O *C. elegans* é um nematelminto. Teve seu genoma sequenciado em 1998. Cf. Szigeti et al. (2014).

Entretanto, durante a exposição, ficou mais claro que aquilo era resultado de um extenso trabalho que o próprio pesquisador definiu como “*um dos principais desafios da modelagem de organismos vivos*”. A medida da extensão do trabalho com a produção dessa simulação era dada pela menção ao esforço na coleta de diferentes conjunto de dados, no desenvolvimento de padrões de validação. Um simples animal se contorcendo, do ponto de vista da biologia computacional, é uma importante contribuição para a compreensão do funcionamento do sistema nervoso de organismos vivos e como ele responde à força física¹³¹. À medida que o *C. Elegans* se contorce na tela (Figura 11), ocorrem mudanças de coloração em diferentes pontos da figura, indicando atividades elétricas do sistema nervoso e outras células do animal.



Figura 9: Pesquisador exhibe o modelo que simula o E. Elegans.

Fonte: Arquivo do autor.

¹³¹ A biologia computacional busca simular o funcionamento de organismos abrangendo aspectos celulares, moleculares e mecânicos (Op. cit.).

Enquanto interagía com a figura utilizando o cursor do mouse, Gleeson explicou com empolgação que o modelo simulava o comportamento de tecidos “*moles*” do verme e buscava recriar a “*rica complexidade biológica real do animal*”, como as interações sensoriais e o movimento. O modelo permitia também “descascar” a pele, olhar as células e girar o organismo para olhá-lo de diferentes ângulos.

Como foi possível perceber na exposição de Gleeson, a criação do *OpenWorm* demandou o desenvolvimento de uma série de outras ferramentas complementares, cada uma cumprindo uma função diferente para fazer o modelo funcionar, conforme Figura 12, a seguir. Em muitos momentos de sua exposição ele ressaltou que o desenvolvimento daquele organismo virtual, como a “*fidelidade*” e a “*qualidade*” que o projeto do *OpenWorm* produziu, somente foi possível devido a um trabalho “cooperativo e de Ciência Aberta” que se estabeleceu no desenvolvimento de cada uma das ferramentas e no compartilhamento de informações entre os participantes do projeto.

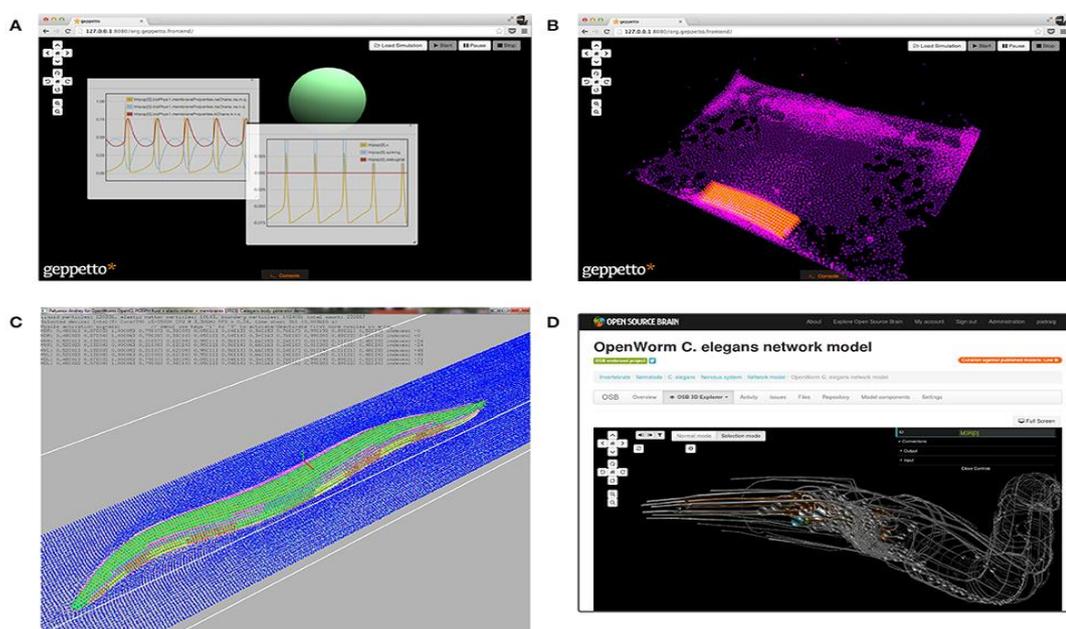


Figura 10: Componentes desenvolvidos para o *OpenWorm*.

Fonte: Szigeti et al (2014).

O que chamou atenção no detalhamento do projeto foi o papel atribuído às linguagens de representação do modelo. Foi empregada uma linguagem de descrição tida como “*consenso na comunidade para descrever os componentes do modelo*”. Aqui, novamente, os padrões aparecem como viabilizadores da modelagem e como parte das infraestruturas de compartilhamento e foi apontado a necessidade de definição de “*terminologias comuns*” para que desenvolvimentos como aqueles sejam possíveis.

Padrões, tais como as terminologias, funcionam como linguagens que agem na redução da fricção entre as diferentes culturas epistêmicas. As linguagens de representação ou de codificação, incluindo os padrões de metadados, têm o papel de estabelecer condições para o entendimento entre pesquisadores, funcionando como uma forma de “expressar” os modelos em uma linguagem compreensível por outros modeladores. De maneira prática, as terminologias permitem a interoperabilidade entre sistemas de informação, possibilitando o trânsito, arquivamento e reutilização dos modelos em diferentes territórios disciplinares, geográficos e epistêmicos. Tecnicamente, essas codificações possibilitam que os modelos sejam mapeados e lidos para aplicativos de validação¹³².

Como um modelo integrativo, o OpenWorm utilizou diferentes fontes de conjuntos de dados e com diferentes formatos. Assim como outros projetos da área de modelagem neuronal, empregou um padrão de descrição de modelos neuronais chamado de “NeuroML”¹³³.

O formato NeuroML é uma linguagem comum para a descrição ou codificação detalhada das propriedades dos modelos neuronais. Permite a interoperabilidade dos modelos e suas propriedades entre diferentes ambientes virtuais como ambientes de simulação. Na página da documentação do projeto OpenWorm, afirma-se que “Uma representação NeuroML é útil para responder a questões sobre morfologia de modelos e parâmetros de simulação” (OPENWORM, 2019).

132 Embora a ideia de “validação” possa ser entendida aqui também como certificação e aceitação dos modelos pelos pares, os modeladores de fato trabalham com ferramentas de validação, isto é, simuladores virtuais que importam e “rodam” os parâmetros dos modelos, conforme explicou o modelador A3. Alguns dos validadores citados por A3 são os simuladores virtuais Genesis (<http://www.genesis-sim.org>) e Neuron (<https://www.neuron.yale.edu/neuron>).

133 O formato NeuroML é um projeto de código aberto baseado na linguagem de marcação *Extensible Markup Language* (XML). Cf. <<https://www.neuroml.org>>. Para um maior detalhamento Cf. Gleeson et al. (2010).

À medida que possibilita que os modelos transitem e sejam reutilizados em diferentes contextos, essa linguagem de representação insere-se nesse lugar de integração da comunidade de modeladores em torno de um esforço de criação de condições para a reutilização e validação dos modelos por uma comunidade mais ampla – visando à criação de modelos precisos e de qualidade (GLEESON et al, 2010).

No momento em que descreve o significado da linguagem para o modelo OpenWorm, o pesquisador afirma que um padrão de descrição comum cria condições para que os modelos sejam detalhados e expostos à comunidade de neurocientistas com a finalidade de validação.

O compartilhamento dos modelos é uma “troca de ideias” [...] porque os modelos carregam visões e conhecimentos de quem os produziu, e quando você compartilha o modelo, você está transmitindo uma ideia, que precisa ser compreendida e fazer sentido para o outro modelador¹³⁴.

Bowker (2005) afirma que os padrões podem ser concebidos como ferramentas para estabilizar as configurações da infraestrutura. No limite, a padronização funciona como um instrumento que visa a agregar valor aos sistemas técnicos de compartilhamento seguindo um processo “evolucionário” típico das infraestruturas, que os torna um “sistema de sistemas” (LAGOZE et al, 2015). Percebe-se que para se estabelecerem de fato como uma infraestrutura de pesquisa, isto é, serem incorporados ao trabalho diário, viabilizando e estruturando práticas desses pesquisadores e respondendo da maneira esperada quando acionados, os sistemas técnicos utilizados por esses pesquisadores precisam incorporar uma série de outros recursos que viabilizem seu “bom” funcionamento, como padrões técnicos estabelecidos e utilizados pela comunidade. Nota-se aí a razão de muitos dos esforços em termos de desenvolvimentos tecnocientíficos de áreas como a neurociência computacional serem voltados para soluções como padrões de metadados, taxonomias e ontologias.

Do mesmo modo, houve um esforço interno em adotar esses padrões no desenvolvimento do NES como forma de integrá-lo a essa infraestrutura de informação mais ampla. No entanto, a ferramenta depara-se com fricções de outra natureza, como abordado no tópico a seguir.

134 Registro de fala durante Workshop.

5.4 O software como produto da atividade acadêmica: dilemas de abertura e fechamento

Alguns entrevistados relataram que um dos desafios enfrentados durante o desenvolvimento do NES foi possibilitar que o mesmo respondesse às necessidades dos pesquisadores durante sua utilização. Os principais envolvidos no desenvolvimento da ferramenta (B2 e A13) apontam a interoperabilidade com outros sistemas e formatos como uma das limitações de softwares de código aberto no contexto da pesquisa experimental.

O NES foi construído segundo os princípios de desenvolvimento do software livre. No entanto, em razão da natureza dos equipamentos da experimentação em neurociências, foi preciso fazer algumas adaptações. A13 explicou que os equipamentos de EEG – que fazem o registro conectado a um computador – produzem imagens em um formato proprietário, o que significa que as imagens devem ser lidas em software específico que acompanha o equipamento. De acordo com a pesquisadora, se as imagens forem exportadas para leitura em software livre, alguns dados acabam se perdendo, prejudicando assim a obtenção de resultados precisos. Essa realidade exigiu do grupo de trabalho o desenvolvimento de uma solução que permitisse que a ferramenta “conversasse” de maneira adequada com esses equipamentos, implementando assim formatos e protocolos proprietários.

De acordo com Levin e Leonelli (2016), o desenvolvimento de ferramentas de softwares acadêmicos depende de uma mistura ou de um remix de aplicações de código aberto com aplicações proprietárias. Essa mistura ocorre justamente porque equipamentos de laboratórios e outros recursos essenciais de pesquisa são geralmente produzidos e comercializados exclusivamente por empresas que utilizam softwares proprietários nos equipamentos e que, ao longo do tempo, tornaram-se indispensáveis em determinadas etapas da pesquisa e para a produção da evidência científica, não sendo possível para o pesquisador prescindir de utilizá-los. Nesse contexto, as ferramentas abertas acabam incorporando determinadas propriedades dos artefatos e protocolos fechados/comerciais para poder interoperar e serem aceitas e utilizadas pela comunidade de pesquisadores.

Ainda que a ferramenta seja concebida como uma ferramenta aberta, os objetivos para os quais foi criada acabam se defrontando com práticas fechadas e formatos proprietários. A esse respeito, trabalhos como o de que Wiley e Michaels (2004) sugerem que esta é uma peculiaridade das ferramentas de softwares para a pesquisa em biologia. Eles argumentam que, por existirem muitas subdisciplinas, foram desenvolvidas ferramentas orientadas para tarefas muito específicas do campo, o que geralmente significa a existência de

uma base de usuários muito pequena para cada software. Ao mesmo tempo, as habilidades de desenvolvimento dos pesquisadores nessas áreas (incluindo a neurociência) – talvez com exceção da genômica e da bioinformática – são mais rudimentares e, portanto, as opções de ferramentas de código aberto são mais limitadas.

