

Andre Luiz Appel

A e-Science e as atuais práticas de pesquisa científica

Dissertação de mestrado
Março de 2014



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
ESCOLA DE COMUNICAÇÃO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Andre Luiz Appel

A E-SCIENCE E AS ATUAIS PRÁTICAS DE PESQUISA CIENTÍFICA

Rio de Janeiro
2014

Andre Luiz Appel

A E-SCIENCE E AS ATUAIS PRÁTICAS DE PESQUISA CIENTÍFICA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

Orientadora: Profa. Maria Lucia Maciel.

Coorientadora: Profa. Sarita Albagli.

Rio de Janeiro

2014

A646 Appel, Andre Luiz.
A e-science e as atuais práticas de pesquisa científica / Andre
Luiz Appel. Rio de Janeiro, 2014.
88f.

Orientadora: Maria Lucia Maciel.
Coorientadora: Sarita Albagli.

Dissertação (mestrado) – Universidade Federal do Rio de
Janeiro, Escola de Comunicação, Programa de Pós Graduação
em Ciência da Informação, 2014.

1. Ciência da informação. 2. e-Science. 3. Tecnologia da
informação. 4. Gestão da informação. 5. Gerenciamento de
recursos de informação. I. Maciel, Maria Lucia. II. Albagli,
Sarita. III. Universidade Federal do Rio de Janeiro. Escola de
Comunicação. II. IBICT.

CDD: 020

Andre Luiz Appel

A E-SCIENCE E AS ATUAIS PRÁTICAS DE PESQUISA CIENTÍFICA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

Orientadora: Profa. Maria Lucia Maciel.
Coorientadora: Profa. Sarita Albagli.

Aprovada em: ___ / ___ / ____.

Profa. Dra. Maria Lucia Maciel – PPGCI-IBICT/UFRJ-ECO
Orientadora

Profa. Dra. Sarita Albagli – PPGCI-IBICT/UFRJ-ECO
Coorientadora

Profa. Dra. Gilda Olinto de Oliveira – PPGCI-IBICT/UFRJ-ECO

Profa. Dra. Lea Maria Leme Strini Velho – PPGPCT-UNICAMP

AGRADECIMENTOS

Às minhas orientadoras Maria Lucia Maciel e Sarita Albagli, por compartilharem suas experiências e seus vastos conhecimentos, pelo exemplo de competência, e por me ajudarem a desenvolver novos olhares frente às questões e aos desafios de pesquisa já enfrentados e aos que ainda estão por vir.

Às Professoras Gilda Olinto e Lea Velho pelas contribuições e sugestões ao longo do processo de construção deste trabalho.

Aos colegas e amigos do PPGCI, pelo companheirismo e pelo apoio nos momentos difíceis, de aflições e pressões mil, e pelos momentos felizes de descontração, de discussões produtivas e de diversas conquistas compartilhadas.

Aos Professores e funcionários do PPGCI, pelos nobres ensinamentos e pelo apoio, carinho e atenção com que sempre nos receberam.

Aos pesquisadores entrevistados, pelo tempo dedicado e pela atenção com que me receberam.

À minha mãe e à minha irmã, Iria Appel e Fabiola Appel, por fazerem parte da minha vida, por estarem sempre me dando forças e apoio e a quem eu espero sempre poder retribuir todo o amor e carinho recebido.

À amiga e eterna Professora Patrícia Zeni Marchiori, por ter me motivado na escolha pelo PPGCI-IBICT/UFRJ, possibilitando que eu entrasse em contato com o melhor dos ambientes para a minha formação como pessoa e como profissional.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pelo financiamento dos meus estudos, sem o qual eu certamente não teria alcançado tão cedo esta conquista.

RESUMO

APPEL, Andre Luiz. *A e-Science e as atuais práticas de pesquisa científica*. Orientadora: Profa. Dra. Maria Lucia Maciel; Coorientadora: Profa. Dra. Sarita Albagli. Rio de Janeiro, 2014. 88 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2014.

Este trabalho teve como objetivo analisar a relação entre novas práticas de produção colaborativa de conhecimento científico e o desenvolvimento e uso de plataformas tecnológicas de amparo à pesquisa colaborativa, movimento conhecido como *e-Science*, levando-se em consideração as diferentes visões, perspectivas e interesses dos atores atuantes nessas práticas, suas opções de uso e adesão às plataformas de pesquisa de *e-Science* e as definições quanto aos direitos de acesso e uso dos dados e resultados de pesquisa em tais práticas. O uso do termo *e-Science* diz respeito a um movimento que prevê a concepção de uma ciência atuante no uso intensivo de dados e na colaboração por meio do uso de plataformas de pesquisa baseadas em computação avançada. Como proposta metodológica deste trabalho, desenvolveu-se uma primeira fase documental, baseada em levantamento na literatura, para a exploração do histórico, de conceitos e práticas relacionados à *e-Science*. Em um segundo momento, desenvolveu-se uma abordagem empírica, com o estudo de uma experiência em *e-Science*, mais especificamente, o caso do *Conseil Européen pour la Recherche Nucléaire* (CERN). Para desenvolvimento desse estudo de caso foram entrevistados pesquisadores e especialistas atuantes em colaborações junto ao CERN e em demais pesquisas e iniciativas relacionadas à *e-Science*. Os resultados das entrevistas foram analisados com base em categorias e conceitos-chave elencados a partir dos questionamentos iniciais da pesquisa e do referencial teórico. Dentre os principais resultados e considerações destacados ao longo do trabalho está a percepção da governança como uma dimensão significativa no contexto da *e-Science*. Ganham destaque também as implicações referentes às condições de financiamento à pesquisa e às formas de organização dos atores e dos grupos de pesquisa para viabilizar a colaboração, além da descoberta dos dados como ativos importantes nos processos de produção de ciência, e como isso afeta as estruturas de avaliação e mensuração de resultados nesses processos. Em linhas gerais, destaca-se convergência dos objetivos propostos e temas abordados, apresentando-se como principal desafio a carência de estudos e publicações brasileiras sobre o assunto. Como perspectivas futuras de estudo, destaca-se a potencialidade de desenvolvimento de um *framework* para análise e estudo da governança em *e-Science*, com foco na abertura dos processos de geração e tratamento de dados derivados de pesquisas colaborativas, ou de iniciativas por parte da comunidade científica brasileira alinhada a esses processos, garantindo-lhe um posicionamento estratégico no campo das colaborações em *e-Science*.

Palavras-chave: *e-Science*. Produção de ciência. Uso intensivo de dados. Pesquisa colaborativa.

ABSTRACT

APPEL, Andre Luiz. *A e-Science e as atuais práticas de pesquisa científica*. Orientadora: Profa. Dra. Maria Lucia Maciel; Coorientadora: Profa. Dra. Sarita Albagli. Rio de Janeiro, 2014. 88 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2014.

This study aimed to investigate the relationship between new practices of collaborative production of scientific knowledge and the development and use of technological platforms for support to collaborative research, a movement known as e-Science. This was done considering the divergent views, perspectives and interests of actors within these practices, also considering their options of use and adherence to the e-Science platforms as well as the definitions related to the rights of access and usage of research data and results in those practices. The concept of e-Science explored here is related to a movement that requires the design of a data-intensive science and collaboration through research-based advanced computing platforms. The methodological approach involved a literature research to explore the history, concepts and examples of practices related to e-Science. As a second step, an empirical approach was developed, with the study of an e-Science experience, the case of *Conseil Européen pour la Recherche Nucléaire* (CERN). We interviewed a group of researchers who collaborate in CERN based projects and other research programs related to e-Science. The interviews were analyzed based to categories and key-concepts evidenced from the primary research questions and the supporting literature. One of the main results highlighted throughout the work is the perception of governance as a significant dimension in the context of e-Science. Another highlighted result is the discovery of implications regarding the conditions of research funding and forms of organization of actors and research groups to enable collaboration, as well as the discovery of data as important assets in the processes of science production and how it affects the structures assessment and measurement of outcomes in these processes. In general, it was possible to observe the convergence of the proposed objectives and covered topics and as major challenge to the work is the lack of Brazilian studies and publications on the subject. As future prospects of study, there is the potential for developing a framework for analysis and study of the governance in e-Science focusing on opening up the processes of data creation and processing, and the proposal of Brazilian scientific community initiatives aligned to these processes, guaranteeing to it a strategic position in the field of e-Science collaborations.

Keywords: e-Science. Science production. Data intensive. Collaborative research.

SUMÁRIO

1	INTRODUÇÃO	8
2	A EMERGÊNCIA DA E-SCIENCE NAS PRÁTICAS DE PESQUISA	11
2.1	Infraestrutura computacional de <i>e-Science</i>: elementos essenciais	12
2.2	A produção e o uso intensivo de dados em <i>e-Science</i>	13
2.3	A colaboração no contexto da <i>e-Science</i>	19
3	DIMENSÕES SOCIAIS E ORGANIZACIONAIS DA CIÊNCIA	24
3.1	Estudos relacionados às práticas científicas	25
3.2	Gestão e organização das práticas científicas	31
3.3	Fluxos e circulação da informação nas práticas científicas	36
4	PRÁTICAS DE COLABORAÇÃO EM E-SCIENCE: O CASO DO CERN	40
4.1	Sobre o CERN	43
4.2	A colaboração em <i>e-Science</i> no CERN	45
4.2.1	Sobre a atuação no CERN.....	46
4.2.2	Sobre as condições de participação no CERN ou em suas colaborações	46
4.2.3	Sobre o acesso aos dados gerados pelas colaborações do CERN	52
4.2.4	Resultados e condições de avaliação das colaborações.....	57
4.2.5	Cenário Brasil.....	61
5	CONSIDERAÇÕES FINAIS	67
	REFERÊNCIAS	71
	APÊNDICE A – Roteiro de entrevistas.....	77
	APÊNDICE B – Quadro de análise das entrevistas	80

1 INTRODUÇÃO

Esta pesquisa partiu do interesse no estudo da relação entre novas práticas de produção colaborativa de conhecimento científico e o desenvolvimento e uso de plataformas tecnológicas de amparo à pesquisa colaborativa. Essa emergência de novas plataformas e práticas de pesquisa tem sido identificada por diferentes autores sob o nome de *e-Science* – entre outras nomenclaturas, com veremos mais à frente –, e compreende, além de outros aspectos, o uso intensivo de dados na pesquisa científica, o uso de infraestrutura tecnológica para a colaboração entre cientistas e os processos de colaboração científica em si, mediados pelas tecnologias.

Com base nessa premissa, passou-se a formular alguns questionamentos relacionados a esse fenômeno, assim como possíveis rotas metodológicas que pudessem auxiliar na elucidação e compreensão de tais questionamentos. Buscou-se privilegiar o entendimento de quem são os principais atores (e seus papéis) envolvidos na pesquisa em *e-Science*, suas formas de interação e organização, e como essa interação interfere ou recebe interferências nas/das condições sob as quais são estruturadas as plataformas para desenvolvimento da *e-Science* e como ocorrem as formalizações dos processos de colaboração em tais plataformas, assim como seus padrões de interoperabilidade.

Da mesma forma, um desses questionamentos envolve a necessidade de identificar onde e como serão armazenados os dados ou os principais serviços/plataformas de *e-Science*, bem como as formas e condições de uso e compartilhamento de dados e resultados de pesquisas que fazem uso de tais plataformas.

Um outro questionamento envolve a necessidade de entender como se manifestam as disputas por prioridade ou recompensa a partir dos resultados de pesquisas desenvolvidas nas plataformas de *e-Science*, ou seja, analisar como se dão os processos decisórios sobre essas e outras questões em que estão presentes os mais variados interesses, sejam eles políticos, econômicos ou epistemológicos.

Esses e outros questionamentos foram tomados como objeto de análise, com base na afirmação de Schroeder (2008) de que o estudo da *e-Science* é um bom indicativo para examinar mudanças nas práticas de pesquisa, pois envolve discussões extensivas no âmbito das políticas de ciência acerca da forma como devem ser organizados os esforços para uma colaboração de crescente complexidade em escala e escopo, distribuída entre diferentes instituições e regiões geográficas e também interdisciplinar.

Nessa direção, formulou-se o objetivo geral dessa pesquisa, que era o de analisar os fatores que influenciam as práticas de colaboração científica a partir da introdução da *e-Science*, levando-se em consideração as diferentes visões, perspectivas e interesses dos atores atuantes nessas práticas, suas opções de uso e adesão às plataformas de pesquisa de *e-Science* e as definições quanto aos direitos de acesso e uso dos dados e resultados de pesquisa em tais práticas.

Esse objetivo geral foi, então, decomposto nos seguintes objetivos específicos:

- a) contribuir para o entendimento das mudanças nas práticas dos cientistas brasileiros decorrentes da sua inserção na colaboração em *e-Science*;
- b) analisar como são estruturadas as plataformas para desenvolvimento da *e-Science* e como ocorrem as formalizações dos processos de colaboração em tais plataformas, assim como seus padrões de interoperabilidade;
- c) identificar e analisar as formas e condições de uso e compartilhamento de dados e resultados de pesquisas que fazem uso de tais plataformas;
- d) verificar quais são e como se manifestam as disputas por prioridade ou recompensa a partir dos resultados de pesquisas desenvolvidas nas plataformas de *e-Science*.

Os questionamentos e objetivos motivaram uma exploração inicial na literatura, com maior disponibilidade no idioma inglês, permitindo uma compreensão inicial do fenômeno, além de gerar subsídios para decisão sobre a forma de se explorar o fenômeno no campo empírico. Elaborou-se, então, um referencial teórico, cuja apresentação se dá de acordo com os capítulos descritos a seguir.

No primeiro capítulo, que trata da *emergência da e-Science nas práticas de pesquisa*, são apresentados conceitos de *e-Science* evidenciados na literatura, suas possíveis dimensões e abordagens de estudo. Envolve conceitos relacionados à formação de uma infraestrutura de amparo à colaboração em *e-Science*, discussões sobre as questões de acesso, uso e compartilhamento de dados de pesquisa e discussões sobre a organização dos atores atuantes nas práticas de pesquisa desencadeadas pela *e-Science*.

O segundo capítulo, *Dimensões sociais e organizacionais da ciência*, apresenta e discute o contexto em que as discussões sobre a *e-Science* e sobre novas práticas se inserem. São discutidos elementos dos novos estudos da ciência, conceitos de governança

em meio às práticas científicas, além de questões ligadas aos fluxos e à circulação de informações em meio a redes sociotécnicas.

O terceiro capítulo, *Práticas de colaboração em e-Science: o caso do CERN*, apresenta o estudo de uma experiência em *e-Science*, mais especificamente, o caso do *Conseil Européen pour la Recherche Nucléaire* (CERN). Para desenvolvimento desse estudo de caso foram entrevistados pesquisadores e especialistas atuantes em colaborações junto ao CERN e em outras pesquisas e estudos relacionados à *e-Science*. Os resultados dessas entrevistas foram analisados com base em categorias e conceitos-chave elencados a partir dos questionamentos iniciais da pesquisa e do referencial teórico trabalhado nos capítulos anteriores.

O último capítulo apresenta considerações finais, destacando-se que foram elucidados vários dos questionamentos levantados no início da pesquisa, assim como novos elementos de análise também foram observados, como a percepção da governança como uma dimensão significativa no contexto da *e-Science*.

Destaca-se aqui também a evidência de implicações referentes às condições de financiamento à pesquisa e às formas de organização dos atores e dos grupos de pesquisa para viabilizar as colaborações, além da descoberta dos dados como ativos importantes nos processos de produção de ciência, e como isso afeta as estruturas de avaliação e mensuração de resultados nesses processos.

2 A EMERGÊNCIA DA E-SCIENCE NAS PRÁTICAS DE PESQUISA

O termo *e-Science*¹ foi introduzido por John Taylor no ano de 2001, enquanto diretor geral do Conselho de Pesquisa do *Office of Science and Technology* (OST) do Reino Unido, para denominar novas formas de colaboração e o trabalho multidisciplinar em áreas-chave da ciência e referindo-se à infraestrutura necessária para possibilitar essas novas formas de trabalho (HEY; TREFETHEN, 2002).

Entre outros fatores que impulsionam a *e-Science*, constata-se a gênese de uma ciência que passou a se valer de uma vertente quase que exclusivamente computacional em algumas áreas do conhecimento, com forte apelo à pesquisa por meio da simulação de fenômenos complexos. Vislumbra-se, então, um momento da exploração de grandes volumes de dados, em que se busca a interlocução de teorias, experimentos e simulações, caracterizando o momento da ciência de uso intensivo em dados ou *data-intensive* (GRAY, 2007; BELL; HEY; SZALAY, 2009).

Todos esses dados derivam de uma ampla variedade de tecnologias e da colaboração entre pessoas, realizada em larga escala. A aplicação de sensores, equipamentos de leituras e medições automatizadas, experimentos com conjunto de dados retroalimentáveis, sistemas de informação e bases de dados alimentadas por cientistas que registram e realizam anotações em dados de pesquisa são apenas alguns exemplos de possíveis fontes geradoras de dados. O trabalho na formulação de modelagens ou metodologias para análise desses dados também intensifica os fluxos de colaboração entre instituições e cientistas.

A armazenagem e análise de tais dados requer, por sua vez, a preexistência de uma infraestrutura computacional robusta, expansível e, preferencialmente, que possa ser organizada ou acessada de forma distribuída, para que os cientistas possam contribuir, por exemplo, na análise dos dados a partir de qualquer parte do planeta desde que disponham de acesso à internet.

Em linhas gerais, baseando-se em tais evidências conceituais, é possível apontar ao menos três fatores básicos para o desenvolvimento da *e-Science*:

¹ A literatura revela que a *e-Science*, além de ter grafias diferenciadas como *eScience*, *escience*, *EScience* etc., também pode receber o nome de *Cyberinfrastructure*, *cyberscience*, *eInfrastructure* e *eResearch* (MEDEIROS; CAREGNATO, 2012, p. 315). Para esta pesquisa, optou-se pelo uso do termo *e-Science*.

- a) a construção de uma infraestrutura computacional para uso distribuído ou para processamento de larga escala;
- b) a produção e o uso intensivo de dados; e
- c) a colaboração entre cientistas, grupos de cientistas ou instituições, pelo compartilhamento de esforços, dados e/ou recursos computacionais.

Schroeder (2008) reforça essa ideia apontando como possíveis dimensões do conceito de *e-Science* o uso de tecnologias computacionais para o compartilhamento distribuído de recursos digitais na pesquisa científica ou acadêmica e a promoção da colaboração global, não se aplicando apenas à questão do uso da internet, mas considerando também a dimensão institucional da *e-Science*. É com base, então, nessa categorização que os conceitos relacionados à *e-Science* são abordados nas seções subsequentes.

2.1 Infraestrutura computacional de *e-Science*: elementos essenciais

A implementação de infraestrutura tecnológica para *e-Science* é baseada na construção e disponibilização de *Grids*, *Middleware*, *Workbenches*, *Webservices*, *Virtual Research Environments* (VRE), tecnologias de notação e armazenamento de dados concebidas em padrões como XML², entre outros elementos técnicos.

O termo *Grid*, computação em grade, foi utilizado pela primeira vez no final dos anos 1990 (FOSTER; KESSELMAN, 1998), para denotar uma infraestrutura computacional distribuída servindo a atividades de ciência e engenharia avançadas. À época, a concepção dessa ideia foi direcionada pela vontade de se utilizar (um *cluster*³ de) recursos de computação distribuídos como um metacomputador. O nome *grid* é derivado da analogia com a rede/grade de energia elétrica (do inglês *electricity power grid*), compreendendo a ideia de que o poder de computação poderia estar disponível para qualquer pessoa em qualquer lugar (HEY; TREFETHEN, 2005).

² *eXtensible Markup Language* “é um formato simples de marcação, baseado em texto, para representar informações estruturadas: documentos, dados de configuração, livros, transações, faturas e muito mais. É derivado de um formato padrão mais antigo denominado SGML (ISO 8879), a fim de ser mais apropriado para o uso na Web” (QUIN, c2010).

³ Conjunto de computadores conectados.

Os *grids* compreendem a disponibilização de uma infraestrutura de *hardware* e *software* que fornece acesso a recursos computacionais de alto nível (com alto poder de processamento de dados e informações), confiáveis, consistentes, abrangentes e de baixo custo (FOSTER; KESSELMAN, 1998). A questão da confiabilidade está ligada ao fato de os usuários requererem altos níveis de previsibilidade, sustentabilidade e desempenho a partir dos recursos presentes nos *grids*. Esses fatores são associados, em geral, à velocidade de banda larga, ao poder de computação e aos serviços de *software* e segurança. A consistência envolve a implementação de serviços, interfaces e operações padrão, para que todos os usuários tenham um mesmo entendimento das formas de uso dessas tecnologias. A abrangência requer que os serviços oferecidos via *grid* sejam acessíveis a partir de qualquer ambiente em que se deseje implementá-lo. O baixo custo é um fator que se desenvolve na medida em que as tecnologias se tornam amplamente aceitas e utilizadas.

O *Middleware* corresponde ao *software* disponível entre a infraestrutura computacional dos *grids*, distribuída em rede, e as aplicações de uso individual dos cientistas ou projetos e instituições, e permite que cada usuário possa compartilhar seus recursos ou ter acesso aos recursos de outros usuários (SCHROEDER, 2008). De forma literal, *middleware* corresponde ao *software* que “está no meio” (*middle*), fazendo a ponte entre as aplicações e a rede (HEY; TREFETHEN, 2005). Um conjunto de *software* com módulos, funções e aplicações pré-definidas é comumente chamado de *Workbench*, grosso modo chamada de bancada de trabalho.

Webservice é um sistema de *software* projetado para suportar interações interoperáveis máquina-a-máquina em uma rede⁴. Possui uma interface descrita em um formato processável por máquina/computador, ou seja, é uma camada intermediária que permite a comunicação entre o *software* ou sistemas presentes nos *grids*.

2.2 A produção e o uso intensivo de dados em *e-Science*

Um dos objetos da implementação da infraestrutura anteriormente descrita compreende a exploração de grandes quantidades de dados que são geradas ao longo de processos ou atividades de pesquisa. O fenômeno da geração de grandes volumes de

⁴ World Wide Web Consortium (W3C). *Web Services Glossary*. 11 fev. 2004. Disponível em: <<http://www.w3.org/TR/ws-gloss/#webservice>>. Acesso em: 3 jun. 2013.

dados vem sendo recorrentemente identificado como *Big Data*, dados que excedem as capacidades convencionais de processamento dos sistemas de bases de dados. Esses dados variam em volume, velocidade de crescimento e variabilidade (sofrendo transformações), dificultando sua estocagem nas arquiteturas tradicionais de banco de dados, sendo que, para se obter algum valor a partir dos mesmos, existe a necessidade de se apelar para processos alternativos de armazenagem e processamento (DUMBILL, 2012).