Para Halchenko e Hanke (2012, p. 1):

À medida que os cientistas, estudantes e grupos de pesquisa responsáveis pelas ferramentas passam para novas tarefas, os softwares desenvolvidos são deixados de lado, sem suporte permanente para correções de *bugs* e manutenção suficientemente coordenada. Consequentemente, as ferramentas terão pouca ou nenhuma atualização ao longo do tempo. Estes, entre outros fatores, levam ao predomínio dos softwares comerciais proprietários e de empresas que fornecem suporte a essas ferramentas nessas áreas.

O projeto de desenvolvimento do NES se deu em uma realidade particular. De acordo com A13, os recursos do Programa CEPID são relativamente altos e permitem a contratação de desenvolvedores e outros profissionais capacitados para criar e manter as ferramentas computacionais do projeto por meio de bolsas de treinamento técnico. Ele destaca que outros projetos semelhantes que requerem o desenvolvimento do mesmo tipo de ferramenta de software não costumam dispor dos mesmos recursos, o que muitas vezes trazia a descontinuidade dessas ferramentas após a finalização do projeto. Embora existissem recursos disponíveis no período de vigência do programa CEPID, a manutenção desse tipo de software tem custos incorporados, de modo que A13 afirmou que não sabe qual seria o futuro da ferramenta após esse período, uma vez que *“a FAPESP não disponibiliza recurso para manter o software depois que acabar o projeto”*.

O ponto levantado pelo pesquisador – que colocava em suspenso o futuro do software – relaciona-se a um aspecto que vai além do papel do software como ferramenta viabilizadora da Ciência Aberta, colocando-a também como parte das ações de transferência de conhecimento e tecnologia do CEPID, sendo, portanto, também um produto tecnológico.

O que me chamou a atenção, logo no início da pesquisa, durante a leitura do projeto de pesquisa do NeuroMat submetido à FAPESP, foi a ideia de Ciência Aberta como parte da estratégia de pesquisa, inovação e transferência de conhecimento, tendo como objetivo o desenvolvimento de ferramentas computacionais abertas para gerenciamento, curadoria e compartilhamento de dados científicos e criação de um banco de dados de acesso aberto. O relatório de atividades do centro de pesquisa de 2015 reforça ainda esses objetivos¹³⁵.

A abordagem fundamental deste desenvolvimento tem sido o que geralmente é chamado de ciência aberta, ou seja, um conjunto de diretrizes para garantir o acesso, uso e confiabilidade de dados, resultados e métodos de pesquisa. Assim, o objetivo do NeuroMat em relação à transferência de tecnologia é gerar ferramentas para que todos os dados produzidos pelo CEPID sejam armazenados em um banco de dados que será disponibilizado publicamente para a comunidade científica, de acordo com os requisitos de privacidade para informações confidenciais (NEUROMAT, 2015, p. 6).

A proposta inicial era de que os esforços de transferência de tecnologia e a produção de inovações estariam voltados para o desenvolvimento de produtos para programas de saúde pública em reabilitação e disponibilizados na rede de hospitais voltados para reabilitação do Governo do Estado de São Paulo, que também forneceria protocolos de pesquisa para serem disponibilizados no banco de dados público. Com auxílio do CCSL, a transferência de tecnologia seria implementada a partir da produção de ferramentas computacionais para pesquisas em neurociências e uso clínico, disponibilizadas em código aberto.

A proposta apresentada pelo NeuroMat à FAPESP se destacava pelo contraste que estabelecia em relação à abordagem comumente dada à transferência de tecnologia na academia e incentivada por agências de fomento – como a FAPESP – e pelo próprio programa CEPID. Se por um lado parecia ser comum uma grande ênfase no patenteamento ou no estreitamento de relações com a indústria, no NeuroMat circulava uma compreensão de que a transferência de tecnologias pode ocorrer por meio de formas de apropriação características de práticas comuns ao FLOSS.

135 Cf. F16

Esta flexibilidade interpretativa sobre a transferência de tecnologia foi, inclusive, origem de alguns conflitos internos e de críticas ao NeuroMat durante a avaliação periódica do CEPID realizada pela FAPESP envolvendo o NES.

Desde o momento de fundação do centro houve uma série de controvérsias em relação ao eixo de transferência de conhecimento e tecnologia. Um dos principais problemas estava relacionado a diferentes compreensões, entre os pesquisadores, sobre transferência de tecnologia e conhecimento. Segundo o histórico feito por um dos informantes, desde o início existiu uma disputa interna que, direta ou indiretamente, envolvia questões referentes a formas de conceber a transferência de tecnologias e inovações para a sociedade.

A primeira pessoa que ocupou a coordenação do Eixo de transferência de tecnologia¹³⁶ tinha uma postura contrária à Ciência Aberta e defendia uma “*política de fechamento*”. Inclusive recusava-se a fazer reuniões com o desenvolvedor (B2) responsável pela parte de desenvolvimento das ferramentas e bases de dados. Por conta desses conflitos, a pessoa deixou o projeto e algumas mudanças na coordenação dos eixos foram efetuadas. O eixo de transferência de tecnologias passou para a coordenação do docente que era responsável até então pelo Eixo de Difusão Científica, como forma de mitigar atritos que vinham sendo causados, inclusive com a FAPESP, já que as ações de transferência de tecnologia não estavam ocorrendo como esperado¹³⁷.

No início do trabalho de campo, um dos meus questionamentos dizia respeito às formas de apropriação tecnológica em áreas duras, como a física e a matemática. No início, a percepção era a de que essas áreas não têm o patenteamento como um padrão de comunicação, ou seja, não originam produtos facilmente patenteados/licenciados. Ao colocar essa afirmação para o coordenador do Eixo de transferência de tecnologia do NeuroMat, ele disse que a forma como as agências de financiamento encaram os desenvolvimentos realizados na área de computação neuronal não é algo tão simples:

Não é assim que funciona... na prática não é tão simples, tudo depende da forma como a FAPESP encara o que a gente faz. Se ela achar que um algoritmo ou software ou mesmo uma fórmula matemática é de interesse da indústria, ela pedirá que o pesquisador licencie aquilo de algum jeito.

136 O nome do pesquisador foi suprimido a pedido da equipe do NeuroMat.

137 Em razão desse remanejamento, o Eixo de difusão ficou um período sem um coordenador, até que um docente da UNICAMP foi convidado para assumir a coordenação das atividades.

O mesmo pesquisador, que é da área de modelagem computacional, relatou uma experiência com a FAPESP nesse sentido. O seu laboratório havia produzido um software de código aberto para auxiliar a elaboração de diagnósticos médicos. A proposta era que fosse disponibilizado na web para utilização gratuita por médicos e instituições médicas credenciados. O projeto era financiado pela FAPESP e os avaliadores da FAPESP recomendaram que o software fosse licenciado, embora não fosse do interesse desse pesquisador e dos seus alunos envolvidos no desenvolvimento licenciar a ferramenta.

Nota-se, assim, uma grande presença da FAPESP no cotidiano do centro, inclusive na forma como os pesquisadores se relacionam com o software de código aberto enquanto produto da atividade científica¹³⁸. Essa influência da FAPESP também é notada em outras atividades da equipe. Em uma das minhas visitas em 2017, nas conversas e circulação pelo corredor, acabei por me dar conta do quanto a dinâmica das atividades da equipe direta ou indiretamente se dão em torno da produção científica do NeuroMat. Por ser um CEPID, periodicamente ocorrem avaliações por parte da FAPESP, dirigidas principalmente à análise da produtividade científica e tecnológica do centro, nos três eixos que o sustentam.

De acordo com B1, que era responsáveis por elaborar o relatório de atividades bianual, houve uma “discussão violenta” na avaliação do CEPID em 2015 em relação às atividades do projeto porque o eixo da transferência foi muito mal avaliado e a estratégia de transferência de tecnologia baseada na abordagem de Ciência Aberta foi duramente criticada pelo comitê científico. Na percepção dos avaliadores, o NeuroMat não produz inovação no que diz respeito às ferramentas de software desenvolvidas. Na avaliação, o NES aparece como

138 A avaliação periódica realizada pela FAPESP consiste na entrega de um relatório de atividades bianual e em uma reunião fechada com um comitê de avaliadores, entre eles cientistas estrangeiros de renome. A produção do relatório influencia muitas das dinâmicas internas. Tive oportunidade de fazer reuniões no NeuroMat em um período próximo da entrega do relatório bianual e muitas das conversas e atividades tinham o relatório como centro ou então ele era mencionado de alguma maneira. Boa parte do trabalho de alguns técnicos envolvia organizar a produção acadêmica dos pesquisadores com essa finalidade. Em uma das visitas, acompanhei um pouco do trabalho de B4, que era responsável por fazer esse controle e organizar a produção do centro. Ele montava uma listagem das publicações feitas pelo NeuroMat com o objetivo de compor o relatório e também divulgar as publicações no site. Na ocasião, ele mostrou uma tabela impressa com dados referentes aos artigos que citavam o NeuroMat indexados no Google Acadêmico. Na tabela estavam listadas 175 publicações (entre artigos, capítulos de livro, anais, e outros) com a identificação do canal onde fui publicado, os autores, data e se a publicação já havia sido divulgada no site do NeuroMat. Seu trabalho envolvia atualizar a listagem diariamente e fazer pedidos reiterados aos pesquisadores para que informassem as publicações recentes. B4 utilizava uma forma considerada por ele mais confiável e precisa para recuperar as publicações dos pesquisadores – a partir do número de processo FAPESP no Google Acadêmico. Como os pesquisadores são obrigados a citar o número do processo, ficava mais fácil encontrar as publicações indexadas inserindo o código entre aspas no Google, além do que a FAPESP somente considera como produção vinculada ao projeto se houver a citação do número de processo na produção.

um ponto fraco do projeto. Isso porque não é um software licenciado e não havia interesse em licenciá-lo. O coordenador do NeuroMat e o coordenador do Eixo de transferência de tecnologia relataram que uma das críticas feitas pelos avaliadores do comitê de avaliação – entre eles um ganhador do prêmio Nobel – foi justamente no sentido de questionar o porquê do centro investir esforços no desenvolvimento de uma ferramenta como o NES, uma vez que já existem softwares comerciais proprietários que cumprem as mesmas funções a que se propõe o software do NeuroMat.

Em entrevista concedida em julho de 2017, B1 e A1 explicaram que se comparado ao início do projeto, o NES estava caindo em prioridade naquele momento para a estratégia de transferência de tecnologia. Segundo ele, a medida que as atividades do projeto estavam se desenvolvendo e diversificando, essa ferramenta estava ficando “no final do relatório FAPESP”, o que mostrava que sua centralidade vinha diminuindo no aspecto de produtor a ser transferido. Para ele, a FAPESP entende a transferência de conhecimento como uma “forma de negócio”.

Na percepção de B1, o NES só não foi “cortado” do projeto porque parecia haver um interesse da FAPESP em transformá-lo em algo comercializável, algo semelhante ao que já existe em outros softwares de pesquisa do gênero. Em uma de nossas interações, ocorrida logo em seguida à avaliação realizada bisanualmente, B1 comentou que um dos pontos da avaliação FAPESP foi de que o NES poderia ser transformado em uma *startup* para oferecer os serviços para a comunidade científica de forma paga¹³⁹. O outro exemplo ilustrativo disso é que também foi proposto que o projeto de difusão científica – que, ao contrário do eixo de transferência, havia recebido uma boa avaliação da FAPESP em razão dos números apresentados com as atividades centradas em aplicações *wiki* – fosse transformado em uma *start-up* que oferecesse serviços baseados em aplicações *wiki* para a área de educação.