Esse movimento/fenômeno do *big data* vem sendo explorado, além do campo científico, também por empresas e demais segmentos que trabalham com dados e informação, com finalidades que extrapolam a pesquisa científica. Os dados podem ser coletados a partir de textos de discussões e interações em redes sociais, *logs*⁵ de servidores da internet, sensores de tráfego, imagens de satélite, transações bancárias, dados do mercado financeiro, entre outros (DUMBILL, 2012), com as mais diversas finalidades, tais como estudos preditivos, estudos de comportamento dos consumidores, estudos de planejamento e gestão urbanos etc.

No campo da ciência, a terminologia *e-Science* é mais comumente utilizada, pois por si só já denota o uso intensivo de dados, e sua finalidade é estritamente a pesquisa científica. Outro fator que diferencia a *e-Science* da simples exploração de *big data* está na pesquisa colaborativa e no uso de recursos compartilhados para a exploração de dados, tal como se discute na sequência.

Na *e-Science*, ao menos três ações são tidas como chave quanto ao uso de dados: captura, curadoria e análise (GRAY, 2007). No que diz respeito à captura, vislumbram-se processos nos quais os dados são gerados diretamente por instrumentos de captura ou por meio de simulações computacionais prévias, processados por *software* e armazenados em bancos de dados. Em meio a esse processo, o papel de muitos cientistas é o de codificar as suas informações para que as mesmas possam ser intercambiadas com outros cientistas. Para isso a informação precisa estar representada de forma algorítmica, por meio de representações padrão de objetos em análise – tais como genes, galáxia, temperatura etc. (GRAY, 2007). A curadoria compreende a “gestão atuante e a

⁵ Um arquivo de *log* contém um registro de eventos gerados por *software* ou a saída de um dispositivo de *hardware*. A maioria dos arquivos de *log* são salvos em um formato de texto simples, permitindo sua visualização em um editor de texto básico. Podem também ser gerados por servidores *Web*, instaladores de *software*, ou uma variedade de aplicações. Os dados armazenados em um arquivo de *log* podem ser utilizados para solucionar problemas de *software* ou para a geração de relatórios (LOG..., 2010).

preservação de recursos digitais”, durante seu ciclo de vida, “tendo como perspectiva o desafio temporal de atender a gerações atuais e futuras de usuários” (SAYÃO; SALES, 2012, p. 184). A análise envolve o uso de tecnologias computacionais distribuídas e compartilhadas, os *grids*, as quais permitem o uso de recursos tecnológicos ou habilidades analíticas de cientistas posicionados em diferentes localidades.

Sobre os aspectos que dizem respeito à gestão e ao amadurecimento de soluções práticas voltadas para o uso e compartilhamento de dados científicos, algumas questões emergem, tais como: onde os dados são armazenados? Como serão organizados? Quem terá acesso? Quem irá pagar pela infraestrutura de armazenamento, compartilhamento e análise desses dados? (OHNO-MACHADO, 2012).

Nesse sentido, as iniciativas de *e-Science* costumam privilegiar o desenvolvimento de uma ciência aberta (*open Science*), ligada principalmente à disponibilização e manutenção de bases de dados abertos, de acesso público, que subsidiem o trabalho de pesquisa tanto no âmbito individual como no colaborativo. A formação desses conjuntos de dados é, em geral, decorrente de resultados de projetos de pesquisa financiados com recursos públicos, sendo disponibilizados para acesso público após um período de carência, ficando disponíveis para outros pesquisadores.

Buscando-se garantir a reprodutibilidade de posteriores estudos, os conjuntos de dados são “congelados” no momento da publicação e raramente sofrem modificações. Os novos resultados são incorporados como novas versões, e servindo a pesquisas posteriores (SÀNCHEZ-ARTIGAS; GARCÍA-LÓPEZ, 2010). Ohno-Machado (2012) acrescenta que a reprodutibilidade é de difícil realização, pois demanda os mesmos ambientes de *software* usados nos experimentos originais, requerendo instalações e configurações extensivas. A ideia do congelamento subentende a necessidade de os dados serem salvos ou publicados com um nível máximo de detalhamento e padronização, dispensando a necessidade de alterações posteriores.

No caso das ciências biomédicas, em especial, Ohno-Machado (2012) salienta que muitos periódicos dessa área ainda não estão equipados para implementar processos de revisão ou manutenção de grandes bases de dados anotados⁶ ou aplicações de *software*, e pequenos grupos de pesquisa podem não ter recursos para manutenção de dados,

⁶ Um dado anotado recebe uma descrição, comentário, marcação ou um conjunto de descritores (metadados, ou dados sobre dados) que facilitam a sua organização, arquivamento e recuperação, além do seu uso por outros pesquisadores.

metadados e recursos de *software* desenvolvidos para uso primário *in loco*. O acesso a dados gerados em outras localidades é igualmente difícil, pois mesmo que existam repositórios públicos, uma grande variedade de dados não está corretamente representada, tornando-se pouco útil. Nas ciências da saúde, existe ainda a falta de padronização dos dados advindos de diferentes instituições, além da necessidade de ações que permitam a proteção da privacidade institucional e pessoal (de pacientes), tal como um quadro de políticas que garantam níveis apropriados de consentimento e concordância com as regulações estabelecidas nos âmbitos institucional, estadual, federal e internacional (OHNO-MACHADO, 2012).

Sobre a disponibilização de dados abertos para a pesquisa científica, David, Besten e Schroeder (2006) argumentam que o crescimento dos estoques de tais dados depende da disponibilização de recursos de dados e informações como ‘bens públicos’ (*public goods*), que permitam o acesso e o compartilhamento irrestritos. Segundo os autores, esse panorama contrasta com as ações de controle e restrição do acesso às informações que geralmente são requeridas em favor da obtenção de benefícios materiais privados pela posse de conhecimentos científicos e tecnológicos, como os registros de propriedade intelectual. Em meio a um regime de propriedade, em que se busca a exploração comercial e rentável de descobertas e invenções, os investimentos de capital privado em pesquisa e desenvolvimento são realizados com base na expectativa de obtenção de lucro. É bom lembrar, no entanto, que boa parcela dos conhecimentos utilizados pelo setor privado advém de pesquisas oriundas de financiamento público.

Por outro lado, os adeptos da ciência aberta advogam pela disseminação de novas informações, considerando-se a expectativa de que com o aumento dos estoques de conhecimento público ocorrerá uma elevação das taxas de retornos sociais dos investimentos em pesquisa. Esse cenário, contudo, não elimina a rivalidade desencadeada em função das prioridades na revelação de descobertas e invenções, fazendo com que determinados grupos possam suspender atividades de cooperação em suas áreas do conhecimento (DAVID; BESTEN; SCHROEDER, 2006), estimulados principalmente pela necessidade de reconhecimento individualizado, que pode ser revertido em garantia de apoio e financiamento por parte das agências, ou seja, materializando o ditado “quem é visto é lembrado”.

Embora exista um grande número de ferramentas da *e-Science* para suporte a projetos de natureza aberta e distribuídos, não significa necessariamente que toda a

pesquisa colaborativa identificada como *e-Science* será qualificada como *open Science*. Mesmo projetos acadêmicos de *e-Science*, cuja intenção é trabalhar com dados abertos, falham em um ou mais aspectos, especialmente no que diz respeito às normas de divulgação de informações (DAVID; BESTEN; SCHROEDER, 2006).

David, Besten e Schroeder (2006, p. 3) apresentam dois grandes conjuntos de questões significativas na condução de projetos *open Science*. Um primeiro conjunto diz respeito aos termos sob os quais cada integrante pode ingressar ou deixar um projeto, abrangendo:

- a) quem está autorizado a juntar-se à colaboração?
- b) todos os participantes estão autorizados a receber acesso total às bases de dados do projeto e outras ferramentas-chave de pesquisa?
- c) o quão fácil ou difícil é para novos ingressantes definirem agendas distintas de investigação no contexto do projeto em andamento, e qual o nível de controle que os mesmos detêm sobre novas descobertas?
- d) quais restrições são colocadas (formal e informalmente) com relação ao uso de dados, informações e conhecimentos na posse de pesquisadores antes de os mesmos deixarem a colaboração?

Um segundo conjunto de questões apresentado pelos autores (DAVID; BESTEN; SCHROEDER, 2006) diz respeito às normas e regras que regulam a divulgação de dados e informações:

- a) em que proporção e quão rapidamente as informações sobre procedimentos e dados de pesquisa são divulgados?
- b) o quão completos os dados e procedimentos são documentados – de forma a que não estejam apenas acessíveis, mas também utilizáveis para quem não participa diretamente do grupo do projeto?
- c) em quais termos e a partir de quanto tempo pesquisadores externos poderão ter acesso a material e dados resultantes do projeto?
- d) em que momento a propriedade intelectual é considerada (pelos membros dos projetos ou pelas instituições de origem)? Haverá licenciamento para usuários externos ou será uma base de uso exclusivo?
- e) os direitos de uso desses recursos “protegidos” estarão condicionados ao pagamento de *royalties*?

- f) os acordos de transferência de materiais impõem o pagamento de taxas (para as pesquisas celulares, reagentes, espécies) substancialmente mais altas por parte de pesquisadores externos?

Pela observação de ambos os conjuntos de questões, percebe-se a possibilidade de formação de um possível *framework* e de um conjunto de documentação de processos e protocolos de projeto para nortear as ações de tomada de decisão em uma série de etapas ou situações da pesquisa, por parte dos pesquisadores, na suas relações com a colaboração e por parte da colaboração como um todo.

O compartilhamento de dados no plano da *e-Science* pode incorrer na busca de modelos ou exemplos de práticas que melhor atendam às necessidades de cada instituição, pesquisador ou grupo de pesquisadores participantes da colaboração. Ohno-Machado (2012), nesse sentido, descreve três possíveis modelos de compartilhamento que envolvem: a) *download* de dados por usuários; b) acesso e análise de dados remotos pelos usuários; c) usuários importam todo um conjunto de ambiente de *software*.

No caso do *download* de dados, usuários responsáveis pela busca identificam fontes de dados relevantes a partir de servidores distribuídos ou centralizados e os baixam diretamente para computadores locais, do tipo estação de trabalho – de uso individual. Tal processo incorre em alguns problemas, tais como a geração de quantidades massivas de dados, necessidade de uma alta frequência de acesso aos servidores para a atualização dos dados e necessidade de conexão de alta velocidade à internet.

No modelo em que usuários acessam e analisam dados remotamente, ainda persiste a necessidade de conexão com a internet, contudo as responsabilidades dos usuários sobre a perda de dados são reduzidas no caso de perda ou roubo de um computador, já que os dados encontram-se replicados em outros computadores ou servidores.

O modelo em que usuários importam todo um conjunto de ambiente de *software* é similar ao modelo anterior, contudo, ao invés de os usuários usufruírem de recursos computacionais externos, eles fazem o uso de máquinas virtuais (*virtual machines* – VMs) que permitem a manipulação de dados locais (armazenados no computador do usuário), assim como dados de sites remotos. Esse modelo também permite a formação de redes de

centros de colaboração, mesmo que políticas institucionais não permitam o compartilhamento de dados no nível individual.

Destaca-se ainda, no contexto da *e-Science*, que a distribuição, a heterogeneidade e o volume dos dados representam três dimensões fundamentais que precisam ser consideradas durante as fases de projeto e implementação de novas infraestruturas orientadas a dados, serviços e sistemas. Da mesma forma, a questão da heterogeneidade desempenha um papel chave devido à grande variedade de plataformas, fontes de dados, sistemas, formatos, interfaces, políticas, abordagens etc. É argumentável que a mesma amplia a complexidade no ambiente de trabalho e direciona os esforços para a busca de interoperabilidade e padronização (FIORE; ALOISIO, 2011).

2.3 A colaboração no contexto da *e-Science*

A colaboração científica em ciência é definida por Sonnenwald (2007, p. 645, tradução nossa) como “a interação que se situa em um contexto social entre dois ou mais cientistas, a qual facilita a partilha do propósito e da ação de execução de tarefas, respeitando a um objetivo superordenado e compartilhado mutuamente”. A autora lança mão de características de contextos científicos descritas na literatura para categorizar ou classificar a colaboração a partir de categorias disciplinares, geográficas ou organizacionais.

De acordo com a autora, novos tipos de colaboração estão sendo e ainda podem ser gerados ou estimulados a partir do uso ou implementação de Tecnologias da Informação e da Comunicação (TICs) especialmente quando os cientistas não podem, ou não devem, estar alocados/fixados geograficamente (SONNENWALD, 2007). A alocação de recursos digitais geograficamente distribuídos, o poder de computação (capacidade de processamento), as tecnologias de visualização de dados e a capacidade de armazenamento são colocados como os principais esforços para a construção de TICs como infraestrutura de suporte à pesquisa colaborativa (BARJAK et al., 2013).

Em meio à gênese de inúmeras ferramentas e plataformas colaborativas da internet, e com finalidades igualmente múltiplas, Dutton (2008) destaca, no âmbito da pesquisa científica, o conceito de “organização rede”. Segundo o autor, tal conceito é útil para distinguir esta forma organizacional daquelas mais formais ou localmente/fisicamente baseadas, sendo que a organização em redes colaborativas

promove configurações dinâmicas, e na maioria das vezes inter-organizacionais, e entre indivíduos.

Dentre as ações que mediam a formação de redes colaborativas e iniciativas que privilegiam o avanço da pesquisa científica por meio da internet, pode-se destacar (DUTTON, 2008):

- a) a ação de compartilhar, como a habilidade de criar objetos e documentos conectados por meio de uma rede distribuída, desse modo reconfigurando como e com quem determinadas informações são compartilhadas;
- b) a ação de contribuir, incluindo a habilidade de empregar aplicações de *social networking* da internet para facilitar a comunicação entre grupos; e
- c) a ação de cocriar, a qual implica a habilidade de colaborar por meio de redes para facilitar o trabalho cooperativo em direção a objetivos comuns, desse modo reconfigurando o sequenciamento, composição e definição de papéis dos colaboradores.

Por outro lado, as infraestruturas tecnológicas para pesquisa colaborativa podem ser usadas para facilitar a formação de ‘clubes fechados’, embora localizados de forma distribuída – incluindo laboratórios governamentais engajados em projetos secretos de defesa e equipes corporativas de pesquisa e desenvolvimento (P&D) que trabalham com dados e materiais proprietários, cujos resultados são guardados como segredos comerciais. Um sintoma disso é o fato de que muitos pesquisadores empregam *software* proprietário e instrumentos patenteados, e publicam suas descobertas em revistas científicas que cobram taxas elevadas de acesso aos artigos, embora divulguem suas descobertas frequentemente e por completo e colaborem livremente com colegas em bases informais e não contratuais (DAVID; BESTEN; SCHROEDER, 2006).

Sobre as implicações das tecnologias no processo da comunicação científica, tanto formal quanto informal, Meyer e Schroeder (2009) argumentam que esse processo vem ocorrendo crescentemente na internet, impulsionado pela praticidade e rapidez de acesso à informação quando comparado aos recursos tradicionais de busca por informação. Esse fator também pode ser explicado, de acordo com os autores, pelo crescimento do número de jovens pesquisadores, nativos da *web*. Com base em Garvey e Griffith (1967), argumentam que os próprios cientistas tendem a criar elementos que preenchem suas necessidades de informação não satisfeitos pelos canais tradicionais

existentes, fortalecendo a comunicação científica informal na internet, por meio de *blogs*, *wikis*, redes sociais etc. A presença *online* também amplia a visibilidade das pesquisas, possibilitando a aproximação de iniciativas correlatas.

Ainda no plano da comunicação, Vasileiadou e Vliegthartb (2009) examinam o papel da internet e da comunicação face-a-face na pesquisa colaborativa. Mais precisamente, os autores investigam se o uso da internet está relacionado com o aumento da produtividade, usando como estudos de caso dois grupos de pesquisa colaborativos e distribuídos. Fazendo o uso de indicadores quantitativos baseados em dados coletados *online* em tempo real e produtividade fracionada, examinam se, no contexto dos casos analisados, a ligação entre o volume de comunicação e a produtividade da pesquisa é condicionada pelo grau de interdependência entre os colaboradores. Os resultados sugerem que o impacto positivo do uso da internet na produtividade da pesquisa é limitado e pode ser relevante somente quando os esforços colaborativos enfrentam problemas de coordenação. Ao mesmo tempo, os autores observaram que os encontros face-a-face demonstram-se como importantes previsores de produtividade na pesquisa. Dentre os indicadores utilizados na pesquisa estavam:

- a) uso da internet: *e-mail* e troca de arquivos;
- b) encontros: face-a-face, por meio de observação;
- c) nível de participação: mensal, de pessoas por projeto;
- d) produtividade: artigos, capítulos de livros e apresentações fracionados (coautorias).

Diante dos indicadores aplicados pelos autores, faz-se importante a observação de que o *e-mail* pode representar um indicador restrito no que tange o processo de colaboração na internet que, se consideradas as múltiplas dimensões da *e-Science*, se apresenta como um fenômeno mais complexo, ou seja, com maior gama de variáveis. No caso dos encontros, é pertinente também considerar as interações em reuniões virtuais, por videoconferência etc. que, embora não sejam face-a-face, contemplam a sincronicidade dos encontros presenciais. Sobre a análise da produtividade, já se consideram elementos múltiplos para a mensuração das formas de produção e contribuição intergrupos de pesquisa (BJÖRNEBORN; INGWERSEN, 2004), extrapolando-se os limites advindos da mensuração de indicadores bibliográficos e coautorias.

As formas de organização dos recursos e de atores, considerando-se instituições, grupos ou indivíduos atuantes em projetos de *e-Science*, para a colaboração também são aspectos passíveis de análise e discussão. Ressalta-se que a construção e o uso de plataformas e tecnologias de pesquisa no contexto da *e-Science* apresentam novos desafios para a organização de grupos e unidades de pesquisa, permeando questões político-econômicas, que envolvem discussões sobre direito autoral e propriedade intelectual.

Schroeder (2008) chama a atenção para a existência de um forte estímulo para a internacionalização das iniciativas de *e-Science* o que, por sua vez, suscita o desenvolvimento de políticas restritivas de acesso a recursos computacionais acadêmicos presentes nos *grids*, por parte dos diferentes países envolvidos, o embate entre diferentes regimes de direito de propriedade intelectual, a formação de barreiras comerciais (como limitações de acesso em bases de dados e bibliotecas digitais de conteúdos acessíveis apenas para assinantes) e a necessidade de tradução de *software* ou tecnologias para permitir a interação de comunidades isoladas.

A criação de pontos para a formação de redes de colaboração interinstitucionais poderia demandar a formulação de acordos entre países, sendo esses passíveis de engessamento caso fossem tratados no escopo das políticas das relações internacionais, e a revisão das atuais formas de regulação sobre a propriedade intelectual que, segundo Albagli e Maciel (2011, p. 27) “[...] têm se mostrado inadequadas e ineficazes ante a crescente dificuldade de se introduzirem barreiras legais à ampla difusão e ao compartilhamento que os novos meios de circulação digital proporcionam”.

No plano das questões político-econômicas e de financiamento, Barjak et al. (2013) ressaltam a necessidade de se considerar a transição da prática de pesquisa baseada em projeto para a pesquisa de produção em larga escala e de qualidade, criando-se a demanda por fontes de financiamento sustentáveis e que privilegiem o uso de infraestrutura de *e-Science*. Os autores evidenciam alguns aspectos que devem ser considerados nessa transição, incluindo as formas de inserção na comunidade científica, o propósito e responsabilidades das ações e a necessidade de realização de atividades de ensino e aperfeiçoamento, construção e desenvolvimento de comunidades.

Outros aspectos destacados pelos autores, mais voltados para a formalização das ações de colaboração e de gestão, incluem o desenvolvimento de mecanismos de coordenação – com a avaliação dos níveis de coordenação: comunidade, mercado,

hierarquias e reciprocidade (rede), o estudo e implementação da formalidade da governança, com a análise dos incentivos motivadores da conformidade, as relações entre atores e as fontes de legitimidade de ações. Outros aspectos ainda dizem respeito ao tamanho e ao escopo da colaboração – considerando-se o número de organizações participantes, além do escopo de participação: geográfico, disciplinar (áreas do conhecimento) e setorial (acadêmico versus público-privado), e a análise dos fatores de sustentabilidade do financiamento, levando-se em conta o horizonte de tempo dos esquemas de financiamento e a duração e estabilidade dos sistemas de apoio (BARJAK et al., 2013).

Discussões paralelas também levam em consideração os possíveis, e já visíveis, impactos nas vigentes estruturas e sistemas de recompensa e avaliação de indicadores e produtividade em pesquisa no contexto de pesquisas colaborativas, o que, por sua vez, provoca impactos diretos na questão do financiamento. Como proposta de um arcabouço para o estudo de tais impactos, Jonkers, Moya Anegón e Aguillo (2012, p. 1375) esclarecem que a intensidade de uso de tecnologias da informação aplicadas à pesquisa oferece um indicador de um tipo específico de atividade de pesquisa que pode complementar os indicadores clássicos baseados em insumos de pesquisa (*inputs* – número de pesquisadores com dedicação exclusiva, financiamento à pesquisa etc.) e baseados em resultados de pesquisa (*outputs* – número de publicações, número de citações etc.).

Em alguns casos, análises de produtividade por meio de estudos bibliométricos ou cientométricos têm sido realizadas em associação ao método de análise de redes sociais (BJÖRNEBORN; INGWERSEN, 2004), permitindo o mapeamento das formas e do volume de interações entre os atores de uma determinada rede. Vale atentar, contudo, que os estudos webométricos podem não ser eficazes em mensurar publicações que fazem parte da “web profunda”⁷ (*deep web*), mesmo sendo este um canal que concentra – ainda que de forma não estruturada – um volume significativo de registros de atividade científica.

⁷ Também chamada de *web* invisível, concentra os conteúdos da internet que não são facilmente recuperáveis pelas ferramentas de busca tradicionais. Inclui documentos/arquivos não textuais/HMTL, com formatos multimídia (arquivos de áudio e vídeo), entre outros.

3 DIMENSÕES SOCIAIS E ORGANIZACIONAIS DA CIÊNCIA

Além da interferência dos fatores relativos às condições de infraestrutura, de recursos e de fatores condicionantes da colaboração ou coordenação para a *e-Science*, é importante também estudar o contexto mais amplo – social e organizacional – da ciência, que pode estimular ou desencorajar as práticas associadas ao desenvolvimento da *e-Science*. Neste capítulo, procede-se então a uma tentativa de se delinear tal contexto sob a ótica dos estudos sociais da ciência e da tecnologia, das condições de gestão e organização das práticas científicas, além da necessidade de mapeamento e compreensão dos fluxos e da circulação de informações nas práticas científicas.