Os informantes relataram que apesar dessa relativa diminuição da centralidade do NES enquanto produto tecnológico e uma inovação, a ferramenta era muito importante no contexto dos objetivos científicos mais amplos do centro de pesquisa e o seu papel nas práticas de modelagem:

139 O programa CEPID incentiva a geração de empresas a partir das tecnologias desenvolvidas pelos centros. O exemplo é o CEPID Genoma Humano e Células-Tronco, localizado na Universidade de São Paulo, que deu origem a empresas especializadas em aconselhamento genético. Cf. <<http://cepid.fapesp.br/materia/100>>.

Se você olhar o relatório é muito curioso porque o NES ‘vai só caindo’ (sic), ficando nas últimas páginas [...] Mas por que ainda assim o NES é importante? Porque ele está relacionado com os novos projetos: Rede Amparo e ABRAÇO. A rede Amparo é um projeto encabeçado por uma outra professora que entrou no projeto recentemente e pelo prof. A5. A ABRAÇO está encabeçada por A2. A proposta é que o NES interligue em um único repositório os laboratórios que fazem parte desse projeto, algo tipo uma ‘open data base’ [Entrevista com B1].

O resultado da avaliação do NES gerou algumas preocupações, levando os coordenadores a repensarem as estratégias de transferência de tecnologia. Não havia, naquele momento, interesse em licenciar a ferramenta ou torná-la comercializável. A principal estratégia foi firmar um acordo com *International Neuroinformatics Coordinating Facility*¹⁴⁰ – organização internacional da área de neurociência computacional que atua no desenvolvimento de ferramentas computacionais e na promoção dessas ferramentas na comunidade científica – para incluir o NES no repositório de ferramentas da rede, que é disponibilizado globalmente.

Alguns pesquisadores, entre eles o coordenador, apontavam que a FAPESP possuía uma posição ambígua em relação às práticas de Ciência Aberta: ou a agência costuma ficar no meio termo ou não se posicionava a respeito, deixando a cargo do pesquisador decidir. Em congressos que tratam do assunto, percebo que as posições da FAPESP limitam-se normalmente a referenciar a iniciativa SciELO¹⁴¹ e os dados disponibilizados por meio de sua biblioteca virtual. Paralelamente, observa-se que os estímulos ao patenteamento e as recomendações para que os pesquisadores busquem publicar em periódicos de considerável Fator de Impacto são bem explícitas.

Entretanto, observa-se também uma crescente aproximação da FAPESP por meio de algumas políticas que compreendem de alguma forma os dados de pesquisa. Entre 2013 e 2015, a agência sediou dois eventos sobre Ciência Aberta. Em um deles foi apresentada a iniciativa da Comissão Europeia para Ciência Aberta ligada ao Programa Horizonte 2020, com a participação do Comissário Europeu de Pesquisa, Ciência e Inovação da Comissão Europeia, Carlos Moedas.

Destaco também a participação da FAPESP na instalação do laboratório do *Structural Genomics Consortium* na UNICAMP e criação do primeiro banco público de dados

140 A INFC é uma organização internacional da área de neurociência computacional que atua no desenvolvimento de ferramentas computacionais e na promoção dessas ferramentas na comunidade científica.

141 A SciELO é uma rede de periódicos de acesso aberto mantida pela FAPESP. Disponível em: <<http://www.scielo.org>>.

genômicos da América Latina – BIPMed¹⁴². Foi criado ainda o programa *E-science* FAPESP para apoiar, entre outras coisas, o desenvolvimento de infraestruturas de pesquisa, incluindo repositórios digitais para o gerenciamento de dados em diferentes áreas do conhecimento¹⁴³. Em 2017, foi instituída a obrigatoriedade dos pesquisadores que tiverem pesquisadas apoiadas pela FAPESP apresentarem um plano de gestão e compartilhamento dos dados gerados pelos projetos temáticos¹⁴⁴. Foi criado também um grupo de trabalho, coordenado pela FAPESP, com representantes de instituições de pesquisa do Estado de São Paulo para a criação de uma rede de repositórios de dados de pesquisa de acesso aberto¹⁴⁵.

Considerações Finais

O desenvolvimento de ferramentas computacionais está entre as principais atividades do NeuroMat. Nesse contexto, o centro deu início a um projeto para um banco de dados neurofisiológicos de acesso público. Dentro desse projeto, foi criado um software de código aberto para auxiliar o desenvolvimento dessa infraestrutura de compartilhamento. O desenvolvimento do *Neuroscience Experiments System* (NES) passa pela adoção de padrões abertos comuns que permitam a circulação de dados entre diferentes disciplinas. Nesse contexto, padrões de metadados remetem a uma série de aspectos que envolvem, por exemplo, questões sobre colaboração em uma área cuja função é desenvolver ferramentas, como modelos, que precisam ser validadas por uma comunidade mais ampla.

142 BRAZILIAN INITIATIVE ON PRECISION MEDICINE. **Home page**. Disponível em: <<http://bipmed.org/>>. Acesso em 20 mar. 2018.

143 FAPESP. **Programa FAPESP de Pesquisa em eScience e Data Science**. s./d. Disponível em: <<http://www.fapesp.br/8436>>. Acesso em 20 mar. 2018.

144 FAPESP. **Plano de Gestão de Dados**. s/d. Disponível em: <<http://www.fapesp.br/gestaodedados/>>. Acesso em 20 mar. 2018.

145 Cf. Julião (2018).

O NES está inserido em relações heterogêneas e lhe é atribuído diferentes papéis. Em razão dos objetivos e das atividades associados ao NES, percebe-se que o software não é apenas uma ferramenta que viabiliza o acesso a dados experimentais, mas está inserido em uma série de outras dinâmicas que o fazem ser significado como uma ferramenta para viabilizar a modelagem, ora para discutir colaboração científica em nível global ou como um produto da atividade científica que precisa ser transferido para a indústria.

A seguir, apresento as considerações finais da tese a partir da retomada do percurso da pesquisa e dos seus principais resultados.

CONSIDERAÇÕES FINAIS

Esta tese nasceu com o objetivo de abordar a abertura na ciência a partir das configurações e deslocamentos dos significados no espaço onde é produzida, enfatizando, sobretudo, práticas de compartilhamento de dados de pesquisa e o papel das diferentes infraestruturas de informação nesses processos. Ao longo da pesquisa aqui apresentada, busquei tratar da Ciência Aberta e de suas configurações, práticas e infraestruturas a partir de uma etnografia realizada no Centro de Pesquisa, Inovação e Difusão em Neuromatemática (NeuroMat).

Uma escolha fundamental para o desenvolvimento da pesquisa foi dirigir a análise do objeto para os momentos e lugares em que a Ciência Aberta se torna prática e é performada nos discursos e no trabalho dos atores conectados ao centro de pesquisa investigado. Busquei nesta pesquisa deixar de lado a preocupação com pressupostos fixos e estanques relacionados com o objeto para assumir que a ciência e suas práticas não são uniformes ou obedecem a um código único ou seguem uma única direção. Como desfecho para a trajetória trilhada a partir desses objetivos e relatada nos capítulos que compõe a tese, apresento a seguir os principais resultados da investigação, dialogando com os pressupostos teóricos que guiaram a feitura do trabalho.

Apresentar os conceitos e princípios por trás do tema da Ciência Aberta foi objetivo do primeiro capítulo, no qual busquei explorar as bases conceituais da Ciência Aberta como movimento e como abordagem para a pesquisa científica. Inicialmente, apresentei uma revisão acerca das normas sociais que historicamente marcam a ciência como uma organização pautada pela comunicação pública dos resultados da atividade científica – prática que sustenta o sistema de distribuição de crédito e reconhecimento na academia.

Apesar de muitas vezes a comunicação pública e aberta do conhecimento científico ser aceita como uma característica incontestável e incontornável da ciência, a história da C&T é marcada por diversos conflitos e disputas em torno da circulação e apropriação dos resultados de pesquisa. As controvérsias em torno da privatização ou fechamento dos resultados científicos – por meio de mecanismos jurídicos, políticos, técnicos, etc – são fatos que marcam, por exemplo, a história do desenvolvimento da biotecnologia e da informática.

É importante pontuar que a noção contemporânea de Ciência Aberta possui uma grande interface com os movimentos e cultura do Software Livre e de Código Aberto (FLOSS – *Free and Open Source Software*). Essa interface traz novas formas e possibilidades de abordar a produção e a comunicação do conhecimento científico, tomando como referência a produção colaborativa e o compartilhamento aberto. As práticas de abertura na ciência inserem-se também em um contexto de tensões existentes entre as práticas sustentadas no compartilhamento e no livre fluxo de informação com as diferentes formas de privatização dessa informação – sejam elas operadas pelas empresas ou pelas manifestações particulares enraizadas nas disputas por capital simbólico e recursos materiais próprias do campo científico. A filiação do movimento de Ciência Aberta com a cultura aberta e colaborativa própria da Internet traz uma série de novas questões – e ressignificação de outras – para a produção e comunicação do conhecimento científico. Possibilita-se, nessa nova configuração, uma outra forma de desenvolvimento da pesquisa acadêmica e destaca-se a busca por maior dinamismo nas atividades científicas, nos processos de descoberta e inovação, bem como questões diretamente relacionadas com a democratização da ciência.

A Ciência Aberta configura-se na atualidade como um movimento multifacetado, mas também com uma abordagem para o desenvolvimento da pesquisa científica. Albagli (2015) argumenta que os atores envolvidos com as práticas de Ciência Aberta mobilizam múltiplas interpretações e, por vezes, conflituosas, sobre o que é abertura na ciência. De maneira geral, as abordagens recentes sugerem que a Ciência Aberta pode transformar a produção de conhecimento e a própria ciência, tornando-a mais colaborativa, transparente, acessível a diferentes públicos, reprodutível e mais eficiente do ponto de vista das descobertas e soluções que oferece para a sociedade. Coloca-se aí uma outra forma de desenvolvimento da pesquisa acadêmica – que ao mesmo tempo que enfatiza o ethos científico clássico, traz um modo de produção sustentada em uma transparência de instrumentos e processos, colaborativa e maximizada pelo reuso dos dados, em oposição a uma ciência fechada – porque

desenvolvida dentro de um paradigma organizacional da pesquisa centrado na propriedade intelectual e que tende ao secretismo, em razão das disputas próprias do meio científico.

Muitas das políticas e discursos que pautam a Ciência Aberta têm enfatizado os dados de pesquisa como centrais para sua configuração atual. Essa ênfase nos dados representa uma significativa mudança na maneira como os resultados da pesquisa são tornados públicos e reutilizados nos processos de produção de conhecimento e tecnologias. Em essência, surgem novas institucionalidades na ciência, especialmente em relação às formas de produção de pesquisa e sua difusão, incluindo uma reconfiguração nas relações da ciência e dos sujeitos da ciência com o seu entorno, com suas práticas, modos de produzir o conhecimento e comunicá-lo.

O segundo capítulo abrigou o referencial teórico e metodológico que auxiliou no desenvolvimento da pesquisa e das análises apresentadas. A pesquisa vincula-se aos Estudos Sociais da Ciência e da Tecnologia, focando no estudo da ciência enquanto uma prática cultural, o que significa, entre outros aspectos, refletir sobre a ciência como espaço da multiplicidade em que se manifestam diferentes culturas epistêmicas (KNORR-CETINA, 1999). A partir desse referencial, buscou-se, também, compreender a Ciência Aberta inserida no contexto das práticas desenvolvidas no Centro de Pesquisa, Inovação e Difusão em Neuromatemática. A área de atuação do NeuroMat é a modelagem avançada de dados neuronais, coletados em experimentos da área de neurociência. Como parte do seu projeto científico, esse centro de pesquisa passou a atuar como um promotor de Ciência Aberta e a realizar uma série de atividades relacionadas com o desenvolvimento e utilização de ferramentas em processos de compartilhamento aberto.

Por meio de um trabalho etnográfico – realizado entre 2015 e 2017 – estive permanente em contato com a equipe de pesquisadores e equipe técnica do centro, acompanhando as atividades relacionadas à Ciência Aberta que vinham sendo desenvolvidas. A imersão no campo ocorreu, principalmente, por meio do acompanhamento de eventos internos do centro, como workshops, seminários de pesquisa, reuniões da equipe, análise de uma série de materiais produzidos pelo grupo e entrevistas com um grupo de pesquisadores e integrantes da equipe técnica.

Como forma de identificar e interpretar as configurações da Ciência Aberta, recorri ainda aos Estudos de Infraestrutura, corrente dos ESCT que se volta para a análise de elementos heterogêneos que, em relacionamento, formam as infraestruturas que dão suporte ao trabalho de alguém, viabilizando práticas e o funcionamento das instituições e serviços de determinada forma. Em outras palavras, a infraestrutura é considerada para além de sua configuração material, incluindo assim pessoas, normas, padrões, etc. Portanto, tomando como referência a noção de que a infraestrutura é sempre “um relacionamento ou uma regressão infinita de relações, nunca uma coisa” (STAR; RUHLER, 1996, p.4), direcionamos a análise para pensar a Ciência Aberta no NeuroMat inserida em uma rede de relações. Nesse relacionamento se associam uma diversidade de elementos, tais como pesquisadores, técnicos, organizações, códigos de conduta, softwares, banco de dados entre outros elementos do trabalho científico que performam a Ciência Aberta sempre em relação a algo ou alguém.

Ao longo da imersão do trabalho de campo, a minha atenção voltou-se também para comportamentos individuais e coletivos, para as formas discursivas e papéis humanos que pudessem evidenciar o que Susan Leigh Star (1999) chama de elementos “invisíveis” e “intratáveis” da infraestrutura. Seguindo essa maneira de olhar, durante a imersão no campo e no momento da produção do texto etnográfico, guiei-me por questões como: quais os elementos que compõem a Ciência Aberta nesse espaço? De que modo esses elementos operam para que a Ciência Aberta possa funcionar e viabilizar o trabalho de alguém?

À medida que pude seguir e acompanhar as interações e associações, fazendo uso de uma série de recursos discursivos, tecnológicos e imagéticos, foi possível identificar os significados e formas de representação da Ciência Aberta no NeuroMat, seus limites e seus contornos. Foi possível observar, com atenção, as fronteiras e as divisões culturais que emergem em relação ao compartilhamento aberto de dados e ferramentas associadas à constituição da abordagem de Ciência Aberta. Por meio do enfoque analítico dos Estudos de Infraestrutura busquei também perscrutar as práticas de circulação do conhecimento científico que emergem com a Ciência Aberta, bem como os papéis e as formas de divisão do trabalho estabelecidas em torno dela.

O terceiro capítulo adentrou o trabalho de campo, descrevendo as configurações do NeuroMat enquanto centro de pesquisa. O capítulo relatou as origens do envolvimento do NeuroMat com o tema da Ciência Aberta e descreveu as formas de associação com diferentes frentes de atuação do projeto. O NeuroMat é um grande produtor e difusor de informações

diversas, o que, de certo modo, está na sua razão de ser expressa na missão enquanto “Centro de Pesquisa, Inovação, e Difusão”.

Nesse aspecto, a utilização de recursos digitais online, como sites, redes sociais, blogs, para citar alguns exemplos, adquirem grande centralidade nas ações voltadas à disseminação dessas informações. Destacam-se, aqui, os Eixos de Difusão Científica e Educação, nos quais são desenvolvidas ações de popularização da ciência; e o Eixo de Transferência de Conhecimento e Tecnologia, no qual ocorrem atividades que buscam transferir as realizações do NeuroMat para diferentes setores da sociedade.

Em ambas as frentes, ferramentas abertas – como softwares e licenças livres – estão presentes como parte dos produtos e dos processos desenvolvidos em cada eixo, muitas vezes estruturando suas atividades. Tais apropriações possibilitam ao NeuroMat assumir um lugar de organização comprometida com os princípios de Ciência Aberta e reivindicar autoridade nesse aspecto. Chamo atenção aqui para a atuação de um conjunto de profissionais e cientistas responsáveis por “montar” e executar uma complexa estratégia interna e externa de Ciência Aberta que se dá por meio do desenvolvimento de sistemas de informação, treinamentos de outros atores, como pesquisadores e equipe técnica. Nesse cenário, os cientistas da computação e desenvolvedores se destacam como principais detentores dos conhecimentos necessários à implementação das infraestruturas de informação (softwares, bases de dados) que operacionalizam as práticas de abertura.

A importância dos profissionais técnicos na operacionalização dessas atividades sobressai no Eixo de Difusão Científica, no qual são desenvolvidas atividades operacionalizadas por um grupo de profissionais formado por jornalistas, cientistas sociais e linguista. Além das habilidades comunicacionais em jornalismo científico e produção de texto, alguns dos seus integrantes se destacam por terem um importante histórico na participação de grupos envolvidos com o “conhecimento aberto” (por exemplo, o Grupo de Usuários Wikimedia no Brasil, responsável por divulgar as ferramentas colaborativas e abertas da Fundação Wikimedia). Essas conexões e aproximações têm se traduzido em reconhecimento externo do centro no âmbito da popularização da ciência por meio da Wikipédia. Esse reconhecimento, por sua vez, passou a refletir em uma relativa autonomia da equipe de difusão, que conquistou um espaço físico próprio e maior liberdade para desenvolver alguns projetos.

Um outro ponto que merece destaque diz respeito à relação das práticas de Ciência Aberta desenvolvidas no NeuroMat com seu projeto científico e institucional de implementação de um projeto científico em torno da modelagem neuronal de grande escala. Parte do projeto inclui o desenvolvimento de um banco de dados experimentais de acesso aberto para disponibilizar dados de pesquisa a fim de viabilizar processos de modelagem de dados neuronais. A Ciência Aberta, representada aqui nas ferramentas de compartilhamento de dados, aparece como uma forma de sustentar a colaboração entre neurocientistas e modeladores. O projeto do banco de dados de acesso aberto, no entanto, trouxe uma série de discussões internas e alguns atritos entre neurocientistas e modeladores em relação ao compartilhamento dos dados nesses termos. Ao acompanhar algumas etapas do projeto, tive a oportunidade de identificar alguns dos pressupostos, interesses e motivações que moldavam a forma como os pesquisadores naquele espaço se relacionavam com a abertura dos resultados de pesquisa, em especial dos dados experimentais. Esta análise está presente no quarto capítulo da tese, no qual busquei compreender especificamente as práticas de pesquisa e de que maneira os pesquisadores, no domínio da neurociência e da modelagem, se relacionam com o compartilhamento aberto dos resultados de pesquisa.

Atentando-me às características das práticas científicas de neurocientistas e modeladores, busquei identificar as manifestações das culturas epistêmicas, sobretudo em torno das divisões que se estabelecem em relação às formas experimentação, coleta de dados e possibilidades de tornar os dados publicamente disponíveis.

Sobre essas práticas, chamo a atenção para as diferenças na forma como neurocientistas abordam o compartilhamento de dados segundo os métodos utilizados, os tipos de experimentos e área de formação. Os neurocientistas da área clínica estabelecem posturas em relação ao compartilhamento de dados a partir do cuidado com a privacidade do paciente, manifestando questões éticas implicadas para o pesquisador e para o paciente que participa dos experimentos. O secretismo, nesse contexto, é constitutivo dessa comunidade como uma forma de resguardar os produtores dos dados e/ou o paciente.

Entre os neurocientistas experimentalistas, o secretismo está relacionado a uma questão meritocrática e de posse sobre os dados. Nessa relação, as motivações para a obtenção do crédito acadêmico e prioridade sobre os resultados da pesquisa orientam a forma pela qual os dados são “escondidos” ou revelados em pequenas porções conforme a conveniência. Os modeladores se relacionam de forma muito mais aberta à Ciência Aberta. Notou-se que as práticas desses cientistas, incluindo a própria fabricação de modelos, mobiliza uma série de

ferramentas abertas (como softwares), de modo que disciplinas que compõe o grupo de modeladores do NeuroMat estruturaram-se ao longo de sua história em torno do desenvolvimento de ferramentas de código aberto que viabilizam o trabalho colaborativo da modelagem.

Outro aspecto importante da observação foi que os modeladores produzem avanços no seu próprio campo a partir da aplicação das ferramentas matemáticas e computacionais nos dados produzidos no domínio da neurociência, o que se traduz em uma postura dos modeladores muito “proativa” e favorável à abertura dos dados, uma vez que a prática de compartilhar abertamente constitui para esses pesquisadores instrumento viabilizador do seu próprio trabalho.

Nesse sentido, os modeladores de dados experimentais colocam uma série de problematizações em relação aos dados produzidos pelos neurocientistas. Dentre elas são mencionadas a falta de padronização e qualidade dos dados experimentais e como isso impacta a capacidade dos experimentos serem recriados e também a produção de modelos de qualidade. Em seus discursos, os modeladores reforçam a necessidade dos neurocientistas disponibilizarem ferramentas de pesquisa e de descreverem protocolos experimentais e metadados associados aos dados de forma padronizada, pois disso depende a qualidade e viabilidade dos modelos.

Emergem daí questões relativas às infraestruturas de informação e o seu papel como viabilizadoras desses processos. Um aspecto abordado na tese, mais especificamente no quinto capítulo, diz respeito ao papel dos padrões como parte dos processos de compartilhamento aberto e reutilização dos dados de pesquisa. A padronização aqui tem significados e formas variadas. Pode significar um aspecto mais amplo no sentido de adoção generalizada de uma prática pelo grupo de pesquisadores, como a padronização de protocolos experimentais, mas também refere-se à forma de registro dos dados.

O foco no processo de padronização no NeuroMat, associado à infraestrutura ali desenvolvida, revelou a complexidade e a diversidade dos dados neuronais em termos de origem, forma de coleta e registro, o que se apresenta como um problema para os pesquisadores que pretendem reutilizá-los e compartilhá-los. Tais elementos foram analisados no quinto capítulo, no qual trago análises desenvolvidas fundamentadas pelo trabalho etnográfico acompanhando o desenvolvimento de uma ferramenta – o *Neuroscience Experiments System* – como parte do projeto do centro para implementar seu projeto de

banco de dados experimentais de acesso aberto. É atribuído ao software a função de gerir dados experimentais produzidos pelos laboratórios colaboradores do NeuroMat visando o compartilhamento de informações publicamente. Examinei, ao longo do capítulo, diferentes relações estabelecidas em torno do NES.

Uma noção importante colocada por alguns informantes diz respeito ao papel de metadados e protocolos em agregar “valor aos dados”, ou seja, como essa padronização confere, a um conjunto de registros, propriedades que os tornam utilizáveis na produção de conhecimento de uma área – no caso específico da disciplina analisada aqui, a neuromatemática, possibilita sua incorporação à produção dos modelos neuronais. As ferramentas desenvolvidas no domínio da modelagem, entre elas o NES, inserem-se em um processo de construção e de legitimação do conhecimento produzido nesse espaço. Nesse contexto, a padronização das infraestruturas reflete um propósito mais amplo de diluir diferenças, diminuir os atritos, validar e convergir processos em torno de objetivos e questões de pesquisa comuns.

O NES está inserido em um conjunto diverso de relações que o faz ser interpretado de diferentes maneiras. Para alguns pesquisadores, a ferramenta é apenas um software de trabalho ou um viabilizador da neuromatemática; para outros, representa a possibilidade de descobrimento de tratamentos mais eficazes para pessoas com Lesão do Plexo Braquial. Ou ainda, para instituições financiadoras de pesquisa, como a FAPESP, trata-se de um “produto tecnológico” e, como tal, ele pode ser “transferido” para a sociedade por meio do licenciamento, permitindo ser utilizado para fins comerciais. Em muitos momentos, sugeriu-se a existência de uma incompatibilidade entre uma abordagem mais voltada para o fechamento que permeia certas práticas da FAPESP e a abordagem de Ciência Aberta adotada para transferência de conhecimento científico e de tecnologias para a sociedade.