Em linhas gerais, verifica-se que a ciência e as práticas científicas apresentam estágios de desenvolvimento que variam de acordo com a época, contexto sociocultural e econômico, ao mesmo tempo em que interferem e recebem interferência no/do desenvolvimento tecnológico. Tais estágios, por outro lado, não devem ser analisados sob uma perspectiva estritamente evolutiva, pois a gênese ou desenvolvimento de um novo estágio não anula ou substitui um estágio anterior por completo, havendo uma simbiose de práticas ou culturas científicas. Contudo, é importante analisar e procurar entender as principais características manifestadas por cada um desses estágios da ciência para possibilitar, entre outras finalidades, a explicação e a compreensão de fenômenos correntes, tal como a *e-Science*, que afetam as formas de “fazer ciência”.

Nesse sentido, buscou-se elencar algumas das características marcantes dessas práticas, além de tendências ou abordagens de estudos das ciências sociais ou estudos da ciência que permitiram a identificação de distinções e a formação de categorias de análise de tais práticas. Realizou-se, então, em uma perspectiva macro-histórica, a partir de meados do século XX até os dias atuais, um apanhado de todos esses estudos, com foco nos novos estudos da ciência. E é a partir dessa última visada, dos novos estudos da ciência, que se dará o exercício de análise e compreensão do fenômeno da *e-Science*, ou seja, serão levados em consideração conceitos como a governança em meio às práticas científicas e a questão dos fluxos e da circulação de informações em meio a redes sociotécnicas.

3.1 Estudos relacionados às práticas científicas

Seja como paradigmas, seja como estágios de características heterogêneas, a ciência apresenta determinadas fases ou ciclos de desenvolvimento que podem ser evidenciados e explicados a partir de correntes de pensamento dos Estudos Sociais da Ciência, ou mesmo a partir dos conceitos dominantes de ciência que vigoram por determinados períodos de tempo (VELHO, 2011).

Um primeiro marco desses estudos compreende a visão de uma ciência que está além das interferências sociais. A área da Sociologia da Ciência, nesse caso, “ocupa-se com o estudo das condições sociais e efeitos da ciência, e com as estruturas e processos sociais decorrentes da atividade científica” (BEN-DAVID, 1975, p. 203, tradução nossa). Baseando-se no trabalho de Ben-David (1975), Collins (1983) traduz o entendimento da Sociologia da Ciência como a elucidação de um conjunto de normativas e demais arranjos institucionais que permitem à ciência – entendida como responsável pela formulação de perguntas e respostas acerca da natureza – existir e funcionar de maneira eficiente. Essa concepção coloca a natureza como o último elemento capaz de emitir respostas sendo que o homem, nesse processo, aparece apenas como mediador. De acordo com o autor, “pré-requisitos institucionais adequados devem abreviar os efeitos mundanos dos desacordos e vieses, além de instituir um sistema de recompensas para encorajar a perseguição vigorosa de respostas” (COLLINS, 1983, p. 266-267, tradução nossa).

Desenvolveu-se, então, uma noção de ciência como tradição e instituição, inspirada pelos rigores e normas universais do trabalho do cientista – universalismo, comunismo, desinteresse e ceticismo organizado – descritos no ensaio *The Normative Structure of Science* de Robert Merton, publicado em 1942, e retomadas no livro *The Sociology of Science* do mesmo autor (MERTON, 1973). Tais normas, “modelam e normatizam o comportamento esperado dos membros da comunidade de pesquisa para garantir a produção de conhecimento livre de valores e de influências sociais” (VELHO, 2011, p. 137).

Nesse contexto, a ciência é vista como o motor do progresso (VELHO, 2011) e a intenção dos analistas sociais, em geral, é a de compreender, explicar e efetivamente reforçar o sucesso da ciência, ao invés de questionar as suas bases, posicionamento este defendido por Collins e Evans (2002), autores responsáveis pela adoção de uma abordagem reconhecidamente mais crítica em relação aos primeiros estudos da ciência.

Coincidem, nesse mesmo período, as ações de fortalecimento dos estados-nação, incluindo a injeção de fortes investimentos públicos em sistemas de ciência, tecnologia e inovação, sendo que, “nessa fase, o Estado, como principal financiador da ciência, delegava esta tarefa aos conselhos de pesquisa e esses delegavam as principais decisões aos próprios cientistas – os únicos juízes competentes” (VELHO, 2011, p. 138) que, recorrentemente, pautavam-se na objetividade autoaferida e no processo de revisão por pares para a legitimação de suas ações e escolhas. As ciências, então, adquiriram um cunho esotérico e autoritário, tornando inconcebível a tomada de decisões sobre questões envolvendo ciência e tecnologia que não viessem de cima para baixo, tal como explicam Collins e Evans (2002), autores que fazem parte de uma corrente que, reconhecidamente, atua no plano da crítica aos primeiros estudos sobre ciência e tecnologia.

Segundo Collins e Evans (2002), o que provocou uma quebra nessa “onda”, que eles classificam como positivista, foi o livro de Thomas Kuhn tratando da evolução da ciência e suas revoluções em um panorama histórico e introduzindo o conceito de paradigma científico (KUHN, 1998). Fuller (1999) argumenta que o trabalho de Kuhn foi um catalisador para investigações posteriores sobre os estudos sociais da ciência já que permitiu um ponto de aproximação entre a visão normativa e novas abordagens que viriam a reivindicar o papel da sociedade na construção e desenvolvimento da ciência.

A noção de paradigma científico de Kuhn (1998, p. 3) engloba a ideia de que existem “realizações científicas universalmente reconhecidas que, durante algum tempo, fornecem problemas e soluções modelares, para uma comunidade de praticantes de uma ciência”. Admite a existência de uma ciência normal que frequentemente suprime novidades fundamentais – que o autor define como arbitrariedades, porque essas subvertem necessariamente seus compromissos básicos de formulações acerca da realidade observada. A normalidade é afetada, com conseqüente ruptura paradigmática, com o surgimento das anomalias, “com o reconhecimento de que, de alguma maneira, a natureza viola as expectativas paradigmáticas que governam a ciência normal (KUHN, 1998, p. 78)”, que resultam nas descobertas ou, como o autor designa, revoluções científicas.

Observando-se o fenômeno da *e-Science* sob a perspectiva kuhniana, faz-se significativa a noção de que “conscientemente ou não, a decisão de empregar um determinado aparelho e empregá-lo de um modo específico baseia-se no pressuposto de que somente certos tipos de circunstâncias ocorrerão” (KUHN, 1998, p. 81). Os defensores

da *e-Science* como um novo, ou quarto, paradigma da ciência⁸ (GRAY, 2007) – sucessor dos modelos empíricos, teóricos e baseados em computação para a explicação da realidade – têm em pauta esse pressuposto, já que se espera, pela aplicação e uso das tecnologias e plataformas de pesquisa da *e-Science*, a concretização do trajeto: dados transformados em informação que se converte em conhecimento – ou em teorias.

Observam-se, nos constructos da *e-Science*, assim como nas movimentações que Kuhn destaca como sendo de reforço de um determinado paradigma, as ações de atualização da promessa de sucesso das tecnologias de pesquisa colaborativa para as grandes questões contemporâneas da ciência, ampliando-se o conhecimento sobre os fatos que o paradigma apresenta como particularmente relevantes (KUHN, 1998, p. 44), ou seja, no caso da *e-Science*, de que a ciência expandirá seu ciclo de descobertas somente pela exploração de imensas quantidades de dados.

Com a gênese de novos estudos, abriu-se caminho para o desenvolvimento da Sociologia do Conhecimento Científico, sob a perspectiva do Construtivismo Social, a qual está preocupada com “o que pode vir a contar como conhecimento científico e de que forma” (COLLINS, 1983, p. 267, tradução nossa). O “vir a contar”, nessa perspectiva, está associado à ideia do relativismo, sendo que as investigações levam em conta o quanto certas visões sobre o mundo descrito física ou matematicamente serão consideradas como corretas pela sociedade, ao invés de se pensar em como a sociedade pode ser arranjada para possibilitar o surgimento de visões corretas (COLLINS, 1983).

Em meio a esse panorama, Collins e Evans (2002) esclarecem ainda que é necessário recorrer a fatores extracientíficos para enquadrar os debates científicos e técnicos, sendo que método científico, experimentos, observações e teorias, tais como os pré-requisitos essenciais da ciência normativa, não são suficientes. Contudo, o mais distintivo e mais sustentável tema no plano dos estudos da sociologia do conhecimento científico relativistas/simétricos (primordialmente europeus) não derivara de uma reação contra as análises existentes. Quando os relativistas criticam a tradição, eles o fazem para distinguir seu campo em ascensão de uma linha ortodoxa que poderia confundi-los ao invés de servir de trampolim para novas ideias (COLLINS, 1983).

⁸ De acordo com Gray (2007), os paradigmas anteriores da ciência contemplam, respectivamente: a ciência em um ciclo de observações empíricas, pela descrição de fenômenos naturais; uma vertente teórica, que passou a valer-se de modelos e generalizações, compreendendo os dois últimos séculos; e, nas últimas décadas (século XX), a ciência passa a valer-se de uma vertente computacional, com forte apelo à pesquisa por meio da simulação de fenômenos complexos.

Os autores e trabalhos mais significativos para o surgimento e desenho dessa nova área de estudo são apresentados no trabalho de Collins (1983). Entre esses está o de Mulkay (1969), segundo o qual as normas técnicas e cognitivas são de maior importância e têm maior poder explicativo do que as normas elencadas por Merton (1973). O referido artigo traz o conhecimento científico para análise, discutindo restrições à inovação científica e as circunstâncias sob as quais a inovação ocorre apesar de restrições (COLLINS, 1983). Outro trabalho destacado foi escrito por David Bloor (1973), autor responsável por dar forma ao “programa forte” dos estudos da ciência, e defensor da ideia de que os sociólogos devem analisar teorias simétrica e imparcialmente sem perder o foco na questão da verdade ou da racionalidade (COLLINS, 1983). Outro autor é Whitley (1972), o qual defende que o processo de produção do conhecimento científico é tratado como uma caixa-preta, e somente suas entradas ou saídas é que podem ser estudadas (COLLINS, 1983).

Collins (1983) também destaca a contribuição de Latour e Woolgar (1979) e de Karin Knorr-Cetina (1983) ao movimento ou conjunto de pesquisas que ficou conhecido como estudos de laboratório. Uma das principais contribuições de Latour e Woolgar (1979) são as descrições de como um “fato” científico é gerado pelas atividades do dia-a-dia de um laboratório. Isso ocorre, por exemplo, quando uma série de dados gerados por instrumentos de medição dos laboratórios – aparentemente desconexos – uma vez unidos, parecem apontar para existência de um mesmo fato, por exemplo a existência de uma nova droga (COLLINS, 1983). Todos esses dados culminam no que os autores chamam de inscrições, uma série de documentos, gráficos, tabelas etc., que contêm representações da realidade elaboradas pelos cientistas e são capazes de reforçar ou descartar determinados fatos científicos.

No plano dos estudos mais recentes sobre ciência e sociedade, a partir da última década do século XX, os quais consideram uma participação social mais significativa nas decisões sobre os rumos da ciência, Velho (2011, p. 145) esclarece que

a nova concepção de ciência que está sendo delineada admite que existem muitas formas diferentes de conhecimento e que estas se relacionam de forma variável e assimétrica. Isto não quer dizer que não exista qualquer forma de autoridade epistêmica, mas esta será sempre o resultado de uma produção coletiva, com os seus momentos de conflito, que permitirá determinar, de maneira situada, a hierarquia dos saberes e da respectiva autoridade em função da situação, dos problemas, das prioridades e das consequências esperadas de intervenções associadas a esses saberes.

Ou seja, passam a contar também como autoridades epistêmicas os governos e a própria sociedade, de forma direta ou pela representação por meio de seus governantes. As motivações e os fatos científicos deixam de contar como os únicos elementos de apoio nos processos de tomada de decisão sobre as áreas de foco da pesquisa e sobretudo acerca do volume de investimentos em ciência e tecnologia.

Nesse ponto, Collins e Evans (2002) destacam a ocorrência de um embate na construção da legitimidade política das decisões técnicas no domínio público. Para os autores, o conceito de decisões técnicas compreende o processo de tomada de decisões nos pontos em que a ciência e a tecnologia intersectam o domínio político, pois as questões a serem decididas são de visível relevância para o público e para a sociedade. Os autores esclarecem, nesse sentido, que existe um problema de legitimação *versus* extensão no processo decisório. No caso da legitimação, entende-se que a base do processo técnico de tomada de decisões deve ser formada a partir das opiniões (certificadas) de um conjunto de especialistas. Já na questão da extensão, existe o problema de se identificar até que ponto a participação (de setores não acadêmicos/especialistas) deve interferir no processo técnico de tomada de decisões (COLLINS; EVANS, 2002).

Os autores propõem ainda o que eles chamam de uma “terceira onda” dos estudos sociais do conhecimento científico, sugerindo a construção de uma teoria normativa da expertise, pelo argumento de que se a expertise está ao alcance de todos, portanto, o processo decisório também deveria estar. Para Collis e Evans (2002), já que todas as pessoas são dotadas de grande expertise para se comunicarem por meio da linguagem e para a realização de quaisquer outras tarefas que dependem da compreensão do contexto social, e se a expertise – que tem relação com a experiência adquirida – tem alguma relação com esse contexto, ela também pode ser aplicada a contextos especializados, como nas decisões sobre ciência e tecnologia, por exemplo.

Tomando-se como exemplo o processo de tomada de decisão sobre a adesão ou não às plataformas de *e-Science*, sob o prisma da legitimação, entende-se que essa decisão deve ser *legitimada* com base na opinião de especialistas, os quais podem lançar mão de argumentos epistemológicos, em geral, salientando a importância de tais plataformas para o avanço das pesquisas científicas. Por outro lado, sob o prisma da extensão, entende-se que essa decisão deve privilegiar também um contexto sociopolítico amplo,

considerando-se os impactos ou desdobramentos tecnológicos para a sociedade derivados da produção, comercialização e do uso das plataformas.

Ganha destaque, então, o discurso da relevância social para a validação do apoio e estímulo às tecnologias emergentes de pesquisa, em um processo chamado por Simakova (2012) de “ciência governável”, ou o espaço no qual os participantes negociam noções de responsabilidade, boas práticas e resultados esperados. Inclui a ideia de uma ciência cujos rumos e decisões são abertas para um debate amplo, considerando-se as contribuições de diversos setores da sociedade, seja por participação direta ou por representação. Velho (2011, p. 143) destaca que, no plano da Sociologia do Conhecimento Científico, “a noção de ciência como socialmente construída está cada vez mais aceita, se não pelos próprios cientistas, pelos analistas de ciência, tecnologia e inovação (CTI) e pelos tomadores de decisão”. Nesse modelo de análise, adota-se uma visão da ciência como socialmente construída, cujo conteúdo está sujeito à análise sociológica, assim como seu foco, sua direção e seu uso (VELHO, 2011).

De forma geral, observou-se que as discussões que contemporaneamente emergem dos estudos sobre ciência e tecnologia também se manifestam no contexto da *e-Science*, uma vez que suas práticas demandam, em função de sua complexidade tecnológica, significativos investimentos por parte das agências financiadoras, flexibilização dos contratos, normas e regulações das relações entre países, elementos com potencial interesse público. A abertura para o público implica a necessidade de novas formas de ordenar, regular ou mesmo mediar a interação de atores e seus interesses, o uso de recursos e resultados de pesquisas nas práticas de *e-Science*, dando forma a discussões que tratem da questão da governança nesse contexto.

3.2 Gestão e organização das práticas científicas

Neste estudo, buscou-se privilegiar o entendimento dos processos de gestão e organização das práticas científicas sob a ótica da governança, ou seja, buscando-se levar em consideração a multiplicidade de atores e recursos que mobilizam ou são mobilizados por tais processos.

A governança, de acordo com Hufty (2011, p. 405, tradução nossa), refere-se a “uma categoria de fatos sociais, mais precisamente os processos de interação e tomada de decisão entre atores envolvidos em um problema coletivo orientado à criação, reforço ou reprodução de normas e instituições sociais”.

Em outras palavras, pode-se afirmar que a governança procura evidenciar as formas como os indivíduos constituintes de um determinado contexto social se organizam e interagem para tomarem suas decisões, seja para a regulação do comportamento social, seja para a criação de órgãos ou instituições para os quais será transferida a responsabilidade acerca de determinadas decisões. Observa-se que o autor não chega a fazer distinção ou referência a atores (ou organizações) específicos, ou seja, entende-se que todo e qualquer indivíduo componente de um coletivo é responsável pela governança.

Em outro trabalho, Hufty (2009) identifica ao menos três possíveis visões ou abordagens do conceito de governança: um sinônimo para governo; um quadro analítico normativo; e um quadro analítico para sistemas de coordenação não hierárquicos. Alguns elementos dessas abordagens são explorados mais à frente neste estudo apresentando-se, por fim, um enfoque direcionado à governança no plano da ciência.

Na concepção de Stoker (1998) um *governo* é caracterizado pela sua capacidade de tomar decisões e de aplicá-las, ou seja, fazer com que as mesmas sejam cumpridas. O termo governo, na teoria política anglo-americana, é aplicado “para se referir às instituições formais do estado e o monopólio destas sobre o poder coercitivo legítimo” (STOKER, 1998, p. 17, tradução nossa). Ou seja, sob uma ótica de governo, entende-se que somente as instâncias executivas, legislativas, jurídicas etc. do poder público são formalmente ‘autorizadas’ a deliberar acerca de questões sociais, políticas e econômicas num plano hierarquizado (*top-down*).

Contudo, esse mesmo autor ressalta que trabalhos teóricos sobre o tema da governança, por sua vez, “refletem o interesse da comunidade das Ciências Sociais em um padrão de mudança nos estilos de governar” (STOKER, 1998, p. 17, tradução nossa).

Nesse sentido, ganha destaque o conceito de *governança*, cujo valor de uso “reside na sua capacidade de fornecer um quadro conceitual para a compreensão de mudanças nos processos de governo” (STOKER, 1998, p. 18, tradução nossa), e que pode ser traduzido, segundo Stoker, a partir dos seguintes aspectos:

- a) governança se refere a um conjunto de instituições e atores que são concebidos para o governo e para além deste;
- b) governança permite identificar as barreiras e responsabilidades difusas para lidar com as questões sociais;
- c) governança permite identificar a dependência de poder nas relações entre instituições envolvidas em ações coletivas;
- d) governança trata de redes de atores autônomas e autogovernáveis;
- e) governança reconhece a capacidade de realizações, as quais não dependem do poder de comando dos governos ou da autoridade governamental para acontecerem. Considera o governo a partir de sua capacidade de usar novas ferramentas e técnicas para conduzir ou guiar as ações (STOKER, 1998, p. 18, tradução nossa).

Rhodes (2007, p. 1246, tradução nossa), por sua vez, dá ênfase ao conceito de governança como “o ato de governar com as, e por meio das, redes”. Segundo o autor, essa noção permite integrar o qualificador *rede* ao governo, criando o conceito de redes de governança. Tais redes derivam da fragmentação governamental, ou o esvaziamento do estado, decorrente das reformas políticas dos anos 1980 e, tomando-se como base a experiência do governo britânico, derivam também da baixa aceitação aos modelos de governo hierárquicos. As redes, em contrapartida, contam com a existência de normas e valores compartilhados, e têm isso como uma vantagem de negociação e incentivo ao comportamento cooperativo (RHODES, 2007).

Uma fragmentação prévia, que fortalece agrupamentos posteriores descentralizados, com a conseqüente formação das redes, é uma característica específica da governança. Nesse plano, ocorre também a formação de redes de políticas, as quais Rhodes definiu, com base na análise do cenário governamental Britânico, como:

conjuntos de relações [redes] institucionais formais e informais entre o governo e outros atores estruturados em torno de interesses compartilhados acerca da formulação e implementação de políticas públicas. Essas instituições [esses

conjuntos] são interdependentes. As políticas emergem da barganha entre as redes (RHODES, 2007, p. 1244, tradução nossa)

Pela ótica da administração pública e das políticas públicas, Rhodes (2007) resgata o conceito de governança de um trabalho anterior de sua autoria (RHODES, 1997), e o sintetiza a partir dos seguintes elementos:

- a) existência de interdependência entre organizações. A governança é mais ampla que o governo, cobrindo atores não-estatais. Deslocar/alterar os limites de atuação do estado implica tornar difusos os limites entre os setores público, privado e voluntário;
- b) existência de interações contínuas entre os membros da rede, alimentadas pela necessidade de se intercambiar recursos e negociar objetivos e propósitos comuns;
- c) existência de interações como em um jogo, enraizadas na confiança e reguladas por regras do jogo, as quais são negociadas e acordadas pelos participantes da rede;
- d) existência de um grau significativo de autonomia em relação ao estado. As redes não prestam contas ao estado; elas se auto-organizam; embora o estado não ocupe uma posição soberana ou privilegiada, ele tem a capacidade de orientar as redes indiretamente.

A partir das conceituações observadas, torna-se evidente a importância das redes para a concretização de um cenário de governança, sendo que tais redes caracterizam o conjunto de relações interorganizacionais e

“compreendem organizações que intercambiam recursos (por exemplo, financeiros, de informação, expertise) para atingirem seus objetivos, para maximizarem sua influência na busca por resultados e para evitarem a dependência de outros competidores” (RHODES, 1996, p. 658, tradução nossa).

Para Stoker (1998), ocorre a formação de redes autônomas e autogovernáveis, as quais se voltam para as políticas de uma comunidade específica, sendo compostas de “atores e instituições que ganham a capacidade de agir pela combinação de seus recursos, habilidades e propósitos em uma coalisão de longo-termo: um regime” (STOKER, 1998, p. 23, tradução nossa), em uma base informal, porém sustentável, e sem a necessidade de uma estrutura de comando. O autor destaca, contudo, o problema residente no fato de que as redes podem apresentar certo grau de exclusividade, pois podem decair na dificuldade

de representarem apenas os interesses dos seus membros, negligenciando questões mais abrangentes, de interesse público, por exemplo.

Ao explorar o conceito de governança no contexto da ciência, Fuller (1999) esclarece que o mesmo pode ser usado em lugar de políticas de ciência, para resolver o conflito que ocorre entre “os interesses sociais, tipicamente em matéria de tecnologia, e que têm sérios impactos na vida das pessoas, no meio-ambiente ou econômicos” – geralmente condensados nas políticas – e os interesses voltados à produção do conhecimento científico das pessoas envolvidas nesse processo. Fuller (1999) ainda expressa a preferência pelo termo governança, em detrimento de governo (da ciência), em função da informalidade expressada pelo primeiro, em analogia às ações políticas ou de prestação de contas na ciência, tais como eleições, referendos, ensaios e auditorias.

A pesquisa sobre métodos de geoengenharia aplicados na compensação das mudanças climáticas decorrentes do efeito estufa também é objeto de estudo de Parson e Keith (2013) sob a ótica da governança. Os autores argumentam que esse campo ainda não obteve sucesso no desenvolvimento de um sistema de governança, particularmente no que se refere a duas questões básicas. A primeira delas diz respeito à dificuldade de se definirem os limites entre grandes e pequenas intervenções no direcionamento das pesquisas, considerando-se que as primeiras demandam maior controle. A segunda questiona se a autorregulação científica é adequada para pesquisas de pequena escala ou se estas demandam regulação governamental e, em caso afirmativo, qual deveria ser a relação entre os processos regulatórios e científicos (PARSON; KEITH, 2013).

Em outra forma de exploração da abordagem da governança na ciência, Simakova (2012) toma o exemplo das pesquisas em nanotecnologia, a partir de um estudo junto a cientistas, porta-vozes de universidades e escritórios de transferência tecnológica da Costa Oeste dos Estados Unidos envolvidos nesse tema, considerando que

os participantes do meio acadêmico, da política e da indústria idealizam o campo nano[tecnológico] como relevante para a sociedade, devido à sua produção institucionalizada, suas possibilidades de comercialização e governança da pesquisa no campo de estudo (SIMAKOVA, 2012, p. 205, tradução nossa).

Segundo a autora, os documentos fundacionais dessa área de estudo costumam apontá-la como um diferencial de competitividade – nutrindo uma ideia de relevância para a sociedade – contudo, está longe de existir uma visão homogênea sobre o tema (SIMAKOVA, 2012). As noções de relevância das nanotecnologias para a sociedade diferenciam-se na medida em que deliberações acerca de implicações sociais vão sendo

definidas, e têm sido construídas, no contexto de análise da autora, sobre os questionamentos dos benefícios sociais das nanotecnologias para os consumidores americanos.

Nesse contexto, é possível perceber que “a relevância social de campos de pesquisa similares ao das nanotecnologias podem ser vistos como parte de um discurso mais amplo e altamente visionário acerca de iniciativas tecnocientíficas emergentes” (SIMAKOVA, 2012, p. 606-607, tradução nossa), argumento que se poderia também aplicar aos discursos sobre o estímulo às práticas de *e-Science*. Segundo a autora,

análises sobre políticas acadêmicas e de pesquisa, muitas vezes assumem a existência independente de campos de pesquisa emergentes, como o ‘nano’, que possuem trajetórias evolutivas particularmente descritíveis de aceitação ou rejeição social. Desse ponto de vista, o trabalho analítico desenvolve-se, em grande parte, sobre a acumulação e classificação de declarações (prospecções, críticas, descrições), que supostamente nos dizem algo relacionado ao ‘campo’ em questão. Essa cultura despolitiza o empreendimento científico, assumindo que as ambivalências permanecem em contexto (como o retorno do público ou a polaridade das implicações sociais), em vez de colocar ‘nano’ como um objeto constituído de acordo com as expectativas do público em geral (SIMAKOVA, 2012, p. 607, tradução nossa).

A autora destaca ainda, com base nas entrevistas realizadas em seu estudo, que o uso de argumentos que procuram ressaltar aspectos “qualitativamente diferenciados” da área de pesquisa – especialmente o de constituição de novos paradigmas científicos – podem estimular a liberação de recursos junto aos organismos de financiamento. Dessa forma, “os participantes se envolvem na tática de persuadir uns aos outros, e a seus financiadores, no novo status paradigmático de sua ciência, e também demonstrando sua relevância para a sociedade (SIMAKOVA, 2012, p. 615, tradução nossa).

Mais especificamente, na análise da governança em *e-Science*, outros elementos podem ser significativos. Tomando-se como base o trabalho de Stoker (1998), tais elementos dizem respeito à configuração de diversos centros de poder, derivados das práticas de *e-Science*, e às relações entre tais centros, compreendendo níveis locais, regionais, nacionais e supranacionais. É importante definir quais são os elementos que permitem avaliar a centralização ou a distribuição do poder, entre eles, o número de participantes no projeto, a quantidade de recursos – financeiros ou tecnológicos – investidos, o volume de interações ou de resultados decorrentes das atividades de pesquisa. Nesse sentido, Stoker (1998) argumenta que o investimento no compartilhamento é determinado não apenas pela quantidade de recursos dos

participantes, mas também pelas regras do jogo e pelo contexto da troca, que se definem em sistemas de governança.

As aberturas (*openness*) e fechamentos (*enclosures*) dos sistemas de colaboração em *e-Science* vinculam-se, assim, a condições como conformidade no estabelecimento de regras e justificação das regras pela referência às crenças compartilhadas, que, ao final, expressam o claro consentimento da subordinação a relações de hierarquia e poder (BEETHAM, 1991 apud STOKER, 1998). A construção de padrões de interoperabilidade, o estabelecimento de regras de compartilhamento e de benefício de resultados derivados de projetos de *e-Science* expressam referência a tais condições.

3.3 Fluxos e circulação da informação nas práticas científicas

A exploração ou mobilização de recursos informacionais por parte dos cientistas e pesquisadores ocorre em meio a redes tecnoeconômicas, como um “conjunto coordenado de atores heterogêneos que interagem com mais ou menos sucesso para desenvolver, produzir, distribuir e difundir métodos para a geração de produtos e serviços” (CALLON, 1991, p. 133, tradução nossa).

A visão de Callon, e de outros autores da corrente de estudos da ciência conhecida como relativista, que sucede e complementa a perspectiva sociológica estruturalista por meio da qual os atores só podem ser compreendidos quando inseridos em (estudados a partir de) um espaço comum, construído por eles próprios. Callon procura agregar a visão dos economistas, levando em consideração os chamados elementos *intermediários* – textos (literatura científica), artefatos técnicos (máquinas e instrumentos científicos), seres humanos (conhecimentos e habilidades) e recursos financeiros – considerados atores que colocam, por sua vez, outros intermediários em circulação (CALLON, 1991, p. 141, tradução nossa).

Tais atores, humanos e não humanos, passam a interagir nas chamadas redes sociotécnicas, objetos de análise da Teoria Ator-Rede (TAR ou ANT, do inglês *Actor-Network Theory*). Conforme esclarece Latour (2006), outro autor de destaque nos estudos relativistas, a TAR tem aplicabilidade em estudos em que as coisas se transformam rapidamente, nos quais as fronteiras são terrivelmente fluidas, tais como os estudos organizacionais, da informação, marketing, ou nos estudos de ciência e tecnologia. Tais características, expressas pela TAR, são também observáveis no contexto da *e-Science*,

dado o seu 'poder' de colocar cientistas e recursos em circulação e de formar grupos ou instituições a seu propósito, garantindo-lhe um *status* de ator central para o desenvolvimento atual da ciência.

Nesse sentido, amplia-se o campo de reflexão acerca da *e-Science*, buscando-se compreender o seu papel na redistribuição e na reconfiguração de atores e recursos presentes no campo da ciência contemporânea. Para tanto, faz-se importante o aprofundamento em alguns conceitos pertinentes ao estudo de redes e alguns específicos da TAR, tais como mediação, tradução, inscrições, entre outros, tal como segue.

No campo da produção e reprodução social, Latour (2012) considera os conceitos de intermediários e mediadores. Para o autor, um intermediário é aquele/aquilo que transporta significado, forçando-o sem transformá-lo. Tudo o que é usado/tomado para constituir um intermediário serve-lhe igualmente para sua definição/expressão. Mesmo que este seja constituído de várias partes, ele será sempre visto/compreendido como único. Os mediadores, pelo contrário, não podem ser vistos como únicos. Eles transformam, traduzem, distorcem e modificam o significado ou os elementos que supostamente carregam (LATOUR, 2012). A exploração desses conceitos nesse momento faz-se relevante uma vez que permite a compreensão da interação dos atores sociais no plano das redes sociotécnicas.

Nesse contexto, a informação tem papel crucial, servindo como um elemento estabelecedor da relação entre dois mundos ou duas realidades supostamente distintas. Latour (2010) descreve o conceito de informação como relação, recorrendo ao exemplo de amostras de plantas, espécimes, desenhos etc. que podem ser objetos referentes de uma determinada realidade, permitindo ao centro acumular conhecimentos sobre um lugar que até então ele não podia representar (periferia). De acordo com o autor, a informação, ao invés de um signo, é

[...] uma *relação* estabelecida entre dois lugares: o primeiro, que se torna uma periferia; e o segundo, que se torna um *centro*, sob a condição de que entre os dois circule um *veículo* que denominamos muitas vezes de forma, mas que, para insistir em seu aspecto material, eu chamo de *inscrição* (LATOUR, 2010, p. 40, grifos do autor).

Tal relação permite que o centro negocie o que pode retirar da periferia, criando representações específicas do mundo periférico de acordo com interesses específicos provenientes do centro.

Dessa forma, “[...] a informação [...] é sim uma relação muito prática e muito material entre dois lugares, o primeiro dos quais negocia o que deve retirar do segundo, a fim de mantê-lo sob sua vista e agir à distância sobre ele” (LATOURE, 2010, p. 42). Essa concepção abre espaço para a ideia de que o controle intelectual, o domínio erudito, não se exerce diretamente sobre os fenômenos, mas sim sobre as inscrições que lhes servem de veículos, sob a condição de circular continuamente, e nos dois sentidos, através de redes de transformações – laboratórios, instrumentos, expedições, coleções (LATOURE, 2010, p. 51).

Ainda sob a perspectiva da TAR, faz-se presente o conceito de *tradução*, que significa deslocar objetivos, interesses, dispositivos, seres humanos por meio das redes. Esse movimento implica um desvio de rota, invenção de um elo que antes não existia e que de alguma maneira modifica os elementos imbricados. As cadeias de tradução referem-se ao trabalho pelo qual os atores modificam, deslocam e transladam os seus vários e contraditórios interesses (FREIRE, 2006). Para Latour (2012), esse conceito deriva da compreensão do conceito presente na raiz etimológica do termo social (*socius*), que se refere a alguém que segue outro alguém, um seguidor, um associado. Não se trata de um ator entre muitos e tampouco uma força que transporta todos os atores, mas uma conexão que transporta, à qual o autor denomina tradução, sendo que as redes tomam forma a partir dos traços deixados por esse movimento de tradução (LATOURE, 2012).

Para Callon (1991), a noção de tradução implica definição, sendo que tais definições estão inscritas em intermediários das mais diferentes formas. O termo tradução não pode ser usado de forma genérica, necessitando sempre uma associação ao meio/material no qual está inscrito. Em todo caso, a operação elementar de tradução é triangular, envolvendo um tradutor, algo que é traduzido e um meio no qual a tradução é inscrita.

Em linhas gerais, para além de uma visão funcional estruturalista do social, que privilegia o olhar sobre os indivíduos ou atores estritamente agrupados em estruturas sociais rígidas e imutáveis, a abordagem relacional implica uma tentativa de se organizar o mundo a partir de uma dinâmica que lhe é inerente, que não se resume a regularidades que possam ser traduzidas em leis científicas (FONTES, 2012).

Na representação social dos indivíduos em rede,

“a margem de decisão do indivíduo, no entanto, está sempre presa à distribuição do poder, à estrutura da dependência e das tensões no interior do grupo. Seja o

que for decidido, isto o afastará de uns e o aliará a outros” (MARTELETO, 2000, p. 79).

Segundo a autora, é esta “interdependência funcional” dos indivíduos em relação (ou o conjunto destes) que permite a compressão do social.

É possível considerar a existência e manutenção dos jogos de poder também no contexto da *e-Science*, uma vez que a alocação de recursos financeiros e tecnológicos para pesquisa, a adesão de pesquisadores ou instituições a grupos de colaboração e as condições de acesso a dados e ferramentas são mediados por meio de redes que compreendem atores com maior ou menor poder de negociação, que pode ou não ser revertido em resultados favoráveis aos seus detentores. Faz-se necessária então a compreensão desses processos de interações e negociações por meio das redes colaborativas desencadeadas pela *e-Science*, sendo essa uma das possíveis condições de adesão a esse movimento.

4 PRÁTICAS DE COLABORAÇÃO EM E-SCIENCE: O CASO DO CERN

A partir deste capítulo são descritos os passos metodológicos e resultados da pesquisa de campo realizada neste estudo.

Com base no exame de qualificação desta pesquisa realizado em agosto de 2013, definiu-se que seria tomado como estudo de caso a **Organização Europeia para a Pesquisa Nuclear** (CERN, do Francês *Conseil Européen pour la Recherche Nucléaire*, e do inglês *European Council for Nuclear Research*), e mais especificamente a experiência brasileira na colaboração em *e-Science* com o CERN. Esta é uma experiência reconhecidamente característica da *e-Science*, em função das grandes quantidades de dados geradas em seus experimentos, além da formatação de suas colaborações em *grids*, a partir de unidades distribuídas. Um dos experimentos associados a esse projeto, o LHCb, congrega uma colaboração de cerca de 700 cientistas de 52 diferentes universidades e institutos (CERN, 2012).

Nesse sentido, optou-se pela realização de entrevistas junto a um grupo de pesquisadores de instituições brasileiras – nas quais estão implementados *grids* conectados ao CERN – que estivessem ou já estiveram envolvidos diretamente em colaborações do CERN, investigando como se dá o processo de colaboração em suas interações com instituições e com outros pesquisadores, além do ponto de vista desses pesquisadores sobre os seus papéis ou suas funções nesse contexto. Como complementação, optou-se também pela entrevista com um segundo grupo de pesquisadores não relacionados diretamente ao CERN, mas com atuação em outros projetos ou iniciativas relacionadas à *e-Science*, buscando-se captar suas experiências com a *e-Science* ou com estudos sobre pesquisa colaborativa, sob a perspectiva do tratamento dos dados ou da infraestrutura computacional.

Foram entrevistados oito pesquisadores, sendo este conjunto delimitado ao momento em que se observou a saturação teórica dos conceitos abordados. Cinco desses pesquisadores são da área de Física e três são das áreas de Ciências da Computação e Engenharias, identificados a seguir:

- a) **Alberto Franco de Sá Santoro**⁹: Físico, atua como Professor Titular da Universidade do Estado do Rio de Janeiro (Uerj). Membro da colaboração CMS do CERN;
- b) **Daniel de Oliveira Tavares**¹⁰: Engenheiro de Automação, atua como Engenheiro Eletricista do Laboratório Nacional de Luz Síncrotron (LNLS). Atua em iniciativas para desenvolvimento de tecnologias Open Hardware;
- c) **Jesús Pascual Mena Chalco**¹¹: Engenheiro de Sistemas, atua como Professor Adjunto da Universidade Federal do ABC (UFABC). Pesquisador do Núcleo de Pesquisa em eScience da USP;
- d) **Leandro Salazar de Paula**¹²: Físico, atua como Professor Associado da Universidade Federal do Rio de Janeiro (UFRJ). Membro da colaboração LHCb do CERN;
- e) **Marco Aurelio de Lisboa Leite**¹³: Físico, atua como Físico da Universidade de São Paulo (USP). Membro da colaboração ATLAS do CERN;
- f) **Roberto Marcondes Cesar Jr**¹⁴: Cientista da Computação, atua como Professor Titular da Universidade de São Paulo (USP). Diretor do Núcleo de Pesquisa em eScience da USP;
- g) **Ronald Cintra Shellard**¹⁵: Físico, Pesquisador Titular do Centro Brasileiro de Pesquisas Físicas (CBPF). Membro da colaboração Pierre Auger. Atuou como membro da colaboração Delphi no CERN;
- h) **Sérgio Ferraz Novaes**¹⁶: Físico, atua como Professor Titular da Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp). Membro da colaboração CMS do CERN.

Para orientação no processo de realização das entrevistas, foi elaborado um tópico guia (Apêndice A) com base nos objetivos específicos definidos para esta pesquisa, o qual envolveu, essencialmente, os seguintes conceitos ou temas:

⁹ <http://lattes.cnpq.br/2474284374033405>.

¹⁰ <http://lattes.cnpq.br/0678552601071354>.

¹¹ <http://lattes.cnpq.br/4727357182510680>.

¹² <http://lattes.cnpq.br/7666576260448993>.

¹³ <http://lattes.cnpq.br/1055961687614089>.

¹⁴ <http://lattes.cnpq.br/2240951178648368>.

¹⁵ <http://lattes.cnpq.br/9624642322598710>.

¹⁶ <http://lattes.cnpq.br/7924815022468819>.

- a) **Atuação no CERN:** questões relativas às formas de interação com o CERN, tempo de interação e questionamentos sobre o uso de plataformas/recursos da e-Science no trabalho com o CERN;
- b) **Condições de participação:** condições de acesso e uso de plataformas e recursos, incluindo dados; questões sobre a agenda de pesquisa ou objetivos da colaboração; e questões sobre investimento financeiro ou de infraestrutura;
- c) **Resultados:** indagações a respeito dos procedimentos para publicação e divulgação de resultados de pesquisa, avaliação e atribuição de créditos pela colaboração; indagações sobre condições ou orientações específicas acerca de propriedade intelectual, patenteamento de produtos ou processos, ou licenciamento de *software* derivado da colaboração;
- d) **Cenário Brasil:** indagações acerca de benefícios ou barreiras/dificuldades decorrentes da colaboração e o uso de plataformas ou recursos da *e-Science* junto ao CERN; indagações sobre as condições de acesso a recursos de pesquisa na área de Física de Altas Energias no Brasil, em iniciativas não relacionadas ao CERN.

Para as entrevistas com os pesquisadores que não atuam diretamente nas colaborações do CERN foram feitos alguns ajustes no tópico guia, porém sem fugir muito dos grandes temas de interesse, a exemplo das condições e resultados de participação em colaborações ligadas à *e-Science*.

As entrevistas foram realizadas no período de 18/10/2013 a 10/12/2013, sendo que apenas uma das entrevistas foi feita no mês de janeiro de 2014, em função de alterações não previstas em calendário oficial da instituição de um dos pesquisadores. Sete das oito entrevistas foram realizadas presencialmente e uma delas ocorreu à distância por meio do aplicativo Google Hangouts¹⁷. O áudio de todas as entrevistas foi gravado, com anuência dos entrevistados, e os registros encontram-se preservados em formato digital.

Para a análise das entrevistas não foi realizada a transcrição completa de cada entrevista, realizando-se, por outro lado, a seleção dos principais trechos de interesse, de

¹⁷ <http://www.google.com/hangouts/>.

acordo com as categorias de análise e os conceitos-chave elencados a partir dos questionamentos iniciais da pesquisa e do referencial teórico, e que também nortearam a aplicação do tópico guia durante as entrevistas. Essa seleção gerou um quadro de análise o qual é apresentado como Apêndice B, ao final deste trabalho. Eventualmente foram selecionados alguns trechos mais significativos durante a análise para serem transcritos tal qual foram expressos pelos pesquisadores. Tais trechos terão destaque como citações diretas na seção de resultados neste trabalho. Os dados coletados nas entrevistas foram confrontados e complementados, durante a análise, com documentação disponível no site oficial do CERN¹⁸, além de informações disponíveis nos sites das colaborações tomadas como exemplo, descritas no item a seguir, e nos sites das instituições às quais os pesquisadores entrevistados estão vinculados.

O próximo item apresenta informações que contextualizam o tipo de atuação do CERN. Em seguida, são apresentados os resultados das entrevistas acompanhados de análise que foi norteada pelas indagações de pesquisa iniciais, bem como por elementos de destaque identificados no referencial teórico trabalhado nos capítulos anteriores.

4.1 Sobre o CERN

O CERN se estabeleceu como organização no ano de 1954, na cidade de Genebra, na Suíça, em uma região de fronteira com a França, tendo como objetivo inicial o estudo do interior do átomo, de onde deriva o termo ‘nuclear’ presente no nome da organização. Atualmente, os estudos da Física para a compreensão da matéria vão além da preocupação da compreensão do núcleo do átomo, sendo que a principal área de interesse do CERN é a física de partículas, ou o estudo dos constituintes fundamentais da matéria e das forças que agem sobre ela (CERN, c2014a).

Com o intuito de estudar os constituintes básicos da matéria, ou a física de partículas elementares, no CERN são realizados experimentos por meio dos quais as partículas colidem a uma velocidade próxima à velocidade da luz, dando evidências aos cientistas de como tais partículas interagem e gerando *insights* sobre as leis fundamentais da natureza. Para que tais experimentos possam ser colocados em prática, o CERN

¹⁸ <http://home.web.cern.ch/>.

trabalha no desenvolvimento e aperfeiçoamento de instrumentação científica, correspondente a processos e instrumentos ou equipamentos técnicos, que englobam:

- a) aceleradores de partículas: criam feixes superenergizados (ou de altas energias) de partículas para que as mesmas sejam colocadas em colisão;
- b) detectores: ‘observam’ e gravam os resultados das colisões.

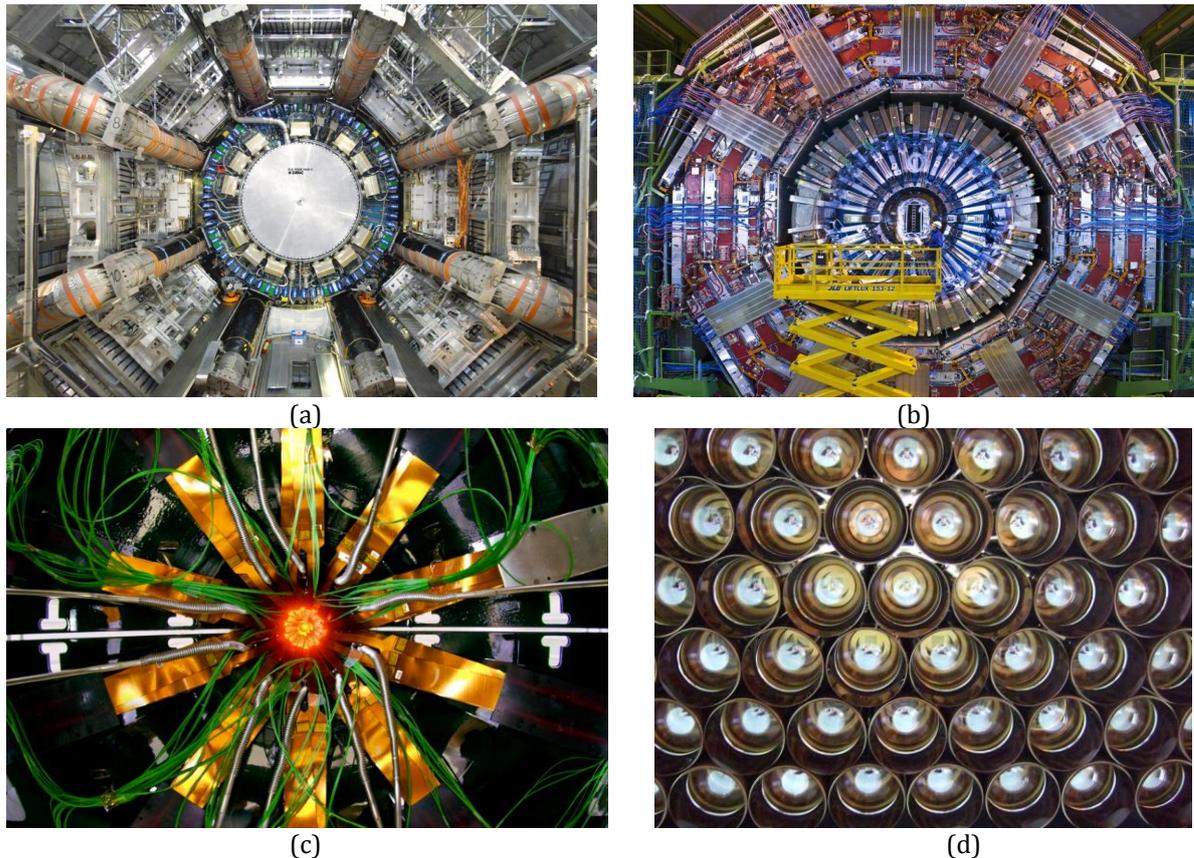
O principal desses aceleradores é o LHC, do inglês *Large Hadron Collider*, correspondente a um túnel de 27 km de extensão, em formato circular/anel, composto de ímãs supercondutores que direcionam os feixes ao longo do túnel, além de uma série de estruturas de aceleração para aumentar a energia das partículas ao longo do caminho.