É importante reiterar que as ferramentas abertas são centrais para nas práticas desenvolvidas no NeuroMat e estão incluídas em processos de legitimação e organização do campo de pesquisa da neuromatemática como disciplina. Além das ferramentas de software desenvolvidas para dar suporte à gestão e o compartilhamento de dados, destacam-se também apropriações das ferramenta *wiki* mantidas pela Fundação Wikimedia em atividades colaborativas e de difusão de informações. Por meio da inserção ou melhoria de termos próprios ou relevantes para a neuromatemática, as atividades de edição da Wikipédia passaram a ocorrer como uma forma de registro da neuromatemática nesse espaço de grande visibilidade na Internet.

Os achados empíricos possibilitam interpretar o uso dessas ferramentas como uma forma particular do NeuroMat desenvolver Ciência Aberta. Nesses processos, essas ferramentas – como as ferramentas colaborativas da Wikipédia – participam de um processo de definição de fronteiras, terminologia, conteúdo e como meio de circulação do conhecimento da neuromatemática. É igualmente interessante observar que, em muitos momentos, há um esforço dos pesquisadores principais em estabelecerem a organização da disciplina da neuromatemática em torno desses recursos abertos e estimular as práticas de compartilhamento entre os pesquisadores mais jovens. Busca-se, por meio dessas atividades, treinar os jovens pesquisadores e formar uma geração de pesquisadores aderentes a essas práticas de compartilhamento. Esse segundo ponto se manifestou, principalmente, em situações em que as ferramentas e elementos associados à Ciência Aberta eram inseridos em processos pedagógicos voltados a pesquisadores em estágio de formação.

Em muitos momentos, os papéis e configurações da Ciência Aberta no espaço do NeuroMat se confundem com as questões e problemas científicos próprios que perpassam as atividades dos pesquisadores, de modo que os desafios que são colocados para essa Ciência Aberta não se separam daqueles que se apresentam para a neuromatemática enquanto disciplina científica que pretende estudar os fenômenos do sistema neuronal.

A neuromatemática, como uma nova disciplina, propõe a integração das teorias sobre o funcionamento do cérebro a partir das questões científicas comuns e do desenvolvimento de uma abordagem teórica capaz de explicar a pluralidade de fenômenos observados no cérebro e de relacionar suas manifestações em diferentes escalas da atividade neuronal. Notou-se que essa “integração” passa não apenas pela aplicação dos instrumentos matemáticos e computacionais para tratar o grande volume de dados experimentais produzidos pela neurociência, como está imbricada com ferramentas e com o compartilhamento aberto de dados entre disciplinas e laboratórios em um circuito de compartilhamento de informações e de colaboração mais amplo.

Nesse aspecto, a Ciência Aberta no NeuroMat – formada por ferramentas, compartilhamento de dados, padrões diversos, práticas, normas e pessoas – funciona como substrato dessa integração pretendida, sendo acionada com a função de reduzir a fragmentação e atritos entre diferentes domínios. Desse modo, viabiliza que as disciplinas, teorias, pesquisadores e infraestruturas, códigos possam cada vez mais convergir em torno da produção de modelos de “qualidade”, “representativos” ou “realísticos” – para usar alguns dos adjetivos usados pelos pesquisadores.

Nesse sentido, concluímos que a Ciência Aberta, no contexto das práticas desse centro, configura-se como um dos muitos componentes associados da infraestrutura que dá suporte ao trabalho da modelagem e outros processos desenvolvidos no NeuroMat direcionados à produção de evidências sobre o funcionamento do cérebro humano, participando da constituição da neuromatemática como a “nova ciência do cérebro”, com suas especificidades e controvérsias específicas.

REFERÊNCIAS BIBLIOGRÁFICAS

ALBAGLI, Sarita. Ciência Aberta em questão. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud. (Orgs.). **Ciência Aberta, Questões Abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015.

ALVES, David; SANTOS, Jean Carlos Ferreira dos; PESCHANSKI, João Alexandre. Tecnologias colaborativas na difusão científica: um relato dos usos e apropriações da Wikipédia no CEPID Neuromat. In: **IX Simpósio Nacional da Associação Brasileira de Pesquisadores em Cibercultura**, Dez. 2016. Disponível em: <http://abciber2016.com/wpcontent/uploads/2016/trabalhos/tecnologias_colaborativas_na_difusao_cientifica_um_retrato_dos_usos_e_apropriacoes_da_wikipedia_no_cepid_neuromat_david_fernando_levon_alves.pdf>. Acesso em: 20 mar. 2017.

ARAÚJO, Daniela de. **Feminismo e cultura hacker**: intersecções entre política, gênero e tecnologia. 2018. 147f. Tese (doutorado em Política Científica e Tecnológica) – Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2018.

ASCOLI, Giorgio A. Mobilizing the base of neuroscience data: the case of neuronal morphologies. In: **Nature Reviews Neuroscience**, v. 7, n. 4, p. 318–324, abr. 2006.

BAKER, Karen S.; MILLERAND, Florence. Infrastructuring ecology: Challenges in achieving data sharing. **Collaboration in the new life sciences**, p. 111-138, 2010.

BEAR, Mark F.; CONNORS, Barry W.; PARADISO, Michael A. **Neurociências**: Desvendando o Sistema Nervoso. Porto Alegre: Artmed, 2017. 2ª ed.

BEAULIEU, Anne. From brainbank to database: the informational turn in the study of the brain. In: **Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences**, v. 35, n. 2, p. 367–390, jun. 2004.

_____. Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity. In: **Social Studies of Science**, v. 31, n. 5, p. 635–680, 1 out. 2001. Disponível em: <<https://annebeaulieu.files.wordpress.com/2011/04/social-studies-of-science-2001-beaulieu-635-802.pdf>>. Acesso em: 27 mai. 2015.

BILL & MELINDA GATES FOUNDATION. **Open Access Policy**. 2011 Disponível em: <<https://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy>>. Acesso em: 26 fev. 2018.

BOLLIER, David. **Pensar desde los comunes: una breve introducción**. Traficantes de Sueños, 2016.

BORGMAN, Christine L. The conundrum of sharing research data. In: **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/doi/10.1002/asi.22634/epdf>>. Acesso em: 13 jul. 2015.

_____. **Big data, little data, no data: scholarship in the networked world**. Cambridge, Massachusetts: The MIT Press, 2015.

_____. **From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World**. Cambridge, Mass: The MIT Press, 2000.

BORGMAN, C. L. EDWARDS, Paul N.; JACKSON, Steven J.; CHALMERS, Steven J. et al. **Knowledge Infrastructures: Intellectual Frameworks and Research Challenges**, 2013. Disponível: <<https://cloudfront.escholarship.org/dist/prd/content/qt2mt6j2mh/qt2mt6j2mh.pdf?t=p7ctt1>>. Acesso em: 09 jul. 2018.

BOULTON, Geoffrey et al. Science as an open enterprise. Londre: In: **The Royal Society**, 2012. Disponível em: <https://royalsociety.org/~media/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf>. Acesso em: 13 fev. 2016.

BOWKER, Geoffrey C; Baker, Karen; MILLERAND, Florence; RIBES, David. Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In: HUNSINGER, J.; KLASTRUP, L.; ALLEN, M. (Orgs.). **International Handbook of Internet Research**. Dordrecht: Springer Netherlands, 2010.

BOWKER, Geoffrey C; STAR, Susan Leigh. **Sorting things out: Classification and its consequences**. MIT press, 2000.

BOWKER, G. C. **Memory Practices in the Sciences**. Cambridge, Mass: MIT Press, 2005.

BRAZILIAN INITIATIVE ON PRECISION MEDICINE. **Home page**. Disponível em: <<http://bipmed.org/>>. Acesso em 20 mar. 2018.

BRUYNINCKX, Joeri. Trading twitter: Amateur recorders and economies of scientific exchange at the Cornell Library of Natural Sounds. In: **Social Studies of Science**, v. 45, n. 3, p. 344–370.

BULLMORE, Ed; SPORNS, Olaf. Complex brain networks: graph theoretical analysis of structural and functional systems. In: **Nature Reviews Neuroscience**, v. 10, n. 3, p. 186–198, mar. 2009.

CARDOSO, Gustavo; JACOBETTY, Pedro; DUARTE, Alexandra. **Para uma ciência aberta**. Lisboa: Editora Mundos Sociais, 2012.

CAULFIELD, Timothy; HARMON, Shawn H.; JOLY, Yann. Open science versus commercialization: a modern research conflict? In: **Genome Medicine**, v. 4, p. 17, 27 fev. 2012. Disponível em:

<<https://genomemedicine.biomedcentral.com/track/pdf/10.1186/gm316?site=genomemedicine.biomedcentral.com>>. Acesso em: 02 mar. 2017.

CCSL. **Home page**. Disponível em: <<http://ccsl.ime.usp.br/>>. Acesso em 20 mar. 2018.

CHAVES, Gabriela Costa; OLIVEIRA, Maria Auxiliadora. Direitos de Propriedade intelectual e acesso a medicamentos. In: REIS, Renata; TERTO JR, Veriano, PIMENTA, Cristina; MELHO, Fátima. **Propriedade Intelectual: Interfaces e Desafios**. Rio de Janeiro, RJ, Brasil: ABIA, 2009.

CHUBIN, Daryl E. Open Science and Closed Science: Tradeoffs in a Democracy. **Science, Technology, & Human Values**, v. 10, n. 2, p. 73–80, 1 abr. 1985.

CLINIO, Anne. Por que open notebook science? Uma aproximação às ideias de Jean-Claude Bradley. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud. (Orgs.). **Ciência Aberta, Questões Abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015.

COLEMAN, Gabriella. Revoluções Silenciosas: o irônico surgimento do software livre e de código aberto e a constituição de uma consciência legal hacker. In: COLEMAN, Gabriella. **Do regime de propriedade intelectual: estudos antropológicos**. Porto Alegre: Tomo Editorial, 2010.

COMISSÃO EUROPEIA. **Data management - H2020 Online Manual**. 2013 Disponível em: <http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm>. Acesso em: 26 mai. 2014.

_____. **Melhorar o acesso à informação científica:** rentabilizar o investimento público em investigação. 2012. Disponível em: <http://ec.europa.eu/transparency/regdoc/rep/1/2012/PT/1-2012-401-PT-F1-1.Pdf>. Acesso em: 26 mai. 2014.

COSTA, Sely. M. S. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso livre à informação científica. In: **Ciência da Informação**, Rio de Janeiro, v. 35, n. 2, p. 39-50, mai./ago. 2006.

CURTY, Renata. G.; AVENTURIER, Pascal. **O paradigma da publicação de dados e suas diferentes abordagens.** XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO ENANCIB 2017 GT-7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação. **Anais...** Marília, Brasil: out. 2017. Disponível em: <<https://hal.archives-ouvertes.fr/hal-01637987>>. Acesso em: 15 jan. 2019.

DAHLANDER, Linus; GANN, David. M. How open is innovation? In: **Research Policy**, v. 39, n. 6, p. 699–709, jul. 2010.

DAVID, Paul A. Common agency contracting and the emergence of "open science" institutions. In: **The American Economic Review**, v. 88, n. 2, p. 15-21, 1998.

_____. From keeping 'nature's secrets' to the institutionalization of 'open science'. In: GHOSH, Rishab Aiyer A. (Org.). **CODE: collaborative ownership and the digital economy.** Cambridge, Mass: MIT Press, 2005.

_____. The Historical Origins of 'Open Science': an essay on patronage, reputation and common agency contracting in the scientific revolution. In: **Capitalism and Society**, v. 3, n. 2, 2008.

_____. A. Understanding the emergence of "open science" institutions: functionalist economics in historical context. In: **Industrial and Corporate Change**, v. 13, n. 4, p. 571–589, 1 ago. 2004.

DELFANTI, Alessandro. **Biohackers: The Politics of Open Science.** Londres: Pluto Press, 2013.

_____. Beams of particles and papers. How digital preprint archives shape authorship and credit. In: **arXiv:1602.08539 [physics]**, 26 fev. 2016. Disponível em: <https://arXiv.org/pdf/1602.08539.pdf>. Acesso em: fev. 2016.

DELFANTI; PITRELLI, Nico. Ciência aberta: revolução ou continuidade? In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud. (Orgs.). **Ciência Aberta, Questões Abertas.** Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015.

EDWARDS, Paul. Knowledge infrastructures for the Anthropocene. In: **The Anthropocene Review**, v. 4, n. 1, p. 34–43, 1 abr. 2017.

EDWARDS, Paul; Bowker, Geoffrey; Jackson, Steven; Williams, Robin. Introduction: An Agenda for Infrastructure Studies. **Journal of the Association for Information Systems**, v. 10, n. 5, 28 maio 2009. Disponível em: <<http://aisel.aisnet.org/jais/vol10/iss5/6>>. Acesso em: 08 jan. 2018.

EDWARDS, Paul; MAYERNIK, Matthew S. et al. Science friction: Data, metadata, and collaboration. In: **Social Studies of Science**, v. 41, n. 5, p. 667–690, 1 out. 2011. Disponível em: <<https://cloudfront.escholarship.org/dist/prd/content/qt2tj4g7sc/qt2tj4g7sc.pdf>>. Acesso em: 06 nov. 2017.

EOSC. European Open Science Cloud. **Home Page**. Disponível em: <<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>>. Acesso em: 18 jan. 2019.

EVANGELISTA, Rafael de Almeida. Traidores do movimento: política, cultura, ideologia e trabalho no Software Livre. 2010. 250f. Tese (doutorado em Antropologia Social) - Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas. 2010.

FAPESP. Fundação de Amparo à Pesquisa do Estado de São Paulo. **Centro de Pesquisa, Inovação e Difusão (CEPID)**. 2018. Disponível em: <<http://cepid.fapesp.br/>> . Acesso em: 25 fev. 2018.

_____. **CEPID - Edital CEPID 2011**. Disponível em: <<http://www.fapesp.br/6335>> . Acesso em: 25 fev. 2018.

_____. **Relatório de atividades 2016 – versão executiva**, 2017. Disponível em: <<http://www.fapesp.br/publicacoes/relat2016.pdf>>. Acesso em: 27 dez. 2017.

FAUSTO, Sibeles de. **Polêmica PLOS expõe intolerância ao acesso aberto**. SIBiUSP - Sistema Integrado de Bibliotecas da Universidade de São Paulo., 2016. Disponível em: <<http://www.sibi.usp.br/noticias/polemica-plos-aconteceu/>>. Acesso em: 27 mai. 2017.

FECHER, Benedikt; FRIESIKE, Sascha. Open science: One term, five schools of thought. In: BARTLING, Sönke; FRIESIKE, Sascha. (Orgs.). **Opening Science: the evolving guide on how the internet is changing research, collaboration and scholarly publishing**. Heidelberg: Springer Open, 2014.

- FORD, Emily. Open peer review at four. Stem journals: an observational overview. In: **F1000Research**, 9 jan. 2015. Disponível em: <<https://f1000research.com/articles/4-6/v1>>. Acesso em: 26 jun. 2017.
- FOSTER. **Home Page**. Disponível em: <<https://www.fosteropenscience.eu/>>. Acesso em: 24 jan. 2019.
- FRANZONI, Chiara; SAUERMAN, Henry. Crowd science: The organization of scientific research in open collaborative projects. In: **Research Policy**, v. 43, n. 1, p. 1–20, 2014. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0048733313001212>>. Acesso em: 2 mar. 2017.
- FRIESIKE, Sascha et al. Opening science: towards an agenda of open science in academia and industry. In: **The Journal of Technology Transfer**, v. 40, n. 4, p. 581–601, 1 ago. 2015. Disponível em: <<https://link.springer.com/content/pdf/10.1007%2Fs10961-014-9375-6.pdf>> Acesso em: 3 nov. 2015.
- FRONTIERS. **Home page**. Disponível em: <<https://www.frontiersin.org/>>. Acesso em 20 mar. 2018.
- GALVES, Antônio; LÖCHERBACH, Eva. Infinite Systems of Interacting Chains with Memory of Variable Length—A Stochastic Model for Biological Neural Nets. In: **Journal of Statistical Physics**, v. 151, n. 5, p. 896–921, 1 jun. 2013. Disponível em: <<https://link.springer.com/article/10.1007/s10955-013-0733-9>>. Acesso em: 10 out. 2017.
- GARDNER, Daniel et al. The Neuroscience Information Framework: A Data and Knowledge Environment for Neuroscience. In: **Neuroinformatics**, v. 6, n. 3, p. 149–160, 1 set. 2008. Disponível em: <<https://link.springer.com/content/pdf/10.1007%2Fs12021-008-9024-z.pdf>>. Acesso em: 27 dez. 2017.
- GEIGER, R. Stuart.; RIBES, David. Trace Ethnography: Following Coordination through Documentary Practices, 44th. In: **Anais...** Hawaii International Conference on System Sciences. Jan. 2011.
- GIBBONS, Michael et al. **The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies**. Newbury Park : Sage, 1994.
- GLEESON, Pdraig. et al. NeuroML: A Language for Describing Data Driven Models of Neurons and Networks with a High Degree of Biological Detail. In: **PLOS Computational Biology**, v. 6, n. 6, p. e1000815, 17 jun. 2010.

GONZALO RAY, Hernan; PEDREIRA, Carlos; QUIROGA, Rodrigo. Past, present and future of spike sorting techniques. In: **Brain Research Bulletin**, Advances in electrophysiological data analysis. v. 119, n. Part B, p. 106–117, 1 out. 2015. Disponível em:

<<https://doi.org/10.1016/j.brainresbull.2015.04.007>>. Acesso em: 30 dez. 2017.

HALCHENKO, Yaroslav O.; HANKE, Michael. Open is Not Enough. Let's Take the Next Step: An Integrated, Community-Driven Computing Platform for Neuroscience. **Frontiers in Neuroinformatics**, v. 6, 29 jun. 2012.

HAGSTROM, Warren O. **The scientific community**. Carbondale: Southern Illinois: University Press, 1965.

HILGARTNER, Stephen; BRANDT-RAUF, Sherry I. Data Access, Ownership, and Control Toward Empirical Studies of Access Practices. In: **Science Communication**, v. 15, n. 4, p. 355–372, 6 jan. 1994.

_____. **Access to data and intellectual property: Scientific exchange in genome research. Intellectual property rights and research tools in molecular biology**, 1997. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK233532/>>. Acesso em: 6 jan. 2015.

_____. Intellectual property and the politics of emerging technology: inventors, citizens, and powers to shape the future. In: **Chicago-Kent Law Review**, v. 84, p. 197, 2009.

_____. Selective flows of knowledge in technoscientific interaction: information control in genome research. In: **The British Journal for the History of Science**, v. 45, n. 2, p. 267–280, jun. 2012.

HINE, Christine. **Virtual ethnography**. London: SAGE, 2000.

HINE, Chistine. Databases as scientific instruments and their role in the ordering of scientific work. In: **Social Studies of Science**, v. 36, n. 2, p. 269–298, 1 abr. 2006.

_____. **From Virtual Ethnography to the Embedded, Embodied, Everyday Internet**. [s.l.] Routledge Handbooks Online, 2016.

_____. Multi-sited ethnography as a middle range methodology for contemporary STS. In: **Science, Technology, & Human Values**, v. 32, n. 6, p. 652-671, 2007.

_____. **Systematics as cyberscience: computers, change, and continuity in science**. Cambridge: MIT Press, 2008.

- HOFSTAD, Remco Van Der. **Random graphs and complex networks**. Vol. I. Disponível em: <<https://www.win.tue.nl/~rhofstad/NotesRGCN.pdf>>. Acesso em: 27 set. 2017.
- HOPE, Janet. **Biobazaar: the open source revolution and biotechnology**. Cambridge, Mass: Harvard University Press, 2008.
- HUGHES, Thomas Parke. The evolution of large technological systems. In: BIJKER, Wiebe E.; HUGHES, Thomas Parke; PINCH, Trevor (Orgs.). **The Social construction of technological systems: new directions in the sociology and history of technology**. Cambridge, Mass: MIT Press, 1987. p. 45–76.
- IZIQUÉ, Claudia. **FAPESP anuncia 17 novos CEPIDs com investimentos de US\$ 680 milhões, AGÊNCIA FAPESP**. Disponível em: <http://agencia.fapesp.br/fapesp_anuncia_17_novos_cepids_com_investimentos_de_us_680_milhoes/17273/>. Acesso em: 10 abr. 2015.
- JOHNSON, Gareth J. Cultural, ideological and practical barriers to open access adoption within the UK Academy: an ethnographically framed examination. **Insights**, v. 31, n. 0, 11 jun. 2018.
- JULIÃO, Andre. **Abertura de dados de pesquisa torna uso de recursos mais eficiente, AGÊNCIA FAPESP**. Disponível em: <<http://agencia.fapesp.br/abertura-de-dados-de-pesquisa-torna-uso-de-recursos-mais-eficiente/28799/>>. Acesso em: 30 jan. 2019.
- KANFER, Alaina G. et al. Modeling distributed knowledge processes in next generation multidisciplinary alliances. In: **Information Systems Frontiers**, v. 2, n. 3-4, p. 317-331, 2000.
- KNORR-CETINA, Karin. **Epistemic cultures: how the sciences make knowledge**. Cambridge, Mass: Harvard University Press, 1999a.
- KNORR-CETINA, Karin. A comunicação na ciência. In: GIL, Fernando (Org.). **A ciência tal qual se faz**. Tradução de Paulo Tunhas. Lisboa: Edições João Sá da Costa, 1999b. p.375-393.
- KOULOCHERI, E.; KOULOCHERI, E. **What is the EC Open Research Data Pilot?, 2018**. Disponível em: <<https://www.openaire.eu/what-is-the-open-research-data-pilot>>. Acesso em: 22 jan. 2019.
- KOWALCZYK, Stacy; SHANKAR, Kalpana. Data sharing in the sciences. **Annual Review of Information Science and Technology**, v. 45, n. 1, p. 247–294, 2013.

KULCZYCKI, Emanuel. Rethinking open science: the role of communication. In: **Analele Universitatii din Craiova, Seria Filosofie**, v. 37, n. 1, p. 81–97, 2016.

LAFUENTE, Antonio et al. Los cuatro entornos del procomún. Archipiélago. In: **Cuadernos de crítica de la cultura**, n. 77-78, p. 15-22, 2007.

LAGOZE, Carl et. al. Should I Stay or Should I Go? Alternative Infrastructures in Scholarly Publishing. In: **International Journal of Communication**, v. 9, n. 0, p. 20, 30 mar. 2015. Disponível em: <<http://ijoc.org/index.php/ijoc/article/view/2929>>. Acesso em: 26 jul. 2017.

LATOURE, Bruno; WOOLGAR, Steve. **A vida de laboratório a produção dos fatos científicos**. Rio de Janeiro: Relume Dumara, 1997.

LATOURE, Bruno. **Reagregando o social**: uma introdução à teoria do ator-rede. Baurú: Edusc; Salvador: Edufba, 2012.

LEONELLI, Sabina. **Data-Centric Biology**: a philosophical study. Chicago; London: University of Chicago Press, 2016.

_____. Why the Current Insistence on Open Access to Scientific Data? Big Data, Knowledge Production, and the Political Economy of Contemporary Biology. In: **Bulletin of Science, Technology & Society**, v. 33, n. 1–2, p. 6–11, 1 fev. 2013. Disponível em:

<<http://journals.sagepub.com/doi/pdf/10.1177/0270467613496768>>. Acesso em: 15 set. 2014.