Os detectores são estruturas responsáveis pela coleta de dados derivados das colisões que ocorrem no LHC. São compostos de uma série de subdetectores, cada um projetado para permitir a observação de propriedades particulares ou de tipos específicos de partículas. Incluem dispositivos de rastreamento que revelam a trajetória de uma partícula; calorímetros, que retêm, absorvem e medem a energia de uma partícula; e detectores de identificação de partículas, que usam uma variedade de técnicas para definir a identidade de uma partícula (CERN, c2014b). Dentre os principais detectores estão o ATLAS, CMS, ALICE e o LHCb (Figura 1). As colaborações do CERN, as quais se iniciaram por volta dos anos 1990, foram formadas para permitir o estudo dos dados gerados por cada um dos experimentos conduzidos nos detectores, os quais correspondem a diferentes ‘tipos de física’ ou fenômenos de interesse.

O experimento **ALICE**, acrônimo para *A Large Ion Collider Experiment*, ou “um grande experimento de colisão de íons”, teve início em 1993, com o objetivo de construção de um detector de íons pesados para a exploração de interações núcleo-núcleo geradas no LHC. Duas instituições brasileiras têm pesquisadores que participam desse experimento: Universidade Estadual de Campinas (Unicamp) e Universidade de São Paulo (USP).

Com base no experimento **ATLAS** (*A Toroidal LHC ApparatuS*) são conduzidas investigações acerca do bóson de Higgs, do estudo de dimensões extra e de partículas que poderiam permitir a formação de matéria escura. Instituições brasileiras com pesquisadores participantes do ATLAS: Universidade Federal do Rio de Janeiro (UFRJ); Universidade de São Paulo (USP); Universidade Federal de Juiz de Fora (UFJF); e Universidade Federal de São João del Rei (UFSJ).

Figura 1. Detectores presentes no CERN



Legenda: (a) ATLAS; (b) CMS; (c) ALICE; (d) LHCb

Fonte: Disponível em: <http://home.web.cern.ch/about/experiments>.

O experimento **CMS** (*Compact Muon Solenoid*) agrega os mesmos objetivos científicos do experimento ATLAS, contudo utiliza diferentes soluções técnicas e um distinto design do sistema de ímãs. Instituições brasileiras que participam do CMS: Universidade Estadual Paulista Júlio de Mesquita Filho (Unesp); Centro Brasileiro de Pesquisas Físicas (CBPF); e Universidade do Estado do Rio de Janeiro (Uerj).

O experimento **LHCb** (*LHC beauty*) procura investigar as pequenas diferenças entre matéria e antimatéria, estudando um tipo de partícula chamada de "quark bela" ou "*b quark*". Instituições brasileiras que participam do experimento LHCb: Centro Brasileiro de Pesquisas Físicas (CBPF); Universidade Federal do Rio de Janeiro (UFRJ); e Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

4.2 A colaboração em *e-Science* no CERN

Nesta seção é feito um relato sobre como ocorrem as colaborações no CERN e iniciativas de *e-Science* similares, a partir com base nas entrevistas realizadas na pesquisa

empírica. A sequência da análise foi organizada de acordo com a ordenação dos temas ou assuntos elencados no roteiro das entrevistas.

4.2.1 Sobre a atuação no CERN

Dos entrevistados que atuam ou colaboram diretamente com o CERN, todos apontaram que iniciaram a colaboração já por volta do início da década de 1990, sendo que as formas de colaboração ou atuação envolviam (ou envolvem) primordialmente a realização de análises de dados, a atuação em serviços da colaboração – como a identificação de partículas e aprimoramento de *software* –, desenvolvimento de detectores e sistemas a estes associados, além da preparação e atualização dos mesmos visando coletas de dados futuras, participação em reuniões presenciais e por videoconferência, além de ações locais de construção e teste de sistemas de detecção e de aceleração e de manutenção de infraestrutura computacional para atendimento à colaboração.

No caso dos pesquisadores entrevistados que não atuam diretamente no CERN, um deles indicou participação em projetos ou ações de mensuração e avaliação da participação de pesquisadores em colaborações científicas. Outro indicou a participação em iniciativas para o desenvolvimento de instrumentação científica – eletrônica para instalações/laboratórios científicos, incluindo circuitos integrados, aceleradores etc. – com base em licenças *Open Hardware*, desenvolvidas pelo CERN. E o terceiro mencionou a participação em núcleo de pesquisa em Computação e Matemática, que realiza parceria com grupos ou projetos de pesquisa de outras áreas voltados para a pesquisa com a utilização de ferramentas da *e-Science*, em especial nas áreas de Biologia e Medicina, no trabalho de integração de bancos de dados de genoma humano.

4.2.2 Sobre as condições de participação no CERN ou em suas colaborações

O CERN funciona como um grande laboratório, com funções científicas e administrativas/burocráticas, tal como laboratórios ou grupos menores que também contam com escritórios de projeto e/ou setores que garantam os recursos essenciais para o andamento das pesquisas, tais como instrumentos, água, energia, material de consumo etc.

Como instância máxima de deliberação, o CERN conta com um Conselho Superior, composto por um representante de cada país participante nas colaborações, tendo direito a voto apenas os representantes dos países identificados como membros ou associados ao CERN. Essa associação implica o pagamento de uma cota, a qual é proporcional ao Produto Interno Bruto (PIB) de cada país associado. No caso do Brasil, o marco das negociações para ingresso do país como membro associado, de acordo com Ronald Shellard (informação verbal)¹⁹, data dos anos de 2008 e 2009, a partir do interesse no aumento do número de países não europeus associados ao CERN. Com isso, montou-se uma comissão para avaliar a possibilidade de ingresso do Brasil que culminou na formulação de uma carta de intenção e, em uma reunião do Conselho do CERN em 12/12/2013, foi formalizado o convite para o Brasil tornar-se membro associado do CERN (CHADE, 2013). A concretização dessa associação, contudo, depende do aporte de recursos orçamentários que deve ser aprovado pelo Congresso Nacional Brasileiro (recursos viriam diretamente do orçamento da União), como sinalizou Sergio Novaes (informação verbal)²⁰. Esse aporte está estimado em 10 milhões de dólares americanos (CHADE, 2013). As implicações dessa associação para o Brasil são discutidas mais à frente na seção 4.2.5.

Na instância científica, o CERN é dividido em colaborações (ALICE, ATLAS, LHCb etc.), sendo que cada uma delas é plenamente autônoma em relação ao Conselho Superior e em relação às demais nas decisões que dizem respeito aos objetivos ou interesses científicos da colaboração. O CERN, nesse sentido, desempenha apenas um papel fiscalizador das colaborações, mas sem deliberar sobre suas ações.

De acordo com o entrevistado Leandro de Paula (informação verbal)²¹, para se criar uma colaboração é necessário que um cientista faça uma proposta de experimento para que então se avalie se a mesma é cientificamente relevante, se ela é tecnicamente viável e se a mesma se sustenta financeiramente.

Cada colaboração possui um estatuto – chamado de ‘constituição’ – e é composta por diferentes conselhos (ou *boards* – *Editorial Board*, *Executive Board*, *Collaboration Board* etc.), sendo que cada um deles discute ou delega sobre diferentes assuntos, por exemplo, recursos (para participação em conferências ou para alocação a longo prazo),

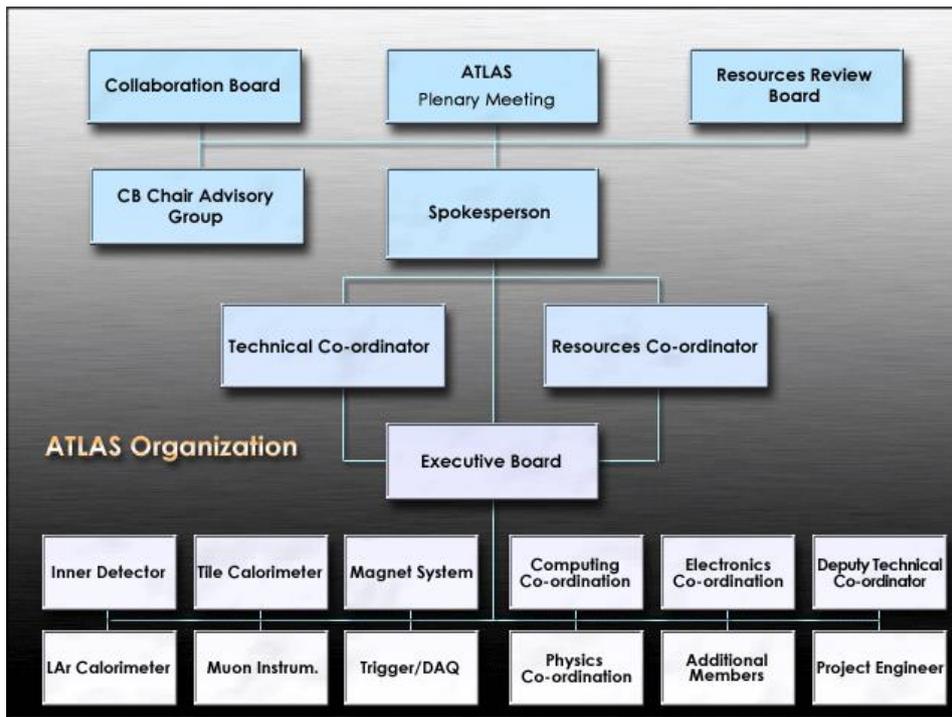
¹⁹ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

²⁰ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

²¹ Entrevista concedida (PAULA, 2013). Físico, Professor Associado da UFRJ.

preservação dos dados etc. Cada colaboração possui também uma espécie de porta-voz (*spokesman*) o qual é eleito por voto de todas instituições participantes. Na Figura 2, apresenta-se o organograma de uma das colaborações do CERN (ATLAS) e sua formatação em *boards*.

Figura 2. Exemplo de organograma de colaboração do CERN



Fonte: ATLAS Experiment. *eTour*. c2013. Disponível em:
 <http://www.atlas.ch/etours_intro/images/organigram.gif>. Acesso em: jan. 2014.

Leandro de Paula (informação verbal)²² destacou que a formação de cada conselho, não é afetada pelo volume de investimentos – financeiros, de infraestrutura e outros exemplos a serem apresentados mais à frente –, mas instituições que aplicam um volume maior de recursos tendem a exercer pressão para que seus integrantes assumam determinados cargos de direção, a exemplo da Universidade de Oxford e seu interesse na direção de um determinado detector.

Sobre a organização dos atores que interagem no contexto da pesquisa no CERN, é importante considerar também a sua dimensão e as implicações desta na distribuição e atendimento a todos esses atores e no entendimento homogêneo dos objetivos da colaboração. Sobre essa questão, Ronald Shellard (informação verbal)²³ explicou que cada

²² Entrevista concedida (PAULA, 2013). Físico, Professor Associado da UFRJ.

²³ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

experimento do CERN agrega um número muito grande de cientistas e seria inviável física e economicamente colocar todos esses cientistas no mesmo local ao mesmo tempo, por isso o trabalho remoto.

Organiza-se o trabalho para funcionar remotamente e com isso criam-se ferramentas para viabilizar essa organização. Da mesma forma, se juntar toda a capacidade computacional necessária ao processamento dos dados gerados no CERN, ela certamente não caberia no espaço do laboratório.

Quanto às hierarquias, vale destacar que as mesmas são pensadas com o intuito de se organizar o trabalho, mas não para a exploração dos resultados práticos. Em muitos casos, não existem tecnologias já desenvolvidas ou soluções prontas que permitam a realização do trabalho remoto – e nos casos em que há soluções o custo é proibitivo, por isso são montadas frentes para desenvolver alternativas, sendo o *open hardware* uma delas. A temática do *open hardware* ganhará destaque mais à frente neste trabalho.

De acordo com os entrevistados, há diferentes possibilidades de atuação no âmbito das colaborações, incluindo investimentos financeiros, trabalho técnico e científico, investimento em infraestrutura etc.

Para países associados ou não ao CERN, cada colaboração requer o pagamento de uma taxa/contribuição anual, paga por cada pesquisador, chamada de Manutenção e Operação (M&O), usada para cobrir os custos operacionais dos experimentos, como água, eletricidade etc., sendo que os custos são rateados entre as instituições. O rateio é proporcional ao número de pessoas de cada instituição atuantes na colaboração, sendo que pesquisadores-autores, cujos nomes aparecem nas publicações em nome da colaboração, correspondem a um valor cheio (100%), pós-doutorandos 75% e estudantes não pagam. Essa contribuição também garante, como afirmou Alberto Santoro (informação verbal)²⁴, o acesso a um *bureau* (escritório/sala) no CERN para os pesquisadores-autores, com acesso a computador, telefone, endereço de *e-mail* do CERN e material de consumo.

Na Tabela abaixo, é possível visualizar um resumo do número de pesquisadores brasileiros que participam de colaborações no CERN. Tais dados alimentam um relatório destinado a um consórcio brasileiro para permitir o financiamento das participações na colaboração com o CERN, o consórcio denominado Rede Nacional de Física de Altas

²⁴ Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

Energias (RENAFAE) foi instituído pela Portaria n. 321 de 28 de maio de 2008 do Ministério da Ciência e Tecnologia e Inovação²⁵, para permitir, além de outros objetivos, o financiamento dos custos de Manutenção e Operação (M&O).

Tabela. Autores brasileiros pagantes da taxa de Manutenção e Operação no CERN

Experimento	2010	2011	2012	2013	2014
ALPHA	2	2	2	2	2
ALICE	6	7	5	5	8
ATLAS	10	10	10	11	11
CMS	17	17	17	22	31
LHCb	12	12	18	18	20
Total	47	48	52	58	72

Fonte: Ministério da Ciência e Tecnologia (Brasil). Rede Nacional de Física de Altas Energias. *ATA do Workshop Anual da RENAFEA e Reunião do CTC*. 19 jun. 2013. Disponível em: <http://www.cbpf.br/~renafae/atas/ATA_wshop-2013.pdf>. Acesso em: 1 set. 2013.

Outra forma de colaboração se dá pela realização de trabalho técnico ou científico, seja pela construção de protótipos ou na melhoria da instrumentação científica, na análise dos dados derivados dos experimentos, ou mesmo na proposição de novas ideias ou experimentos. Marco Leite (informação verbal)²⁶ argumentou que cada pesquisador tem certa liberdade ou autonomia para escolha e realização de suas análises, mas deve buscar sempre demonstrar para a colaboração que a análise é consistente, fundamentada e, sendo um tópico de interesse, possivelmente atrairá a atenção de outros grupos.

Mais uma forma de contribuição diz respeito à infraestrutura computacional compartilhada com as diversas colaborações do CERN para o armazenamento de dados derivados dos experimentos ou para a realização de análises a partir de tais dados. Sergio Novaes (informação verbal)²⁷ argumentou que a Física de Altas Energias trouxe grandes contribuições para o avanço das tecnologias de *Grid Computing*, já que foi necessário criar uma estrutura computacional para dar conta do volume de dados gerados pelos experimentos dessa área. Muitos dos pesquisadores entrevistados também mencionaram o pioneirismo do CERN para o avanço da Internet, com o desenvolvimento da *World Wide Web* (WWW) no CERN, em meados da década de 1990.

²⁵ Publicada no Diário Oficial da União, n. 101, de 29 de maio de 2008.

²⁶ Entrevista concedida (LEITE, 2013). Físico, Físico da USP.

²⁷ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

A estrutura computacional do CERN, ou o *Worldwide LHC Computing Grid* (WLCG), é composta de camadas (Tiers), basicamente 4, sendo (CERN, c2014c):

- a) **Tier 0:** localizado no CERN, é responsável pela coleta dos dados brutos (*raw data*) dos experimentos, as primeiras leituras, realizando uma primeira reconstrução de tais dados para permitir o seu uso e exploração e depois realiza a distribuição desses dados para a segunda camada, Tier 1;
- b) **Tier 1:** agrega atualmente 11 centros de computação (no Canadá, Alemanha, Espanha, França, Itália, Países Nórdicos, Holanda, China, Reino Unido e Estados Unidos) com capacidade suficiente para armazenar os dados gerados pelo LHC. Armazenam proporcionalmente os dados reconstruídos pelo processamento de larga escala.
- c) **Tier 2:** tipicamente universidades ou institutos com poder de computação e capacidade de armazenamento suficientes para a realização de análises específicas;
- d) **Tier 3:** estrutura computacional própria/individual de cientistas ou grupos, sem vinculação formal com o CERN.

Ambos os entrevistados Sergio Novaes e Alberto Santoro (informação verbal)^{28,29} relataram que a participação em cada uma dessas camadas obedece a regras relacionadas à capacidade da estrutura computacional disponível nas instituições participantes e à confiabilidade e estabilidade dessa infraestrutura, ou seja, no caso das Tiers 1 e 2, a estrutura computacional deve estar disponível em regime de 24 horas/dia e nos 7 dias da semana. No Brasil, até o momento, quatro³⁰ instituições mantêm uma estrutura de Tier 2: CBPF, Uerj, Unesp e USP. No caso da Unesp, a administração da Tier 2 fica a cargo do Núcleo de Computação Científica daquela instituição, tendo um *cluster* da estrutura totalmente dedicado aos interesses da colaboração CMS (SPRACE) e o restante da estrutura atendendo a demais projetos da Universidade (GridUnesp).

A Unesp, conforme apontado por Sergio Novaes (informação verbal)³¹, conta com um memorando de entendimento³² assinado entre a Fundação de Amparo à Pesquisa do

²⁸ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

²⁹ Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

³⁰ <http://gstat-wlcg.cern.ch/apps/topology/>.

³¹ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

³² Disponível em: https://espace2013.cern.ch/WLCG-document-repository/MoU/signed/CERN-C-RRB-2005-01_rev_Brazil_PT.pdf.

Estado de São Paulo (Fapesp), agência financiadora e representante no memorando, e o CERN, que formaliza a sua participação no Worldwide LHC Computing Grid (WLCG). Esse documento estabelece as obrigações comuns e as regras de sustentabilidade na colaboração entre os diversos centros ligados ao *grid*. O entrevistado afirmou que durante muitos anos a Unesp foi uma das únicas instituições com essa vinculação formal ao WLCG.

Experiência semelhante à da Unesp, na questão da troca de experiências, pôde ser observada a partir do relato do pesquisador Roberto Marcondes (informação verbal)³³ a respeito do Núcleo de Computação e Matemática da USP. O núcleo realiza parceria com projetos de pesquisa de outras áreas do conhecimento para o desenvolvimento de soluções voltadas para o tratamento e análise de dados de pesquisa. Marcondes afirmou que quando se forma uma colaboração, o núcleo procura evidenciar quais são os fenômenos que o projeto parceiro deseja investigar, quais os tipos de dados envolvidos e como estes são armazenados, que tipo de ruído pode estar presente nos dados etc. Após essa fase é realizado um processo de modelagem, seguido da entrega dos resultados ao parceiro que dará prosseguimento ao processo de análise e interpretação.

É possível observar, em linhas gerais, que as condições extremas demandadas pelo CERN, ou por outras experiências de *e-Science*, aceleram o desempenho e o avanço das equipes e da infraestrutura computacional das instituições parceiras, permitindo a troca de experiências em pesquisas experimentais ou de ponta, além do aproveitamento dos resultados de avanços tecnológicos entre projetos.

4.2.3 Sobre o acesso aos dados gerados pelas colaborações do CERN

Tal como em outras colaborações de natureza distribuída, o acesso aos dados gerados nos experimentos do CERN pode ocorrer de forma local, na própria sede do CERN, ou remota, por meio de qualquer computador com acesso à internet.

Conforme relatou Alberto Santoro (informação verbal)³⁴, antes da internet e da computação distribuída, as detecções, que eram feitas por meio de fotografias, geravam dados que eram analisados *on site*, ou seja, não havia formas de utilizar ou de tratar os dados fora do espaço em que eles eram gerados. Santoro comentou ainda que naquela época havia um certo receio em se distribuir os dados, mas com o crescimento do volume

³³ Entrevista concedida (CESAR JUNIOR, 2013). Cientista da Computação, Professor Titular da USP.

³⁴ Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

e à medida que as colaborações e os experimentos foram ficando mais ‘eletrônicos’ – com os dados gravados imediatamente após as colisões e depois transmitidos para computadores – viu-se a necessidade de distribuição dos mesmos. Atualmente, tanto o acesso aos dados como as análises são feitos por meio de *framework*³⁵ específico de cada colaboração do CERN, muito voltados às necessidades de cada colaboração, de forma que não há soluções comerciais disponíveis, como relatou Sergio Novaes (informação verbal)³⁶.

Nesse sentido, Jesús Mena Chalco afirmou que se passou de uma fase em que os esforços estavam direcionados à geração e obtenção de dados para um momento de concentração dos esforços na análise e exploração de tais dados. Nas palavras do pesquisador:

[...] armazenar os dados processados e anotados é fácil, seria uma coisa rotineira, uma coisa que é possível desde o ponto de vista computacional. Descobrir conhecimento nesses dados, aí é a parte complicada (informação verbal)³⁷.

Ou seja, para cada novo conjunto de dados gerado tem-se o desafio de criação de novas ferramentas e condições de análise, além de um conjunto de normas e regras para regular a divulgação desses dados e estudos derivados, tal como o evidenciado por David, Besten e Schroeder (2006).

De acordo com o relato de pesquisadores relacionados ao CERN (informação verbal)^{38,39,40}, todos os membros da colaboração têm acesso aos dados gerados. Contudo, pôde-se evidenciar a partir do relato dos entrevistados a possibilidade de ocorrência de algumas restrições técnicas ou operacionais quanto ao uso dos dados, não necessariamente como bloqueios, mas como indicativo de que os dados não são públicos ou universalmente acessíveis.

Marco Leite (informação verbal)⁴¹ comentou que o acesso e uso aos/dos dados e informações gerados no CERN é feito por membros qualificados ou em processo de qualificação e estudantes. Os membros qualificados são aqueles cujos nomes figuram na lista de autores de artigos publicados pela colaboração, sendo estes os que têm as

³⁵ Uma plataforma de *software* universal e reutilizável para desenvolvimento de aplicações, produtos e soluções, como programas de computador, por exemplo (SOFTWARE..., 2014).

³⁶ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

³⁷ Entrevista concedida (MENA CHALCO, 2013). Engenheiro de Sistemas, Professor Adjunto da UFABC.

³⁸ Entrevista concedida (PAULA, 2013). Físico, Professor Associado da UFRJ.

³⁹ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁴⁰ Entrevista concedida (LEITE, 2013). Físico, Físico da USP.

⁴¹ Ibidem.

informações e os recursos necessários de como usar os dados para análise. O entrevistado complementou ainda que para o acesso aos dados brutos é necessário que se formalize uma requisição, pois o processo de reconstrução dos dados exige uma capacidade de processamento muito alta, gerando um impacto muito grande no sistema, tal como o tempo e o espaço dos servidores utilizados para cada tomada de dados.

Sergio Novaes (informação verbal)⁴² comentou que se um(a) pesquisador(a) precisa de um determinado *dataset* ele(a) pode importar ou solicitar tais dados, ficando os mesmos disponíveis para quaisquer finalidades, ou seja, não há discriminação entre os membros da colaboração. Contudo, o pesquisador acrescentou:

[...] mas os dados não tornados públicos. O que as pessoas precisam entender é o seguinte: suponha que a gente torne público amanhã os dados, o que vai mudar pra humanidade? Absolutamente nada. Primeiro, vão ser pouquíssimas pessoas que vão conseguir fazer alguma coisa com aquele dado. Porque o dado em si é um monte de bit, pra você manipular aqueles dados você precisa ter as ferramentas corretas, certo? [...] É preciso ter uma estrutura computacional disponível enorme (informação verbal)⁴³.

Ao mencionar as ferramentas corretas, o pesquisador refere-se aos diferentes *frameworks* usados pelas colaborações do CERN para análise e manipulação dos dados. O pesquisador destacou ainda, como um dos problemas decorrentes de uma possível liberação dos dados, a falta de um controle de qualidade na produção dos dados e na atividade científica derivada do uso de tais dados. Segundo ele a validação de qualquer estudo derivado da análise dos dados gerados nas colaborações do CERN depende de uma avaliação prévia dos outros membros da colaboração. Todo o resultado passa por diferentes instâncias da colaboração antes de sua divulgação/publicação como trabalho acadêmico e esse processo envolve validações e explicitação de garantias, por exemplo, de que os dados tenham sido manipulados corretamente, ações que nem mesmo a avaliação por pares tem condições de realizar.

Percebe-se que o tratamento dos dados tem impacto direto nas análises e nos resultados dos experimentos e, com isso, existe uma preocupação para que todas as análises sejam conduzidas a partir do *framework* específico do CERN, criando um espaço equivalente ao ambiente controlado de laboratório. Nesse ambiente controlado, o acesso de pesquisadores é controlado e restrito. Porém, no caso do CERN, isso não ocorre para garantir a proteção de interesses comerciais ou financeiros decorrentes do uso dos dados,

⁴² Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁴³ *Ibidem*.

mas para garantia da confiabilidade e precisão das pesquisas da colaboração, condições que, por sua vez, garantem a continuidade das pesquisas conduzidas pelas colaborações do CERN.

Embora tais ações possam ser vistas como um reforço da qualidade e da responsabilidade na pesquisa conduzida pelas colaborações, por outro lado, elas podem também representar a manifestação do fenômeno de formação de caixas pretas no processo de produção da ciência (WHITLEY, 1972; LATOUR, 2000). Segundo Latour

a expressão *caixa-preta* é usada em cibernética sempre que uma máquina ou conjunto de comandos se revela complexo demais. Em seu lugar, é desenhada uma caixinha preta, a respeito da qual não é preciso saber nada, senão o que nela entra e o que dela sai (2000, p. 14, grifos do autor).

Ou seja, demarca-se um momento em que são descartadas as falhas ou irregularidades da *ciência em construção*, sendo apresentados os resultados de uma *ciência pronta* (LATOUR, 2000) e funcional, cuja avaliação e validação repousam sobre as condições de reprodutibilidade dos dados e dos estudos que, por sua vez, dependem do mesmo ambiente de *software* (OHNO-MACHADO, 2012) implementado no CERN, o qual está além do alcance técnico da comunidade de pareceristas.

Sobre a questão do cerceamento dos dados, Leandro de Paula (informação verbal)⁴⁴ comentou, contudo, que, com a preocupação do final das colaborações, pensa-se na ideia de se deixarem os dados abertos para que qualquer um possa utilizar com qualquer finalidade, não contando como resultado do trabalho das colaborações. O pesquisador destacou que já existe a preocupação sobre como deixar os dados legíveis para outros pesquisadores que não façam parte da colaboração.

Sobre a questão da existência de potenciais riscos envolvendo o compartilhamento ou a cessão de dados, Roberto Marcondes (informação verbal)⁴⁵ afirmou que as comunidades que vêm trabalhando com dados já possuem diretrizes específicas e bem estabelecidas que protegem ou atendem os/aos direitos de qualquer país ou membro da colaboração. Sobre esse aspecto, o pesquisador ainda acrescentou:

Quando um consórcio brasileiro faz parte, quando um grupo brasileiro faz parte de um consórcio que ainda vai criar o sistema, é melhor ainda, porque daí nós fazemos parte, desde o começo, do design do experimento e tudo mais, o que, de certa forma, reconhece os direitos dos grupos brasileiros. Talvez uma limitação brasileira é que nós não temos é... – por várias razões, comunidade pequena, irregularidades do financiamento, falta de ambição de muitos dos pesquisadores

⁴⁴ Entrevista concedida (PAULA, 2013). Físico, Professor Associado da UFRJ.

⁴⁵ Entrevista concedida (CESAR JUNIOR, 2013). Cientista da Computação, Professor Titular da USP.

da nossa comunidade – é difícil imaginar algum projeto desse que tenha sido criado no Brasil ou que já esteja explorando isso daí, liderando isso, seria ótimo se a gente tivesse grupos que procurassem liderar coisas desse tipo.

Observa-se que os elementos de liderança evidenciados pelo pesquisador estão relacionados à iniciativa na proposição de modelos e de soluções e práticas voltadas para o gerenciamento e compartilhamento de dados de pesquisa ou para o desenho e especificação de *frameworks* ou sistemas que viabilizem a exploração de tais dados. Dentre os fatores impeditivos destacados pelo pesquisador, ganha destaque a questão do financiamento irregular, também citado por outros pesquisadores no tocante a outros aspectos das colaborações e que é objeto de análise mais à frente, neste trabalho.

Sobre a formulação de diretrizes para ações de curadoria e necessidade de tratamento de dados constatou-se, pela fala do pesquisador Roberto Marcondes (informação verbal)⁴⁶, que existe a necessidade de uma mudança cultural por parte das comunidades e também um papel proativo dos governos ou agências de financiamento, numa mistura de diretrizes nos sentidos *bottom-up* e *top-down*. Segundo Marcondes,

Tem uma parte que as comunidades vão criando, vão estabelecendo, mas tem uma parte que, por exemplo, [que] o governo e as agências de fomento podem definir. Então, se você pega iniciativas como a NSF [National Science Foundation], por exemplo, tem várias dessas iniciativas, que começam a exigir que um projeto de pesquisa ao ser apresentado ele deve prever o que vai acontecer com os dados, quais dados vão ser gerados e como esses dados vão ser preservados, aí [com] isso você começa a embutir uma cultura junto às diferentes comunidades, porque daí para o pesquisador ter acesso àquele auxílio ele tem que prever o que ele vai fazer com os dados, assim como ele prevê que ele tem que atualizar o currículo Lattes dele porque se não ninguém, simplesmente, vai financiar um novo projeto dele (informação verbal)⁴⁷.

Marco Leite e Jesús Mena Chalco (informação verbal)^{48,49} por sua vez comentaram sobre a possibilidade de uso da computação na nuvem (*cloud computing*) para a hospedagem de dados e para a realização de tarefas de análise que dependam de alto poder computacional. Mena Chalco acrescentou que a computação na nuvem traz um componente de democratização da ciência, pois permite que países que não disponham de tecnologias ou infraestrutura de ponta também possam realizar pesquisas com grandes quantidades de dados. Sobre tal assunto, há que se destacar, contudo, a necessidade de investimento financeiro para a locação de tais serviços ou infraestruturas

⁴⁶ Entrevista concedida (CESAR JUNIOR, 2013). Cientista da Computação, Professor Titular da USP.

⁴⁷ Entrevista concedida (CESAR JUNIOR, 2013). Cientista da Computação, Professor Titular da USP.

⁴⁸ Entrevista concedida (LEITE, 2013). Físico, Físico da USP.

⁴⁹ Entrevista concedida (MENA CHALCO, 2013). Engenheiro de Sistemas, Professor Adjunto da UFABC.

computacionais que, embora seja granular, em muitos casos deve ser garantido em um horizonte de tempo prolongado.

4.2.4 Resultados e condições de avaliação das colaborações

Neste estudo, procurou-se evidenciar que tipos de resultados (publicações, dados etc.) são tomados como elementos para avaliar o desempenho das colaborações, sendo que essa temática foi abordada sob, ao menos, três dimensões: o aspecto das publicações derivadas dos trabalhos colaborativos no CERN e em outras colaborações; implicações relacionadas aos dados derivados da pesquisa colaborativa; e questões relacionadas à possibilidade de registro de propriedade intelectual a partir de processos e/ou produtos ou licenciamento de *software* derivados das colaborações.

Quanto às publicações, os entrevistados atuantes no CERN foram questionados sobre os procedimentos para publicação e divulgação de resultados das pesquisas realizadas e sobre a avaliação e atribuição de créditos pela colaboração.

Constatou-se, com base nos relatos, que todos os membros efetivos das colaborações, ou aqueles que, entre outras regras e requisitos, contribuem com a taxa de Manutenção e Operação, têm os seus nomes divulgados na lista de autores dos artigos publicados por cada colaboração, tal como ilustrado na Figura 3.

É importante ressaltar que a definição dessa lista implica a dotação de orçamento nas instituições de origem dos pesquisadores para a contribuição com a taxa de Manutenção e Operação junto ao CERN. Dessa forma, a possibilidade de fixação de um teto orçamentário poderia interferir na quantidade de pesquisadores que constarão da lista de autores e, conseqüentemente, na seleção ou definição destes. No caso do Brasil, uma proposta para a lista de autores brasileiros para 2014, para cada uma das colaborações, foi disponibilizada em site provisório da RENAFAE⁵⁰, permitindo a produção da Tabela com o quantitativo de autores apresentada anteriormente (p. 47).

A partir dessa diretriz adotada nas publicações do CERN, os pesquisadores entrevistados relataram que toda publicação divulgada em nome da colaboração (com a respectiva lista de autores) deve passar antes pelo conhecimento de todos os membros e

⁵⁰ <http://www.cbpf.br/~renafae/autores/>.

pela revisão de comitês ou comissões, como *Editorial Board*, *Collaboration Board* etc., com o número de instâncias e denominações podendo variar de uma colaboração para outra.

Figura 3. Exemplo de menção a lista de autores em artigo publicado por colaboração do CERN

The screenshot shows a page from the journal *Science*, dated 21 December 2012, volume 338, pages 1569-1575. The article title is "A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider". The author information is listed as "The CMS Collaboration" with a footnote indicating that correspondence should be addressed to cms-spokesperson@cern.ch. A second footnote states that the complete list of authors and affiliations is in the supplementary material. The abstract follows, describing the discovery of a new particle with a mass of about 125 GeV.

21 December 2012 > 338 (6114): 1569-1575

Science 21 December 2012:
Vol. 338 no. 6114 pp. 1569-1575
DOI: 10.1126/science.1230816

< Prev | Table of Contents | Next >

ARTICLES

A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider

The CMS Collaboration^{†,‡}

[†]To whom correspondence should be addressed. E-mail: cms-spokesperson{at}cern.ch.

[‡]The complete list of authors and affiliations appears in the [supplementary material](#) for this paper.

ABSTRACT

The Higgs boson was postulated nearly five decades ago within the framework of the standard model of particle physics and has been the subject of numerous searches at accelerators around the world. Its discovery would verify the existence of a complex scalar field thought to give mass to three of the carriers of the electroweak force—the W^+ , W^- , and Z^0 bosons—as well as to the fundamental quarks and leptons. The CMS Collaboration has observed, with a statistical significance of five standard deviations, a new particle produced in proton-proton collisions at the Large Hadron Collider at CERN. The evidence is strongest in the diphoton and four-lepton (electrons and/or muons) final states, which provide the best mass resolution in the CMS detector. The probability of the observed signal being due to a random fluctuation of the background is about 1 in 3×10^6 . The new particle is a boson with spin not equal to 1 and has a mass of about 125 giga-electron volts. Although its measured properties are, within the uncertainties of the present data, consistent with those expected of the Higgs boson, more data are needed to elucidate the precise nature of the new particle.

Fonte: *Science*, v. 338, n. 6114, 21 Dec. 2012. Disponível em:
<<https://www.sciencemag.org/content/338/6114/1569.short>>. Acesso em: 16 fev. 2014.

As regras, orientações e diretrizes, tanto para a atuação na colaboração em geral, quanto para os procedimentos referentes às publicações, são explicitadas e registradas na forma de *wikis*⁵¹, no caso do CERN, denominadas TWiki⁵². Em cada TWiki é possível visualizar um conjunto de informações de acesso público, enquanto alguns diretórios são de acesso restrito aos membros do CERN, mediante autenticação com usuário e senha.

Sobre o uso das publicações como elemento de mensuração da 'produtividade' das colaborações ou dos pesquisadores que nelas atuam, Sergio Novaes (informação

⁵¹ "Wiki é, usualmente, uma aplicação *web* a qual permite que pessoas possam adicionar, modificar ou deletar conteúdos [na web] em colaboração com outras" (WIKI, 2014).

⁵² <https://twiki.cern.ch/twiki/bin/view/DefaultWeb/WebHome>.

verbal)⁵³ comentou que o artigo não é um bom indicativo da contribuição individual, mas continua como uma boa medida da contribuição efetiva para a produção de conhecimento. Ele também afirmou que já existem mecanismos no âmbito das colaborações que permitem medir a colaboração individual.

Sobre este tema da avaliação por meio de artigos, Jesús Mena Chalco (informação verbal)⁵⁴ comentou que a qualidade dos artigos de determinadas áreas do conhecimento já se encontra em um estado de saturação. O pesquisador tomou como exemplo a área da Física, em que todos os artigos são publicados em revistas internacionais, com grande volume de coautorias e recebendo um volume considerável de citações, ou seja, todas cumprem os requisitos para atingirem avaliação máxima pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), sem que haja possibilidade de se criar novos níveis de distinção entre as mesmas.

O pesquisador Jesús Mena Chalco (informação verbal)⁵⁵ pontuou, contudo, que o objetivo principal do trabalho colaborativo não são as publicações – estas são produtos, mas sim a criação de metodologias, formas de operacionalização do trabalho colaborativo que podem ser aproveitados em projetos da mesma natureza. Sobre esse aspecto, o pesquisador salienta que o aproveitamento das metodologias ainda é muito restrito às questões e aos problemas específicos, porém não exclusivos, de cada área do conhecimento, por exemplo, as metodologias geradas para o mapeamento do genoma humano não se aplicam ao tratamento dos dados da Física ou vice-versa. Ocorre que são criadas ferramentas muito específicas para cada problema e não há quem organize a colaboração interdisciplinar/intergrupos. Ele citou o exemplo de bancos especializados como o NCBI⁵⁶, ArXiv⁵⁷, bancos da área de Astronomia etc. que, embora funcionem de forma parecida, não são interoperáveis, não ‘conversam’.

Sobre a questão da propriedade intelectual ou de licenciamento, não foi possível observar um consenso ou uma diretriz comum a partir dos relatos dos pesquisadores. Houve menção à ausência de procedimentos para registro de propriedade intelectual, à existência de um setor específico no CERN voltado para a transferência de tecnologias e à necessidade de compartilhamento da propriedade entre a instituição de origem e o CERN,

⁵³ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁵⁴ Entrevista concedida (MENA CHALCO, 2013). Engenheiro de Sistemas, Professor Adjunto da UFABC.

⁵⁵ Ibidem.

⁵⁶ <http://www.ncbi.nlm.nih.gov/>.

⁵⁷ <http://arxiv.org/>.

além de relato sobre a experiência de desenvolvimento de instrumentação científica com o uso de modelo de licenciamento *Open Hardware*, desenvolvido pelo CERN.

O CERN conta com um escritório de transferência de tecnologias, cujo *site*⁵⁸ apresenta uma diretriz geral de que a instituição disponibiliza, por meio de diferentes oportunidades de transferência de tecnologia, uma expertise técnica e tecnologias inovadoras para propósitos científicos e comerciais.

Constatou-se, em linhas gerais, que o CERN atua no limiar da capacidade tecnológica, envolvendo aceleradores, detectores, instrumentação científica, tecnologias da informação etc., havendo a necessidade de encomenda ou mesmo o desenvolvimento *in loco* de todo o aparato tecnológico necessário à execução dos experimentos. Com isso, o CERN abre espaço para empresas ou demais instituições de pesquisa, seja por meio de licenças abertas ou outros tipos de acordos que não envolvam *royalties*, para que as mesmas desenvolvam tais tecnologias. No caso de empresas, além do interesse na exploração comercial, as mesmas têm a oportunidade de testarem suas tecnologias e/ou processos produtivos sob as condições extremas existentes no CERN.

No caso específico das licenças abertas, o CERN desenvolveu o modelo de licenciamento *Open Hardware*⁵⁹ que, entre outros aspectos, define que qualquer pessoa ou interessado deve ser capaz de acessar a documentação (contendo especificações, *design* etc.) de um projeto para a construção de *hardware* (máquinas, equipamentos, placas de computador etc.), além de poder estudar, modificar e compartilhar essa documentação de acordo com seu interesse.

Sobre o uso de tais licenças, Daniel Tavares (informação verbal)⁶⁰, que atua no Laboratório Nacional de Luz Síncrotron (LNLS) no desenvolvimento de instrumentação científica, argumentou que o desenvolvimento de placas-mães ou circuitos integrados com o uso desse tipo de licenciamento traz algumas vantagens para o processo, como a possibilidade de customização de especificações já existentes, além da possibilidade de transferência de tecnologia para a implementação/desenvolvimento de tais especificações por instituições parceiras. Entre outras vantagens, Tavares (informação verbal)⁶¹ também relatou a possibilidade de uso do *framework* do CERN para desenvolvimento de *hardware* aberto, o uso de repositório do CERN para

⁵⁸ <http://knowledgetransfer.web.cern.ch/technology-transfer>.

⁵⁹ <http://www.ohwr.org/projects/cernohl/wiki>.

⁶⁰ Entrevista concedida (TAVARES, 2013). Engenheiro de Automação, Engenheiro Eletricista do LNLS.

⁶¹ *Ibidem*.

compartilhamento de código fonte e especificações⁶² (qualquer empresa ou grupo pode participar ou compartilhar), além da troca de experiência com as equipes de desenvolvimento do CERN.

4.2.5 Cenário Brasil

Dentre os pesquisadores entrevistados, muitos deles atuaram em colaborações do CERN desde a sua gênese, e afirmam que a participação brasileira na colaboração é essencial para o avanço no campo da física de altas energias no país, já que no Brasil, ou em qualquer outro país, a construção de novos aceleradores é inviável e o tempo de formação de pesquisadores capacitados para atuar nos experimentos é muito longo. Ronald Shellard (informação verbal)⁶³ destacou que existem grandes desafios científicos para o desenvolvimento de estudos da física de altas energias, estudos climáticos etc., os quais demandam grandes investimentos, tal como uma instituição, um local para organizar experimentos. O custo com a implementação de sensores também é muito alto, sendo necessário desenvolver tecnologias sofisticadas para diminuir os custos desses sensores e, segundo o pesquisador, nenhum país sozinho tem expertise para desenvolver experimentos dessa magnitude – por isso depende de outros participando na colaboração.

Os pesquisadores sinalizaram a questão da aprendizagem na interação com outros pesquisadores no CERN como benefício significativo da participação nas colaborações. Sergio Novaes (informação verbal)⁶⁴ comentou que há um aprendizado, derivado das colaborações, que vai muito além dos conhecimentos a respeito da física de altas energias. Segundo ele ocorre a capacitação em atividades de gerenciamento, de definição de projetos e metas, em computação, eletrônica e tratamento de dados, sendo que para o físico experimental existe a necessidade de desenvolvimento de diferentes habilidades e expertise. Marco Leite (informação verbal)⁶⁵ apontou o próprio ambiente colaborativo como vantagem, já que o pesquisador depende do suporte e do contato com outros pesquisadores, o que é muito bom para os estudantes, por exemplo, pois faz com que eles desenvolvam uma independência, uma pro-atividade.

⁶² <http://www.ohwr.org/>.

⁶³ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

⁶⁴ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁶⁵ Entrevista concedida (LEITE, 2013). Físico, Físico da USP.

No plano das barreiras ou dificuldades apontadas pelos pesquisadores para a participação brasileira nas colaborações com o CERN, ou mesmo em iniciativas de *e-Science*, uma das principais diz respeito à dificuldade de planejamento e provisão de recursos para iniciativas de pesquisa de longo prazo. De acordo com Marco Leite (informação verbal)⁶⁶, existe no Brasil a carência de um projeto de desenvolvimento científico de longo prazo, já que todos as chamadas e editais são muito sazonais.

Roberto Marcondes argumentou:

[...] o que acontece muito com o CERN, é que é uma comunidade muito específica, tem uma outra característica que também é verdade nesse caso que é o seguinte: o CERN em si, por ser um grande experimento mundial, eles precisam se organizar de uma certa maneira pra poder financiar esse experimento e é, normalmente, um projeto que envolve um prazo muito longo, então eles têm que planejar o negócio pra o ano que vem, daqui a cinco anos, daqui a dez anos, daqui a quinze anos... e assim por diante, e que custa muito dinheiro, por isso você tem que ter uma colaboração internacional [...] (informação verbal)⁶⁷.