LESSIG, Lawrence. **Cultura livre**: como a grande mídia usa a tecnologia e a lei para bloquear a cultura e controlar a criatividade. São Paulo: Trama, 2005.

LEVIN, Nadine; LEONELLI, Sabina; WECKOWSKA, Dagmara et al. How Do Scientists Define Openness? Exploring the Relationship Between Open Science Policies and Research Practice. In: **Bulletin of Science, Technology & Society**, v. 36, n. 2, p. 128–141, 1 jun. 2016. Disponível em: <<http://journals.sagepub.com/doi/pdf/10.1177/0270467616668760>>. Acesso em: 20 jun. 2016.

LEVIN, Nadine; LEONELLI, Sabina. How Does One “Open” Science? Questions of Value in Biological Research. **Science, Technology & Human Values**, p. 0162243916672071, 3 out. 2016.

_____; LEONELLI, Sabina. How Does One “Open” Science? Questions of Value in Biological Research. In: **Science, Technology & Human Values**, p. 0162243916672071, 3 out. 2016. Disponível em: <<http://journals.sagepub.com/doi/pdf/10.1177/0162243916672071>>. Acesso em: 20 out. 2016.

LIMA, Bia Ramalho dos Santos. **Investigação da sensibilidade superficial e sensação referida em participantes com lesão e reconstrução cirúrgica do plexo braquial**. 2015. 108 f. Dissertação (Mestrado em Ciências Biológicas) – Instituto de Biofísica Carlos Chagas Filho da UFRJ, Rio de Janeiro, 2015.

LOCKREM, Jessica; LUGO, Adonia. Infrastructure [Editorial]. In: **Cultural Anthropology**, 2012. Disponível em: <https://culanth.org/curated_collections/11-infrastructure>. Acesso em: 31 jan. 2018.

LÓSCIO, Bernadette Farias; BURLE, Caroline; OLIVEIRA, Marcelo lury S.; CALEGARI, Newton. **Fundamentos para publicação de dados na web**. São Paulo: Comitê Gestor da Internet no Brasil, 2018.

LYNCH, Michael. **Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science**. Cambridge: Cambridge University Press, 1993.

LYTTON, William. W. Computer modeling of Epilepsy. In: **Nature reviews - Neuroscience**, v. 9, n. 8, p. 626–637, ago. 2008. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2739976/pdf/nihms-88659.pdf>>. Acesso em: 27 out. 2017.

MARCUS, George. E. Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography. In: **Annual Review of Anthropology**, v. 24, n. 1, p. 95–117, 1995.

MAURER, Stephen M. New institutions for doing science: From databases to open source biology. European Policy for Intellectual Property Conference, Maastricht. In: **Anais...2003**. Disponível em: <https://www.researchgate.net/profile/Stephen_Maurer/publication/237412391_New_Institutions_for_Doing_Science_From_Databases_to_Open_Source_Biology/links/54343a3f0cf2bf1f1f27bcb6.pdf>. Acesso em: 2 mar. 2017.

MERVIS, Jeffrey. Why null results rarely see the light of day. In: **Science**, v. 345, n. 6200, p. 992–992, 29 ago. 2014.

MERTON, Robert K. A note on science and democracy. **J. Legal & Pol. Soc.**, v. 1, p. 115, 1942.

_____. Os imperativos institucionais da ciência. In: DEUS, J. D. A crítica da ciência: sociologia e ideologia da ciência (Org.). Rio de Janeiro: Zahar Editores, 1979. p.37-52.

- MILLERAND, Florence. La science en réseau: es gestionnaires d'information "invisibles" dans la production d'une base de données scientifiques. In: **Revue d'anthropologie des connaissances**, v. 6, n° 1, n. 1, p. 163–190, 25 abr. 2012. Disponível em: <<https://www.cairn.info/revue-anthropologie-des-connaissances-2012-1-page-163.html>>. Acesso em: ago. 2015.
- _____. La science en réseau: les gestionnaires d'information « invisibles » dans la production d'une base de données scientifiques. In: **Revue d'anthropologie des connaissances**, v. 6, n° 1, n. 1, p. 163–190, 25 abr. 2012. Disponível em: <<https://www.cairn.info/revue-anthropologie-des-connaissances-2012-1-page-163.htm>>. Acesso em: 12 nov. 2015.
- _____. Le partage des données scientifiques à l'ère de l'e-science : l'instrumentation des pratiques au sein d'un collectif multidisciplinaire. In: **Terrains & travaux**, n. 18, p. 215–237, 18 ago. 2011. Disponível em: <<https://www.cairn.info/revue-terrains-et-travaux-2011-1-page-215.htm#no2>>. Acesso em: 24 jan. 2018.
- MOLLOY, Jennifer C. The open knowledge foundation: open data means better science. **PLoS Biol**, v. 9, n. 12, p. e1001195, 2011. Disponível: <<http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001195>>. Acesso em: 02 mar. 2017.
- MONTEIRO, Marko Synésio Alves. Reconsiderando a etnografia da ciência e da tecnologia: tecnociência na prática. In: **Revista Brasileira de Ciências Sociais**, v. 27, n. 79, p. 139–151, jun. 2012. Disponível em: <<http://www.scielo.br/pdf/rbcsoc/v27n79/a09.pdf>>. Acesso em: 11 jan. 2015.
- MORENO, Fernanda Passini. Repositórios de dados de pesquisa na Espanha: breve análise. In: **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 23, n. 53, p. 52–63, 6 set. 2018. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p52/37376>>. Acesso em: 12 set. 2018.
- MUNAFÒ, Marcus R.; NOSEK, Brian A.; BISHOP, Dorothy V. M. et al. A manifesto for reproducible science. **Nature Human Behaviour**, v. 1, n. 1, p. s41562- 016- 0021–016, 10 jan. 2017. Disponível em: <<https://www.nature.com/articles/s41562-016-0021>>. Acesso em: 29 jul. 2017.
- NATIONAL INSTITUTES OF HEALTH (NIH). **Final NIH statement on sharing research data**, 2003. Disponível em: <<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>>. Acesso em 20 mar. 2018.

NEUROMAT. Centro de Pesquisa, Inovação e Difusão em Neuromatemática. **CEPID**. NeuroMat Second Report of Activities: July 2014 - July 2015. Disponível em: http://neuromat.numec.prp.usp.br/sites/default/files/reports/second_report_of_activities.pdf > Acesso em: 7 ago. 2017.

_____. **RIDC**. NeuroMat Fourth Report of Activities. Jun 30, 2016 - Jun 30, 2017. Disponível em: http://neuromat.numec.prp.usp.br/sites/default/files/reports/fourth_report_of_activities.pdf . Acesso em: 07 set. 2017.

_____. **Scientific project submitted in 2012**. Disponível em: <http://neuromat.numec.prp.usp.br/pt-br/content/scientific-project-submitted-2012> > . Acesso em: 26 fev. 2015.

_____. **NeuroMat launches Wikipedia initiative**. Disponível em: <http://neuromat.numec.prp.usp.br/content/neuromat-launches-wikipedia-initiative> >. Acesso em: 30 mar. 2018.

_____. **NES. Neuroscience Experiments System**. 2017. Disponível em: <https://github.com/neuromat/nes> >. Acesso em: 25 jan. 2018.

NEUROMORPHO.ORG. **Home page**. Disponível em: <http://neuromorpho.org/byRandom.jsp> > Acesso em 20 mar. 2018.

NEUROSCIENCE INFORMATION FRAMEWORK. **Home page**. Disponível em <https://neuinfo.org/> >. Acesso em 20 mar. 2018.

NOOR, Aisyah Mohd. et al. Big Data: the challenge for small research groups in the era of cancer genomics. **British Journal of Cancer**, v. 113, n. 10, p. 1405–1412, nov. 2015.

NIELSEN, M. A. **Reinventing discovery: the new era of networked science**. Princeton, N.J: Princeton University Press, 2012.

OCDE. Organização para a Cooperação e Desenvolvimento Econômico. **OECD principles and guidelines for access to research data from public funding**. 2007. Disponível em: <https://www.oecd.org/sti/sci-tech/38500813.pdf> >. Acesso em: 13 jul. 2014.

OCDE. Organização para a Cooperação e Desenvolvimento Econômico. **OECD Science, Technology and Innovation Outlook 2016**. Paris: Organisation for Economic Co-operation and Development, 2016.

- OLIVEIRA, Rodrigo Maia de. **Proteção e comercialização da pesquisa acadêmica no Brasil:** motivações e percepções dos inventores, 2011. 167 f. [Tese de Doutorado em Política Científica e Tecnológica]. Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2011.
- PILAT, Dirk; FUKASAKU, Yukiko. OECD Principles and Guidelines for Access to Research Data from Public Funding. In: **Data Science Journal**, v. 6, p. OD4-OD11, 2007. Disponível em: <https://www.jstage.jst.go.jp/article/dsj/6/0/6_0_OD4/pdf>. 12 ago. 2017.
- PLANTIN, Jean-Christophe; LAGOZE, Carl; EDWARDS, Paul N. Re-integrating scholarly infrastructure: The ambiguous role of data sharing platforms. **Big Data & Society**, v. 5, n. 1, p. 2053951718756683, 1 jun. 2018.
- POUZAT, Christophe. **Spike Sorting**. NeuroMat, 2014. Disponível em: <<https://www.youtube.com/watch?v=vSydfDvsewY>> . Acesso em: 12 out. 2017.
- PREMEBIDA, Adriano; NEVES, Fabrício Monteiro; ALMEIDA, Jalcione. Estudos sociais em ciência e tecnologia e suas distintas abordagens. In: **Sociologias**, v. 13, n. 26, 2011. Disponível em: <<http://www.redalyc.org/html/868/86819459003/>>. Acesso em: 24. mar. 2017.
- QUIROGA, Rodrigo Q. Spike sorting. In: **Scholarpedia**, v. 2, n. 12, p. 3583, 21 dez. 2007. Disponível em: <http://www.scholarpedia.org/article/Spike_sorting>. Acesso em: 30 dez. 2017.
- RAYMOND, Eric. S. **The Cathedral & the Bazaar:** Musings on Linux and Open Source by an Accidental Revolutionary. Sebastopol: O'Reilly Media, 1999.
- RE3DATA.ORG. **Home Page**. Disponível em: <https://www.re3data.org/>. Acesso em: 19 jan. 2019.
- RESEARCH DATA SHARING WITHOUT BARRIERS. **Home Page**. Disponível em: <<https://www.rd-alliance.org/>>. Acesso em: 22 jan. 2019.
- RIBEIRO, Fernanda Teixeira. Modelos matemáticos do cérebro. In: **Mente & Cérebro**, Duetto Editorial., 2014.