Sobre esse aspecto, Alberto Santoro (informação verbal)⁶⁸ comentou que no caso da física de altas energias há um tempo, um período, que pode ser considerado oportuno para a ação – o que ele identificou como *timing*. Segundo ele, os recursos de financiamento devem estar à mão/disponíveis de acordo com o *timing* da área, contudo, a administração científica brasileira não tem essa noção, entre outros fatores, pelo fato de a comunidade brasileira de física de altas energias ser muito pequena comparada a outras áreas, o que lhe dá menor poder de barganha frente aos órgãos de financiamento.

Santoro comentou ainda que a escassez e a grande concentração dos laboratórios de física de altas energias demandam que os pesquisadores viajem constantemente e, embora haja recurso suficiente para custeio dos deslocamentos, há sempre um esforço imensurável na apresentação de documentação para a liberação desses recursos.

As afirmações dos pesquisadores sobre as questões de financiamento refletem a manifestação de visíveis mudanças nas práticas de pesquisa, alterando a noção de pesquisa baseada em projetos, com início meio e fim, destacadas por Barjak et al. (2013), e demandando a composição de fontes de financiamento sustentáveis para a manutenção permanente das infraestruturas de *e-Science*, seja para permitir a geração e o tratamento contínuo de dados de pesquisa, seja para possibilitar as interações e o comprometimento ininterruptos entre os membros das colaborações.

⁶⁶ Entrevista concedida (LEITE, 2013). Físico, Físico da USP.

⁶⁷ Entrevista concedida (CESAR JUNIOR, 2013). Cientista da Computação, Professor Titular da USP.

⁶⁸ Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

Sobre os impactos para o Brasil, considerando-se a questão da infraestrutura tecnológica, Sergio Novaes destacou a importância dos benefícios em termos de incremento em infraestrutura computacional para a sua instituição, advindos da experiência na colaboração. Segundo o pesquisador,

A universidade hoje em dia tem uma estrutura computacional respeitável, para qualquer área fazer ciência devido ao fato dela ter tido um grupo participando em física de altas energias que se dispôs a implementar essa estrutura para toda a universidade (informação verbal)⁶⁹.

Outra afirmação do pesquisador diz respeito ao proveito dos benefícios da colaboração em outras áreas ou projetos da instituição:

Tinha um professor, quando eu implementei o GridUNESP, ele veio falar pra mim que ele que tinha uma molécula, uma proteína na gaveta dele há alguns anos, que ele não tirava da gaveta porque era uma coisa extremamente interessante... ele não tirava essa proteína da gaveta pra fazer estudo de enovelamento de proteína porque ele não tinha capacidade computacional pra fazer isso e agora ele iria tirar a proteína da gaveta. Então isto mostra um pouco qual era o interesse, é gente se lançar em áreas... tem áreas [em que] o poder computacional é impeditivo, certo, você não pode fazer nada naquela área se você não tiver uma excelente capacidade computacional (informação verbal)⁷⁰.

Por outro lado, essa condição expressa pelo pesquisador também denota a dependência tecnológica vivenciada por algumas áreas do conhecimento e as consequências disso para o desenvolvimento científico. O ingresso nas colaborações, por um lado, pode representar um incentivo para a gênese de iniciativas em investimento em infraestrutura; por outro, se tais investimentos não forem planejados ou sustentáveis, corre-se o risco de contar com um parque tecnológico estagnado.

Exemplo disso, a Rede Nacional de Pesquisa (RNP), de acordo com Ronald Shellard (informação verbal)⁷¹, surgiu da necessidade de comunicação entre os físicos de altas energias no Brasil, fomentando a ampliação das redes lógicas. Contudo, essa infraestrutura já não mais atende satisfatoriamente a necessidade das colaborações, conforme relatou Alberto Santoro (informação verbal)⁷².

Dentre os benefícios decorrentes do ingresso do Brasil como membro associado ao CERN apontados pelos entrevistados (informação verbal)^{73,74} está o fato de as

⁶⁹ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁷⁰ Entrevista concedida (NOVAES, 2013). Físico, Professor Titular da Unesp.

⁷¹ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

⁷² Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

⁷³ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

⁷⁴ Entrevista concedida (PAULA, 2013). Físico, Professor Associado da UFRJ.

empresas brasileiras da área de tecnologia poderem participar de concorrências em licitações e editais de qualquer natureza do CERN (incluindo licitações para aquisições de material de consumo), juntamente com empresas de demais países associados. Outro fator inclui a possibilidade da contratação de profissionais brasileiros por meio de bolsas de formação de recursos humanos, com contratos de dois anos, para o exercício de funções técnico/administrativas como tradutores, jornalistas etc.

Por outro lado, Alberto Santoro (informação verbal)⁷⁵ ponderou que não há garantia de retorno, ou materialização, dos recursos financeiros investidos com a associação. Santoro questionou se o investimento deveria mesmo beneficiar a indústria brasileira, com a oportunidade de lucrar com as licitações, ou se não traria melhores resultados [como investimento em ciência] caso fosse aplicado diretamente nas instituições brasileiras que já colaboram com o CERN – nas formas anteriormente evidenciadas – intelectualmente, via infraestrutura, via pagamento da taxa de Manutenção e Operação etc.

Sobre essa questão, é importante destacar os papéis e os interesses dos diferentes atores em meio à questão da governança na colaboração e sobre os investimentos em ciência e tecnologia. Em relação ao papel do cientista, vale observar o relato do pesquisador Ronald Shellard:

[...] eu sempre digo, eu sempre enfatizo o seguinte: o que a gente faz não é útil pra sociedade. Um pouco pra chocar a pessoa né? É óbvio que é útil, mas não é minha função torná-lo útil (informação verbal)⁷⁶.

É possível perceber, com base nessa afirmação, um dos elementos que Stoker (1998) relaciona à questão da governança, a saber, a identificação de barreiras e responsabilidades difusas para lidar com as questões sociais. Observa-se que o pesquisador atua no sentido de demarcar a sua função social na prática científica, abrindo espaço para a manifestação de outros atores nesse contexto. Essa observação torna-se mais visível a partir de outro comentário do pesquisador:

No Brasil, por exemplo, você olha agora o ministério de C&T está muito interessado em inovação e o que é inovação? Pra que serve a inovação? Não é? Pensa dois segundos. Claro que inovação, qualquer indústria, [por exemplo], se eu sou um industrial eu tô interessado ali, [mas], tem uma gradação brutal do que seria inovação, eu, por exemplo, se eu sou um pequeno industrial eu quero melhorar um pouquinho a eficiência, como é que eu uso um operário a menos... tem esses pequenos ganhos que é a grande parte da inovação no dia-a-dia das

⁷⁵ Entrevista concedida (SANTORO, 2014). Físico, Professor Titular da Uerj.

⁷⁶ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

empesas. Aí depois você tem inovações que são mudanças um pouco de paradigma nos mecanismos de produção, que é mais de engenharia sofisticada. Tem aí, vamos dizer, as grandes mudanças, a gente trabalha muito mais no investimento que vai esse tipo de ciência, são as grandes mudanças de paradigma. Essas [inovações] não tem como você prever [...] Esses equipamentos que a gente usa pra detectar partículas, que foram inventados lá na década de 1960, hoje em dia estão na..., você vai fazer uma tomografia e ali há um detector de partículas, tem físicos de altas energias envolvidos? Não, hoje em dia não tem, não é? Mas [essas aplicações] são decorrentes dessa, vamos dizer, desse investimento (informação verbal)⁷⁷.

A partir desse ponto já entra em cena uma série de novos atores e interesses, governamentais, da indústria, de usuários de serviços de saúde entre outros, caracterizando o problema da legitimação *versus* extensão no processo de decisões sobre ciência e tecnologia destacado por Collins e Evans (2002). Ou seja, enquanto o governo e a indústria lançam mão da opinião de especialistas (analistas de C&T e engenheiros, por exemplo), os integrantes da sociedade civil realizam avaliação dos investimentos com base em sua percepção sobre as condições dos serviços públicos básicos, como saúde, educação, segurança etc.

De forma geral, com base em todos os aspectos das colaborações em *e-Science* abordados neste trabalho, foi possível a construção de um *framework* básico sobre as principais mudanças relacionadas às práticas de pesquisa dos cientistas atuantes em colaborações desta natureza, disponível no quadro a seguir.

Quadro. Síntese das principais mudanças/impactos nas práticas

(continua)

Dimensão	Cientistas/Pesquisadores		Práticas/Processos	
	Antes	Depois	Antes	Depois
Participação	Acesso por meio de prestígio, status	Acesso por meio da contribuição intelectual	Dispendiosa, consumindo tempo e recursos	Custos compartilhados; Ocorre em tempo real, mediada por tecnologia
Dados e resultados	Publicações individuais ou com poucas coautorias	Publicação coletiva e em forma de <i>wiki</i>	Posse individual, intraprojeto; Compartilhamento na forma de artigo (como produto final)	Compartilhamento e abertura durante todo o processo de geração e tratamento Dados como ativo importante no processo

⁷⁷ Entrevista concedida (SHELLARD, 2013). Físico, Pesquisador Titular do CBPF.

(conclusão)

Dimensão	Cientistas/Pesquisadores		Práticas/Processos	
	Antes	Depois	Antes	Depois
Organização	Intercâmbio de informações e recursos direcionados, 1:1 e 1:n; colégios invisíveis	Intercâmbio de informações e de recursos por meio de <i>grids</i> e em recursos abertos, 1:n e n:n	Localizadas	Globalizadas, distribuídas e descentralizadas
Financiamento	São a principal fonte de consulta sobre as decisões que envolvem investimento em C&T&I	Compartilham o papel decisório com demais atores ou setores da sociedade	Financiamento por projeto, com início, meio e fim	Financiamento sustentável a longo prazo

Fonte: o autor.

Ressalta-se que essa síntese não teve compromisso com a exaustividade, ou seja, buscou-se apenas construir um pano de fundo para facilitar a compreensão das mudanças. No capítulo das considerações finais, a seguir, está disponível um apanhado mais descritivo dessas mudanças, assim como um panorama dos resultados e objetivos alcançados com a conclusão deste trabalho.

5 CONSIDERAÇÕES FINAIS

Pode-se afirmar que um dos resultados mais significativos alcançados no decorrer deste trabalho foi a elucidação das dimensões presentes no contexto das colaborações em *e-Science*, a saber: a construção de uma infraestrutura computacional para uso distribuído ou para processamento de larga escala; a produção e o uso intensivo de dados; e a colaboração entre atores da ciência, pelo compartilhamento de esforços, dados e/ou recursos computacionais. A compreensão de tais dimensões trouxe vantagens à proposta de pesquisa, pois orientou a exploração e a organização do quadro teórico de referência, além da construção de categorias de análise que, por sua vez, amparou a condução da pesquisa empírica.

Percebeu-se, ao longo da pesquisa, que as Tecnologias da Informação e da Comunicação expandiram os limites das colaborações, contudo elas também ampliaram o leque de possíveis barreiras ou restrições para o acesso aos dados ou demais recursos mobilizados nas diferentes colaborações. As TICs também impactaram a ampliação dos níveis de avaliação das colaborações. Enquanto na fase anterior à emergência da *e-Science* as colaborações ou grupos de pesquisa eram avaliados unicamente com base nas publicações, com o advento da *e-Science* ocorre a proposição de novos indicadores, sendo que muitos destes ainda estão sendo estudados/aperfeiçoados e alguns grupos têm mais ou menos sucesso na sua implementação.

Com relação à infraestrutura necessária à viabilização da *e-Science*, observou-se que as colaborações se reúnem em torno de problemas complexos, e essa complexidade é diretamente relacionada à quantidade de dados que esses problemas científicos geram ou demandam. Existe uma produção muito grande de dados, associada à natureza científica do problema abordado, e esse processo traz novos desafios às colaborações, como a questão da armazenagem – possibilidade de estocagem e uso eficaz desses dados em momentos posteriores –, a questão do processamento e análise, sendo que estes devem ocorrer de forma distribuída ou a partir de diferentes localidades.

Sobre a interoperabilidade dos dados e dos processos, verificou-se que esta ocorre com maior frequência no plano interno das colaborações, o que às favorece no sentido de colocá-las em uma posição privilegiada na busca por financiamento, ou seja, não faz sentido ou torna-se inviável a mobilização ou a replicação de recursos e esforços fora das colaborações.

Verificou-se também que os dados, assim como os artigos e demais tipos de publicações científicas, tornam-se ativos importantes da pesquisa, reconfigurando a busca por resultados e os esforços das colaborações em *e-Science*.

Sobre a necessidade de documentação ou de diretrizes que regulem o acesso e as formas de uso dos dados decorrentes das colaborações, evidenciou-se que essas devem ser extensivas e conhecidas por parte de todos os integrantes das colaborações. Por outro lado, comunidades ou pesquisadores que não estejam formalmente vinculados às colaborações podem encontrar dificuldades no acesso aos dados ou à documentação, seja por limites técnicos das colaborações, seja pela dificuldade de compreensão e uso desses recursos. Diante disso, evidenciou-se ainda que a percepção de restrições não se resume a questões comerciais, de registro de propriedade intelectual ou cobrança de *royalties*, mas diz respeito também às dificuldades de acesso inerentes aos processos de uso e análise dos dados gerados pelas colaborações.

Outro desafio que se mostrou presente durante o estudo das mudanças nas colaborações foi a questão da escala ou dimensionamento. Percebeu-se uma mudança de escala, tanto na quantidade de pessoas envolvidas nas colaborações, na distância a ser percorrida/transposta para o intercâmbio de recursos, assim como na escala econômica ou de financiamento. Evidenciou-se que cada projeto ou iniciativa voltada para a colaboração em *e-Science* deve ser planejada com uma provisão de recursos para uma perspectiva temporal muito mais longa. As fases de início, meio e fim de cada projeto tornam-se difusas e com isso as fontes de recursos devem ser contínuas, já que não mais é possível indicar o período de duração de um projeto ou quando os primeiros resultados começarão a surgir.

Com o advento da *e-Science* e das colaborações em larga escala a governança passa por uma mudança significativa nas práticas científicas. Sua importância se faz presente no processo de organização dos atores da colaboração e de como definem seus papéis, na forma como esses atores compartilham os seus recursos e os resultados alcançados pelas colaborações. Nesse sentido, evidencia-se a necessidade de aprofundar o estudo de novas formas de avaliação do reconhecimento pela atuação nas práticas colaborativas, ou seja, poder identificar e analisar os diferentes níveis ou as diferentes formas de atuação e contribuição, de forma que possam ser considerados na partilha de resultados ou na correção de distorções.

Tais preocupações não são significativas em projetos colaborativos locais ou com menor número de integrantes, nos quais a centralidade das decisões pode até ser um fator favorável. Porém com a transição para a colaboração em redes a centralidade das decisões torna-se inviável e o alcance destas cada vez mais limitado. Nesse sentido, a construção da governança pode auxiliar na distribuição dos objetivos a serem alcançados, na definição dos propósitos das colaborações, assim como na identificação de como se dão as relações de poder entre os diferentes atores presentes no interior das colaborações e como alinhá-las para permitir que a colaboração se auto governe e avalie.

A existência de uma organização descentralizada no âmbito das colaborações pode vir a amenizar as tradicionais hierarquias ou relações de poder baseadas em características socioeconômicas, políticas ou de prestígio institucional, por outro lado, ela também abre espaço para a manifestação de novas relações voltadas para mobilização de atores humanos e não humanos (pessoas, recursos financeiros, instrumentos, ideias etc.) que flexionam a rede e as decisões. Tais considerações apontam para a necessidade de desenvolvimento de trabalhos futuros sobre essa temática, buscando-se evidenciar, entre outros aspectos, os elementos que configuram as relações de poder nesse novo contexto organizacional da pesquisa.

Outro aspecto das colaborações que está ligado a essa nova dimensão da governança diz respeito ao comportamento destas frente aos interesses mais amplos da sociedade, além os interesses da comunidade científica ou da comunidade de membros que nelas atuam. Isso reflete a ideia de que os objetivos das colaborações devem ou podem estar alinhados aos interesses sociais.

Dentre as mudanças que afetam as práticas dos cientistas brasileiros, decorrentes da sua inserção nas colaborações em *e-Science*, pode-se destacar a ampliação do leque de capacitação para a pesquisa, com o desenvolvimento de habilidades e competências que só podem ser alcançadas pela vivência e prática nas colaborações.

Verificou-se também que há uma demanda dos pesquisadores para a ampliação da partilha de responsabilidades quanto às tarefas e decisões burocráticas das iniciativas de colaboração, como a provisão de recursos e a necessidade de implementação de escritórios de projeto, por exemplo – o que também pode demandar novas práticas e ações no campo da governança.

Em linhas gerais, foi possível observar a convergência dos objetivos propostos e temas abordados, apresentando-se como principal desafio a carência de estudos e

publicações brasileiras sobre o assunto. Como perspectivas futuras de estudo, destaca-se a potencialidade de desenvolvimento de um *framework* para análise e estudo da governança em *e-Science*, com foco na abertura dos processos de geração e tratamento de dados derivados de pesquisas colaborativas, ou de iniciativas por parte da comunidade científica brasileira alinhadas a esses processos, garantindo-lhe um posicionamento estratégico no campo das colaborações em *e-Science*.

REFERÊNCIAS

ALBAGLI, S.; MACIEL, M. L. Informação, poder e política: a partir do sul, para além do sul. In: MACIEL, M. L.; ALBAGLI, S. *Informação, conhecimento e poder: mudança tecnológica e inovação social*. Rio de Janeiro: Garamond, 2011. p. 9-39.

BARJAK, F.; ECCLES, K.; MEYER, E. T.; SCHROEDER, R.; ROBINSON, S. The emerging governance of e-Infrastructure. *Journal of Computer-Mediated Communication*, v. 18, n. 2, p. 1-24, Jan. 2013. Disponível em: <<http://dx.doi.org/10.1111/jcc4.12000>>. Acesso em: 26 fev. 2013.

BELL, G.; HEY, T.; SZALAY, A. Beyond the data deluge. *Science*, v. 323, n. 5919, p. 1297-1298, 6 Mar. 2009. Disponível em: <<http://dx.doi.org/10.1126/science.1170411>>. Acesso em: 25 mar. 2013.

BEN-DAVID, J.; SULLIVAN, T. A. Sociology of Science. *Annual Review of Sociology*, v. 1, p. 203-222, 1975. Disponível em: <<http://www.jstor.org/stable/2946045>>. Acesso em: 14 jun. 2012.

BJÖRNEBORN, L.; INGWERSEN, P. Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, v. 55, n. 14, p. 1216-1227, Aug. 2004. Disponível em: <<http://dx.doi.org/10.1002/asi.20077>>. Acesso em: 4 jul. 2012.

BLOOR, D. Wittgenstein and Mannheim on the sociology of mathematics. *Studies in History and Philosophy of Science*, v. 4, n. 2, p. 173-91, Aug. 1973. Disponível em: <[http://dx.doi.org/10.1016/0039-3681\(73\)90003-4](http://dx.doi.org/10.1016/0039-3681(73)90003-4)>. Acesso em: 15 jul. 2013.

CALLON, M. Techno-economic networks and irreversibility. In: LAW, J. (Ed.). *A sociology of monsters: essays on power, technology and domination*. Londres: Routledge, 1991. p. 132-164.

Conseil Européen pour la Recherche Nucléaire (CERN). *About CERN*: What is the universe made of? How did it start? Physicists at CERN are seeking answers, using some of the world's most powerful particle accelerators. c2014a. Disponível em: <<http://home.web.cern.ch/about>>. Acesso em: 2 fev. 2014.

Conseil Européen pour la Recherche Nucléaire (CERN). Experiments. *LHCb*. Feb. 2012. Disponível em: <<http://home.web.cern.ch/about/experiments/lhcb>>. Acesso em: 15 jul. 2013.

Conseil Européen pour la Recherche Nucléaire (CERN). *How a detector works*: Just as hunters can identify animals from tracks in mud or snow, physicists identify subatomic particles from the traces they leave in detectors. c2014b. Disponível em: <<http://home.web.cern.ch/about/how-detector-works>>. Acesso em: 2 fev. 2014.

Conseil Européen pour la Recherche Nucléaire (CERN). *The Grid*: a system of tiers. c2014c. Disponível em: <<http://home.web.cern.ch/about/computing/grid-system-tiers>>. Acesso em: 9 fev. 2014.

CESAR JUNIOR, R. M. *Roberto Marcondes Cesar Junior: entrevista* [2 dez. 2013]. Entrevistador: Andre Luiz Appel. São Paulo: [S.n.], 2013. 1 arquivo .mp3 (25 min.).

CHADE, J. Brasil fará parte do maior laboratório de Física do mundo. *O Estado de S. Paulo*, 13 dez. 2013. Disponível em: <<http://www.estadao.com.br/noticias/impreso,brasil-fara-parte-do-maior-laboratorio-de-fisica-do-mundo,1107939,0.htm>>. Acesso em: jan. 2014.

COLLINS, H. M. The Sociology of Scientific Knowledge: studies of contemporary Science. *Annual Review of Sociology*, v. 9, p. 265-285, Aug. 1983. Disponível em: <<http://www.jstor.org/stable/2946066>>. Acesso em: 3 maio 2013.

COLLINS, H. M.; EVANS, R. The third wave of science studies: studies of expertise and experience. *Social Studies of Science*, v. 32, n. 2, p. 235-296, Apr. 2002. Disponível em: <<http://dx.doi.org/10.1177/0306312702032002003>>. Acesso em: 24 maio 2013.

DAVID, P. A.; BESTEN, M. den; SCHROEDER, R. How open is e-Science? In: IEEE INTERNATIONAL CONFERENCE ON E-SCIENCE AND GRID COMPUTING, 2., 4 Dec. 2006, Amsterdam. *Anais...* Amsterdam: IEEE, 2006. Disponível em: <<http://dx.doi.org/10.1109/E-SCIENCE.2006.85>>. Acesso em: 5 jun. 2012.

DUMBILL, E. Getting up to speed with big data. In: O'REILLY MEDIA. *Big data now: 2012 edition*. Sebastopol, CA, USA: O'Reilly Media, 2012. Disponível em: <<http://oreilly.com/data/radarreports/big-data-now-2012.csp>>. Acesso em: 27 out. 2012.

DUTTON, W. Collaborative network organizations: new technical, managerial and social infrastructures to capture the value of distributed intelligence. *Oxford Internet Institute DPSN Working Paper Series*, n. 5, 17 Nov. 2008. Disponível em: <<http://dx.doi.org/10.2139/ssrn.1302893>>. Acesso em: 14 jun. 2012.

FIORE, S.; ALOISIO, G. Special section: data management for eScience. *Future Generation Computer Systems*, v. 27, n. 3, p. 290-291, Mar. 2011. Disponível em: <<http://dx.doi.org/10.1016/j.future.2010.08.012>>. Acesso em: 3 maio 2013.

FONTES, B. A. S. *Redes sociais e poder local*. Recife: UFPE, 2012. 292 p.