- RIBES, David. JACKSON, Steven; GEIGER, Stuart et al. Artifacts that organize: Delegation in the distributed organization. **Information and Organization**, v. 23, n. 1, p. 1–14, 1 jan. 2013.
- RIBES, D.; LEE, C. P. Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories. **Computer Supported Cooperative Work (CSCW)**, v. 19, n. 3, p. 231–244, 1 ago. 2010.
- ROSS-HELLAUER, Tony. **Defining Open Peer Review: Part Two – Seven Traits of OPR** : OpenAIRE blogOpenAIRE Blog, 2 nov. 2016. Disponível em: <<https://blogs.openaire.eu/?p=1410>>. Acesso em: 26 jun. 2017.
- ROYAL SOCIETY (GREAT BRITAIN); POLICY STUDIES UNIT. **Science as an open enterprise**, 2012. Disponível em: <https://royalsociety.org/~media/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf>. Acesso em: 13 fev. 2018.
- ROTH, Wolff-Michael; BOWEN, G. Michael. Of disciplined minds and disciplined bodies: On becoming an ecologist. **Qualitative Sociology**, v. 24, n. 4, p. 459-481, 2001.
- SANTOS, Jean Carlos Ferreira dos. **Estudo sobre o movimento Open Access e de suas implicações para a comunicação na ciência**. 2014. 141 f. Dissertação (Mestrado em Política Científica e Tecnológica) -- Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2014.
- SANTOS, Paula Xavier dos et al. **Livro Verde Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional**. Rio de Janeiro: Fiocruz, 2017. Disponível em: <<https://www.arca.fiocruz.br/handle/icict/24117>>. Acesso em: 07 jul. 2018.
- SENGUPTA, Pratim; SHANAHAN, Marie-Claire. Open science, public engagement and the university. In: **arXiv:1702.04855 [physics]**, 15 fev. 2017. Disponível em: <<https://arXiv.org/pdf/1702.04855.pdf>>. Acesso em: 3 jun. 2016.
- SILVEIRA, Sérgio Amadeu da. Entre Software e Genes: A resistência ao paradigma do conhecimento . In: **Liinc em Revista**, v. 10, n. 2, 2014a. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/view/753>>. Acesso em 10 nov. 2015.
- _____. Coletivos tecnológicos e a produção colaborativa entre pares. MARINHO, Maria Gabriela; SILVEIRA, Sérgio Amadeu da; MONTEIRO, Marko; DIAS, Rafael de Brito; CAMPOS, Cristian de. (Orgs.). **Abordagens em ciência, tecnologia e sociedade**. Santo André: Universidade Federal do ABC, 2014b. p. 251-262.

SisNe. **Laboratório de Sistemas Neurais**. 2008. Disponível em: <<http://sisne.org/>>. Acesso em: 25 fev. 2018.

STAR, Susan. Leigh. The Ethnography of Infrastructure. In: **American Behavioral Scientist**, v. 43, n. 3, p. 377–391, 1 nov. 1999.

_____. Creating an Intellectual Commons through Open Access. In: HESS, Charlotte; OSTROM, Elinor. **Understanding knowledge as a commons: from theory to practice**. Cambridge: MIT Press, 2007.

STAR, Susan; RUHLEDER, Karen. Steps toward an ecology of infrastructure: Design and access for large information spaces. In: **Information systems research**, v. 7, n. 1, p. 111-134, 1996.

SUGIMOTO, Luiz. Evento discute modelo inovador visando à pesquisa de fármacos. In: **Portal da Unicamp**, 2014. Disponível em: <<http://www.unicamp.br/unicamp/noticias/2014/04/28/evento-discute-modelo-inovador-visando-pesquisa-de-farmacos>>. Acesso em: 14 set. 2014.

SZIGETI, Balázs; GLEESON, Pdraig; VELLA, Michael... OpenWorm: an open-science approach to modeling *Caenorhabditis elegans*. **Frontiers in computational neuroscience**, v. 8, p. 137, 2014. Disponível em: <10.3389/fncom.2014.00137>. Acesso em: 13 jul. 2018.

TOLEDO, KAREN. Free software helps to manage neuroscience experimental data. **AGÊNCIA FAPESP**, 2015. Disponível em: <<http://agencia.fapesp.br/free-software-helps-to-manage-neuroscience-experimental-data/20463/>>. Acesso em: fev. 2015.

THE CENTER FOR OPEN SCIENCE. Disponível em: <<https://cos.io/>>. Acesso em: 24 jan. 2019.

TRAÇO DE CIÊNCIA. **CEPID NeuroMat cria nova versão da página Projeto Wikipédia Falada**, 7 fev. 2017. Disponível em: <<https://difusaoneuromat.wordpress.com/2017/02/07/cepid-neuromat-cria-nova-versao-da-pagina-projeto-wikipedia-falada/>>. Acesso em 20 mar. 2018.

VARGAS, Cláudia. D.; KON, Fábio. Em defesa do compartilhamento público de dados Científicos. **Le Monde Diplomatique Brasil**. Disponível em: <<http://www.diplomatique.org.br/artigo.php?id=1653>>. Acesso em: 10/12/2014.

VENTURA, Dora. F. Um Retrato da Área de Neurociência e Comportamento no Brasil. **Psicologia: Teoria e Pesquisa**, v. 26, n. 25, p. 123–130, 30 out. 2010. Disponível em: <<https://revistaptpt.unb.br/index.php/ptp/article/view/474>>. Acesso em: 26 dez. 2018.

- VELHO, Léa. A Ciência da Informação e seu público. **Transinformação**, v. 9, n. 3, 1997. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/1575>>. Acesso em: 09 jan. 2019.
- VICENTIN, Diego Jair. A reticulação da Banda Larga Móvel: definindo padrões, informando a rede. Tese de Doutorado – Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas, Campinas, 2016.
- VERMEULEN, N.; PARKER, J. N.; PENDERS, B. Understanding life together: A brief history of collaboration in biology. **Endeavour**, v. 37, n. 3, p. 162–171, 1 set. 2013.
- WAISELFISZ, Julio Jacobo. **Mapa da violência 2013**: acidentes de trânsito e motocicletas. Rio de Janeiro: Centro Brasileiro de Estudos Latino-Americanos; FLACSO. 2013.
- WIKIVERSIDADE. **Pesquisa em Difusão Científica do CEPID NeuroMat. S/d.** Disponível em: <https://pt.wikiversity.org/wiki/Pesquisa_em_Difus%C3%A3o_Cient%C3%Adfica_do_CEPID_NeuroMat>. Acesso em 20 mar. 2018.
- WILKINSON, Mark D. et al. **The FAIR Guiding Principles for scientific data management and stewardship.** Comments and Opinion. Disponível em: <<https://www.nature.com/articles/sdata201618>>. Acesso em: 29 set. 2018.
- WOMACK, Ryan P. Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics. **PLOS ONE**, v. 10, n. 12, p. e0143460, 4 dez. 2015. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143460>>. Acesso em: 12 set. 2018.
- WOUTERS, Paul; BEAULIEU, Anne. Imagining E-Science beyond Computation. In: HINE, Christine (ed.). **New infrastructures for knowledge production: understanding E-science.** Hershey, PA: Information Science Pub, 2006. p. 48–70, 2006.
- WYKLE, Stacy S. Enclaves of anarchy: Preprint sharing, 1940-1990. In: **Proceedings of the American Society for Information Science and Technology**, v. 51, n. 1, p. 1–10, 1 jan. 2014.
- WILEY, H. Steven; MICHAELS, George S. Should software hold data hostage? **Nature Biotechnology**, v. 22, n. 8, p. 1037–1038, ago. 2004.
- ZOLNERKEVIC, Igor. Conexões dinâmicas: Matemáticos e neurocientistas se unem para entender e prever o funcionamento do cérebro. **Pesquisa FAPESP**, n. 218, abr. 2014.
- ZIMAN, John. **Real Science: Whats it is, and what it means.** Cambridge: Cambridge University Press, 2000. ZIMAN, John. **Real Science: Whats it is, and what it means.** Cambridge: Cambridge University Press, 2000.

APÊNDICES

APÊNDICE A – Roteiro de questões e tópicos utilizado nas entrevistas

Questões

- 1) Comece me falando que tipo de pesquisa você desenvolve aqui.
- 2) Como sua pesquisa relaciona como a proposta do Neuromat?
- 3) Quais os desafios existentes nessa pareceria?
- 4) O que o(a) Sr/Sra entende por Ciência Aberta?
- 5) Quais tipos de dados são produzidos por você?
- 6) O que vem sendo feito na sua área sobre Ciência Aberta?
- 7) Utiliza de recursos como repositórios, periódicos abertos, dados, no seu trabalho etc?

Tópicos a serem abordados

- ✓ Ciência Aberta – opinião, compreensão, críticas.
- ✓ Dados de pesquisa
- ✓ Ciência Aberta no NeuroMat.
- ✓ Neurociências e compartilhamento de dados
- ✓ Relação com a FAPESP
- ✓ Autoria, propriedade intelectual, etc.

APÊNDICE B – Fontes consultadas disponíveis online

FONTES ONLINE	
F1	Scientific project submitted in 2012. Disponível em: < http://neuromat.numec.prp.usp.br/content/scientific-project-submitted-2012 >.
F2	Neuromatemática, a nova ciência do cérebro. Universidade de São Paulo. Disponível em: < https://www5.usp.br/42539/projeto-liderado-pela-usp-investiga-a-neuromatematica-nova-ciencia-do-cerebro/ > .
F3	Neuromat oferece curso de medição de disparos neuronais. Revista Mente e Cérebro. Disponível em: < http://www2.uol.com.br/vivermente/noticias/neuromat_oferece_curso_de_medicao_de_disparos_neuronais.html >
F4	A ciência traz valores fundamentais à sociedade (Entrevista com Cláudia Vargas) 2014. Disponível em: < https://neuromat.numec.prp.usp.br/sites/default/files/entrevista-com-Claudia-Domingues-Vargas.pdf >
F5	Página do NeuroMat no Facebook. Disponível em: < https://pt-br.facebook.com/neuromathematics/ >
F6	Site do NeuroMat. Disponível em: < http://neuromat.numec.prp.usp.br/ >
F7	NeuroMat Newsletter. Disponível em: < http://neuromat.numec.prp.usp.br/newsletter/neuromat >
F8	Canal do NeuroMat - YouTube. Disponível em: < https://www.youtube.com/user/neuromathematics >
F9	ABRAÇO. Disponível em: < http://abraco.numec.prp.usp.br/ >
F10	Pesquisa em Difusão Científica do CEPID NeuroMat. Wikiversidade. Disponível em: < https://pt.wikiversity.org/wiki/Pesquisa_em_Difus%C3%A3o_Cient%C3%Adfica_do_CEPID_NeuroMat. >
F11	Traço de Ciência-Blog de Difusão Científica do CEPID NeuroMat. Disponível em: < https://difusaoneuromat.wordpress.com/ >
F12	NeuroMat é o maior produtor de conteúdo de matemática na Wikipédia. AGÊNCIA FAPESP. Disponível em: < http://agencia.fapesp.br/neuromat-e-o-maior-produtor-de-conteudo-de >

	matematica-na-wikipedia/28179/ >
F1 3	CEPID Neuromat First Report of Activities, August 2013 - September 2014. Disponível em: < http://neuromat.numec.prp.usp.br/sites/default/files/reports/first_report_of_activities.pdf >
F1 4	Collaboration among mathematicians and neuroscientists is a necessity to move brain theory forward: an interview with Remco van der Hofstad, 28 out. 2015. Disponível em: < https://neuromat.numec.prp.usp.br/content/collaboration-among-mathematicians-and-neuroscientists-necessity-move-brain-theory-forward >
F1 5	Eventos do NeuroMat. Disponível em: < http://neuromat.numec.prp.usp.br/events >
F1 6	Relatórios de Atividades do NeuroMat. Disponível em: < https://neuromat.numec.prp.usp.br/reports >
F1 7	Spike Sorting, Christophe Pouzat, NeuroMat. Disponível em: < https://www.youtube.com/watch?v=vSydfDvsewY >
F1 8	NeuroMat - Ciência Aberta (completo). < https://www.youtube.com/watch?v=kNfkSUcedoY >
F1 9	Spike sorting: What is it? Why do we need it... IPTV USP. Disponível em: < https://iptv.usp.br/portal/video.action?idItem=25412 >
F2 0	Free software helps to manage neuroscience experimental data. Disponível em: < http://agencia.fapesp.br/free-software-helps-to-manage-neuroscience-experimental-data/20463/ >.
F2 1	Página de registros de atividades na Wikipédia. Disponível em: < https://pt.wikipedia.org/wiki/Usu%C3%A1rio(a):MGromov >
F2 2	Laboratório de Sistemas Neurais (SisNe). < http://sisne.org/ >.
F2 3	Código fonte do NES no Github. Disponível em: < https://github.com/neuromat/nas#neuromat-nas-neuroscience-experiments-system >.
F2 4	Documentação do NES. < https://nes.readthedocs.io >