FOSTER, I.; KESSELMAN, C. Computational grids. In: FOSTER, I.; KESSELMAN, C. (Ed.). *The Grid: blueprint for a new computing infrastructure*. Los Altos, CA: Morgan Kaufmann, 1999. p. 15-51. cap. 2. Disponível em: <http://www.elsevierdirect.com/companions/9781558609334/appendices/Chapter_02.pdf>. Acesso em: 28 maio 2013.

FREIRE, L. de L. Seguindo Bruno Latour: notas para uma antropologia simétrica. *Comum*, v. 11, n. 26, p. 46-65, jan./jun. 2006. Disponível em: <<http://www.ifcs.ufrj.br/~lemetro/pesquisadores/Leticia%20de%20Luna%20Freire/latour.pdf>>. Acesso em: 13 dez. 2012.

FULLER, S. *The governance of science: ideology and the future of the open society*. Buckingham; Philadelphia: Open University, 1999.

GARVEY, W. D.; GRIFFITH, B. C. Scientific communication as a social system. *Science*, v. 157, n. 3792, p. 1011-1016, Sept. 1967. Disponível em: <<http://dx.doi.org/10.1126/science.157.3792.1011>>. Acesso em: 3 jun. 2013.

GRAY, J. eScience: a transformed scientific method. Transcrição de palestra ministrada por Jim Gray no Conselho Nacional de Pesquisa (EUA), 11 Jan. 2007. In: HEY, T.; TANSLEY, S.; TOLLE, K (Ed.). *The fourth paradigm: data-intensive scientific discovery*. Redmond: Microsoft Research, 2009.

HEY, T.; TREFETHEN, A. E. Cyberinfraestrutura for e-Science. *Science*, v. 308, n. 5723, 6 May 2005. Disponível em: <<http://dx.doi.org/10.1126/science.1110410>>. Acesso em: 5 jun. 2012.

HEY, T.; TREFETHEN, A. E. The UK e-Science core programme and the Grid. *Future Generation Computer Systems*, v. 18, n. 8, p. 1017-1031, Oct. 2002. Disponível em: <[http://dx.doi.org/10.1016/S0167-739X\(02\)00082-1](http://dx.doi.org/10.1016/S0167-739X(02)00082-1)>. Acesso em: 3 maio 2013.

HUFTY, M. Investigating policy processes: the Governance Analytical Framework (GAF). In: WIESMANN, U.; HURNI, H. (Ed.). *Research for sustainable development: foundations, experiences, and perspectives*. Berna: Geographica Bernensia, 2011. p. 403-424. Disponível em: <<http://ssrn.com/abstract=2019005>>. Acesso em: 23 fev. 2014.

HUFTY, M. The governance analytical framework. In: HUFTY, M.; LACROIX, L. (Eds.). *The governance analytical framework: method and case studies (working title)*. National Center for Competence in Research North-South. 2009. In prep. Disponível em: <http://graduateinstitute.ch/files/live/sites/iheid/files/sites/developpement/shared/developpement/mdev/soutienauxcours0809/hufty_Gouvernance/2.1b.Hufty-Eng.pdf>. Acesso em: 23 fev. 2014.

JONKERS, K.; MOYA ANEGON, F. de; AGUILLO, I. F. Measuring the usage of e-research infrastructure as an indicator of research activity. *Journal of the American Society for Information Science and Technology*, v. 63, n. 7, p. 1374-1382, July 2012. Disponível em: <<http://dx.doi.org/10.1002/asi.22681>>. Acesso em: 18 jun. 2012.

KNORR-CETINA, K. The ethnographic study of scientific work: toward a constructivist interpretation of science. In: KNORR-CETINA, K. (Ed.). *Science observed: perspectives on the social studies of science*. London: Sage, 1983. p. 115-140.

KUHN, T. S. *A estrutura das revoluções científicas*. 5. ed. São Paulo: Perspectiva, 1998.

LATOUR, B. *Ciência em ação: como seguir cientistas e engenheiros sociedade afora*. São Paulo: Unesp, 2000.

LATOURE, B. Como terminar uma tese de sociologia: pequeno diálogo entre um aluno e seu professor (um tanto socrático). Tradução de José Glebson Vieira, Leandro Mahalem de Lima e Uirá Felipe. Garcia *Cadernos de Campo*, v. 15, n. 14-15, p. 339-352, jan./dez. 2006. Disponível em: <<http://www.bruno-latour.fr/sites/default/files/downloads/90-DIALOGUE-POR.pdf>>. Acesso em: 13 dez. 2012.

LATOURE, B. *Reagregando o social: uma introdução à Teoria do Ator-Rede*. Salvador: EDUFBA, 2012.

LATOURE, B. Redes que a razão desconhece: laboratórios, bibliotecas, coleções. In: PARENTE, A. (Org.). *Tramas da rede: novas dimensões filosóficas, estéticas e políticas da comunicação*. Porto Alegre: Sulina, 2010. p. 39-63.

LATOURE, B.; WOOLGAR, S. *Laboratory life: the social construction of scientific facts*. Beverly Hills; London: Sage, 1979.

LEITE, M. A. de L. *Marco Aurelio de Lisboa Leite: entrevista* [2 dez. 2013]. Entrevistador: Andre Luiz Appel. São Paulo: [S.n.], 2013. 1 arquivo .mp3 (56 min.).

LOG File. In: CHRISTENSSON, P. *The PC glossary: definitions of computer and internet terms*. PC.net, 9 Apr. 2010. Disponível em: <<http://pc.net/glossary/definition/logfile>>. Acesso em: 7 jun. 2013.

MARTELETO, R. M. Redes e configurações de comunicação e informação: construindo um modelo interpretativo de análise para o estudo da questão do conhecimento na sociedade. *Investigación Bibliotecológica*, v. 14, n. 29, p. 69-94, jul./dic. 2000. Disponível em: <<http://www.ejournal.unam.mx/ibi/vol14-29/IBI02905.pdf>>. Acesso em: 13 dez. 2012.

MEDEIROS, J. da S.; CAREGNATO, S. E. Compartilhamento de dados e e-Science: explorando um novo conceito para a comunicação científica. *Liinc em Revista*, v. 8, n. 2, jul./dez. 2012. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/view/488>>. Acesso em: 29 dez. 2012.

MENA CHALCO, J. P. *Jesús Pascual Mena Chalco: entrevista* [12 nov. 2013]. Entrevistador: Andre Luiz Appel. Rio de Janeiro: [S.n.], 2013. 1 arquivo .mp3 (58 min.).

MERTON, R. K. *The sociology of science: theoretical and empirical investigations*. Chicago: University of Chicago, 1973.

MULKAY, M. J. Some aspects of cultural growth in the natural sciences. *Social Research*, v. 36, n. 1, p. 22-52, Spring 1969. Disponível em: <<http://www.jstor.org/stable/40969943>>. Acesso em: 15 jul. 2013.

NOVAES, S. F. *Sergio Ferraz Novaes: entrevista* [29 nov. 2013]. Entrevistador: Andre Luiz Appel. São Paulo: [S.n.], 2013. 1 arquivo .mp3 (33 min.).

OHNO-MACHADO, L. To share or not to share: that is not the question. *Science Translational Medicine*, v. 4, n. 165, Dec. 2012. Disponível em: <<http://dx.doi.org/10.1126/scitranslmed.3004454>>. Acesso em: 15 mar. 2013. (Seção Comentários).

PARSON, E. A.; KEITH, D. W. End the deadlock on governance of geoengineering research. *Science*, v. 339, n. 6125, p. 1278-1279, 15 Mar. 2013. Disponível em: <<http://dx.doi.org/10.1126/science.1232527>>. Acesso em: 15 mar. 2013.

PAULA, L. S. de. *Leandro Salazar de Paula*: entrevista [10 dez. 2013]. Entrevistador: Andre Luiz Appel. Rio de Janeiro: [S.n.], 2013. 1 arquivo .mp3 (62 min.).

QUIN, L. R. E. XML essentials. *WORLD WIDE WIBE CONSORTIUM (W3C)*. c2010. Disponível em: <<http://www.w3.org/standards/xml/core>>. Acesso em: 2 jul. 2013.

RHODES, R. A. W. The new governance: governing without government. *Political Studies*, v. 44, n. 4, p. 652-667, 1996. Disponível em: <<http://dx.doi.org/10.1111/j.1467-9248.1996.tb01747.x>>. Acesso em: 2 jul. 2013.

RHODES, R. A. W. Understanding Governance: ten years on. *Organization Studies*, v. 28, n. 8, p. 1243-1264, Aug. 2007. Disponível em: <<http://dx.doi.org/10.1177/0170840607076586>>. Acesso em: 27 jun. 2013.

SÀNCHEZ-ARTIGAS, M.; GARCÍA-LÓPEZ, P. eSciGrid: A P2P-based e-science Grid for scalable and efficient data sharing. *Future Generation Computer Systems*, v. 26, n. 5, p. 704-719, May 2010. Disponível em: <<http://dx.doi.org/10.1016/j.future.2009.05.013>>. Acesso em: 3 maio 2013.

SANTORO, A. *Alberto Franco de Sá Santoro*: entrevista [30 jan. 2014]. Entrevistadores: Andre Luiz Appel. Rio de Janeiro: [S.n.], 2014. 1 arquivo .mp3 (45 min.).

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Informação & Sociedade: Estudos*, v. 22, n. 3, p. 179-191, set./dez. 2012. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/12224/8586>>. Acesso em: 28 jan. 2013.

SCHROEDER, R. e-Sciences as research technologies: reconfiguring disciplines, globalizing knowledge. *Social Science Information*, v. 47, n. 2, p. 131-157, June 2008. Disponível em: <<http://dx.doi.org/10.1177/0539018408089075>>. Acesso em: 22 maio 2013.

SHELLARD, R. C. *Ronald Cintra Shellard*: entrevista [18 out. 2013]. Entrevistador: Andre Luiz Appel. Rio de Janeiro: [S.n.], 2013. 1 arquivo .mp3 (71 min.).

SIMAKOVA, E. Making nano matter: an inquiry into the discourses of governable Science. *Science, Technology and Human Values*, v. 37, n. 6, p. 604-626, Apr. 2012. Disponível em: <<http://dx.doi.org/10.1177/0162243911429334>>. Acesso em: 5 nov. 2012.

SOFTWARE framework. In: WIKIPEDIA: the free encyclopedia. [S.l.]: Wikimedia Foundation, 5 fev. 2014. Disponível em: <http://en.wikipedia.org/wiki/Software_framework>. Acesso em: 9 fev. 2014.

SONNENWALD, D. H. Scientific collaboration. In: CRONIN, B. (Ed.). *Annual review of information science and technology*. Medford: Information Today, 2007. v. 41. cap. 14. p. 643-681. Disponível em: <<http://dx.doi.org/10.1002/aris.2007.1440410121>>. Acesso em: 18 out. 2012.

STOKER, G. Governance as theory: five prepositions. *International Social Science Journal*, v. 50, n. 155, p. 17-28, Mar. 1998. Disponível em: <<http://dx.doi.org/10.1111/1468-2451.00106>>. Acesso em: 27 fev. 2013.

TAVARES, D. de O. *Daniel de Oliveira Tavares: entrevista* [26 nov. 2013]. Entrevistador: Andre Luiz Appel. Rio de Janeiro: [S.n.], 2013. 1 arquivo .mp4 (54 min.).

VASILEIADOUA, E.; Vliegenthart, R. Research productivity in the era of the internet revisited. *Research Policy*, v. 38, n. 8, p. 1260-1268, Oct. 2009. Disponível em: <<http://dx.doi.org/10.1016/j.respol.2009.06.005>>. Acesso em: 3 maio 2013.

VELHO, L. Conceitos de ciência e a política científica, tecnológica e de inovação. *Sociologias*, ano 13, n. 26, p. 128-153, jan./abr. 2011. Disponível em: <<http://dx.doi.org/10.1590/S1517-45222011000100006>>. Acesso em: 19 mar. 2013.

WHITLEY, R. Black boxism and the sociology of science: a discussion of the major developments in the field. *Sociological Review Monographs*, v. 18, p. 61-92, Sept. 1972.

WIKI. In: WIKIPEDIA: the free encyclopedia. [S.l.]: Wikimedia Foundation, 1 fev. 2014. Disponível em: <<http://en.wikipedia.org/wiki/Wiki>>. Acesso em: fev. 2014.

APÊNDICE A – Roteiro de entrevistas

Fluxo e Roteiro para Entrevistas

1 Introdução/Apresentação

autoriza a gravação da do áudio da entrevista?

Apresentação do aluno/entrevistador, instituição e orientadoras

Objetivo da pesquisa: analisar as transformações nas práticas de colaboração científica a partir da introdução da e-Science

Objetivo da entrevista: coletar informações sobre a experiência da colaboração CERN – [instituição] como estudo de caso sobre a e-Science e as novas práticas de pesquisa colaborativa

Para efeitos dessa pesquisa, estamos considerando e-Science como:

- **uso intensivo de dados** na pesquisa – obtenção de resultados científicos a partir da análise de grandes volumes dados;
- **uso de computação avançada** – plataformas/recursos, grids/computação distribuída, software para colaboração etc.;
- **colaboração** – colaboração para a produção/geração de dados, assim como no trabalho de análise de dados.

2 Entrevistado(a)

Nome (Instituição) – e-mail

Resumo da biografia

3 Questões

1. Qual sua participação no CERN? [**OPÇÕES DE RESERVA**]
 - a) individual - como pesquisador, grupo, projeto
 - b) institucional – passa pela [instituição], difere da individual...?
2. Desde quando (ano) participa de pesquisa no CERN?
3. Você faz uso de plataformas/recursos da e-Science no trabalho com o CERN? Quais?
 - a) acessa dados armazenados no CERN (remotamente)?
 - b) realiza experimentos/simulações remotamente?
 - c) disponibiliza dados de coletas/experimentos/simulações para outros pesquisadores da colaboração via Internet?
 - d) utiliza equipamento(s) ou software específicos para a colaboração?.....
4. Há quanto tempo usa (usam) as plataformas/recursos? (ou desde quando?)

Condições de participação – pesquisador e grupo

5. Todos os participantes da [instituição] na colaboração utilizam (podem utilizar) essas plataformas/recursos?
6. Existe uma ‘agenda’ de pesquisa da colaboração? Com temas, problemas, objetivos pré-definidos? Quem define/prioriza esses itens? (**VER QUESTÃO 20**)
7. Como se dá a gestão da colaboração? Existe algum contrato, acordo, documento formal etc. que rege(m) a colaboração? Caso exista(m), quem aprova (ou pode sugerir alterações/incrementos nesse(s)) esse(s) documento(s)?
8. Existe alguma cláusula (ou documento) específica sobre o uso dos dados: quem pode acessar o quê? Condições de sigilo? Com quem os dados podem ser compartilhados?
9. Todos os cientistas são informados sobre as condições/contratos de participação? Devem concordar/assinar/manifestar ciência?
10. É necessário algum investimento financeiro para participação na colaboração?
11. É necessário um investimento em infraestrutura (computadores, rede, instrumentos etc.) para participação?

Resultados – partilha/divulgação

12. Quando surgem novos resultados de pesquisa, qual(is) o(s) procedimento(s) para publicação/divulgação? Como é definida a lista de autores? Como se tira proveito do que não gera publicação? Em que momentos é interessante publicar?
13. Existem condições ou orientações específicas acerca de propriedade intelectual, patenteamento de produtos/processos ou licenciamento de software derivados da colaboração?
14. A “partilha” de resultados é afetada pela quantidade investida (pessoas/cientistas, \$\$ e infraestrutura)? [**PERGUNTA RESERVA**]
15. Como você vislumbra a colaboração daqui a 10 anos? Os dados gerados ou as ferramentas/plataformas desenvolvidas continuarão sendo utilizadas? (Pode haver um momento de ‘saturação’ da geração de dados/resultados?)

Cenário Brasil

16. Que “benefícios” a colaboração e o uso das plataformas/recursos da e-Science trazem para o Brasil? (Pesquisadores e país)

17. Que barreiras/dificuldades a colaboração e o uso das plataformas/recursos da e-Science trazem para o Brasil? (Pesquisadores e país)
18. Antes da colaboração, o Brasil tinha recursos ou condições disponíveis para a pesquisa nessa área de física de altas energias?
19. Em termos práticos, o fato de você usar as plataformas/recursos da e-Science altera a forma como você faz/fazia pesquisa? Em que aspectos? [**OPÇÕES DE RESERVA**]
- a) permite o contato com outros pesquisadores?
 - b) gera resultados mais rapidamente?
 - c) precisa aprender a usar equipamento e software?
20. **Remeter à questão 6:** de que forma a e-Science impacta na agenda de pesquisa brasileira e na forma/tipo como o Brasil pesquisa sobre física de altas energias? Quem não tem acesso aos recursos acaba por abandonar esses temas? (Tem condições de ‘concorrer’ com quem está na colaboração? – por recursos/bolsa/apoio/equipamento/laboratório.....)
21. Os recursos ou resultados via e-Science também são aplicados em outros projetos de colaboração/cooperação internos/Brasil?
22. Você utiliza (já utilizou) plataformas/recursos da e-Science (ou similares) em outras experiências/projetos/colaborações?
23. Em quais aspectos a colaboração CERN se diferencia dessas outras experiências? (fora as temáticas/áreas)

APÊNDICE B – Quadro de análise das entrevistas

(1) Elementos de análise:

(a) Tempo e formas de atuação/colaboração

(b) Participação na colaboração

- Acesso: dados, recursos
- Condições de participação, requisitos, expertise...
- Investimento ou contrapartida – também institucional, nacional...
- Questão - da circulação de poder/informação + questão Brasil/Mundo-Centro/Periferia

(c) Resultados

- Publicações
- Dados
- Decisões sobre publicação – processo de análise avaliação, “share” das decisões

(d) Brasil

- Histórico (benefícios, barreiras ou dificuldades), avanços e perspectivas
- Questão centro vs. periferia

(1)	Ronald Shellard (CBPF) – 18/10/2013	Sérgio Novaes (UNESP) – 29/11/2013	Marco Leite (USP) – 02/12/2013	Leandro de Paula (UFRJ) – 10/12/2013	Alberto Santoro (UERJ) – 30/01/2014
(a)	<p>Acelerador que precedeu o LHC – 1988 a 96</p> <p>Políticas científicas – 2008/2009 – expansão do número de países não europeus associados ao CERN – montou-se uma comissão para avaliar a possibilidade de ingresso do Brasil. Coordenou junto com Sérgio Novaes + um diplomata</p>	<p>UNESP – 1989 – Fermilab 2004 – CMS / CERN</p> <p>- Implementaram uma TIER 2</p> <p>- Fazem análises</p> <p>- Serviços da colaboração: identificação de partículas, aprimoramento de software</p>	<p>Desde 1994 – participação não oficial – doutorado – desenvolvimento de instrumentação</p> <p>A partir de 2009 – USP no Atlas</p> <p>- desenvolvimento de detectores e sistemas associados</p> <p>- análise de dados – física de ions pesados</p> <p>- preparação para uma próxima tomada de dados e atualização do acelerador/sistemas de detecção</p>	<p>>> Desde 1993 - Trabalha com Física de Partículas no CERN</p> <p>>> Participa do LHCB</p> <p>- coordena um grupo do IF da UFRJ</p>	<p>>> Estudante de doutorado na França ~1969 – física teórica</p> <p>>> Brasil 1977 – grupo de física teórica no CBPF</p> <p>>> 1982 – colaboração em física experimental de altas energias (Leon M. Lederman) – convite para atuar no Fermilab – 4 experimentos</p> <p>>> consequência – colabora em um experimento do CERN >> CMS – formação de um grupo nacional no CBPF</p> <p>>> 99/2000 – Prof. Titular na UERJ, forma um grupo, e um TIER 2</p>
(b)	<p>Experimentos CERN</p> <p>- São formados por conglomerados de grupos de universidades</p> <p>- contribuem com equipamentos</p>	<p>Física de altas energias grande contribuidora para o avanço da Grid Computing, precisou criar uma estrutura computacional para dar conta do volume de dados gerados pelo LHC</p> <p>TIER 0 – CERN</p>	<p>Atlas – cluster de instituições brasileiras (funciona como um cluster no experimento)</p> <p>Análise dos experimentos do LHC</p> <p>>> computação distribuída</p>	<p>Uso das ferramentas</p> <p>>> Em 1993 trazia os dados em fita para análise no Brasil</p> <p>>> o <i>browser</i> foi criado para guardar a explicação para os cientistas em um local único</p>	<p>Uso de recursos da e-Science</p> <p>>> avanço nas tecnologias tanto de aceleração de partículas quanto dos detectores</p> <p>>> antigamente a detecção era feita por fotografia</p> <p>>> aos poucos as colaborações foram ficando mais ‘eletrônicas’ – com os</p>

<p>- cobre os custos operacionais – eletricidade, água etc. – esse custo é rateado entre as instituições que têm participantes – proporcional ao número de pesquisadores com emprego na instituição? Pesquisadores e pós-doc....</p> <p>– CERN – possui um conselho superior – cada país nomeia um representante – quem paga a cota de associado tem direito a voto – a cota é proporcional ao PNB de cada país</p> <p>- há o investimento direto de países europeus porque o laboratório implica geração de tecnologia</p> <p>- a associação abre perspectiva para a participação nas concorrências de licitações</p> <p>- Grandes desafios científicos (estudos da física de altas energias, climáticos etc.) – dependem de grandes investimentos (uma instituição para organizar experimentos) – a implementação de sensores é muito alto – é preciso desenvolver tecnologias sofisticadas para diminuir os custos dos sensores (da instrumentação) – nenhum país sozinho tem expertise para desenvolver grandes experimentos – por isso dependem de outros participando na colaboração</p> <p>- CERN é um laboratório/grande laboratório</p> <p>>> cada experimento do CERN tem um número muito grande de cientistas, você não consegue colocar todos esses cientistas no mesmo local ao mesmo tempo, por isso o trabalho remoto, organiza-se o trabalho para funcionar remotamente e com isso criam-se ferramentas para viabilizar essa organização. Da mesma forma, se juntar</p>	<p>TIER 1 – Inglaterra, Alemanha, França, Estados Unidos, etc. TIER 2 – centros regionais, como a da UNESP (SPRACE)</p> <p>SPRACE opera junto com Estados Unidos associada à TIER 1 no Fermilab</p> <p>(a estrutura computacional é constantemente avaliada quanto à disponibilidade e estabilidade – tem que estar disponível 98% do tempo... 24/7...)</p> <p>Núcleo/Centro de Computação Científica da UNESP – SPRACE compartilhou com a Universidade a experiência adquirida com o CERN – há uma troca das experiências / dos avanços</p> <p>Um cluster dedicado <i>full time</i> ao CMS (SPRACE) – e uma outra estrutura/espço que atende os projetos da Universidade (GridUNESP)</p> <p>Documento que formaliza a colaboração >> Memorando de entendimento assinado entre a FAPESP e o CERN, estabelece obrigações comuns, sustentabilidade etc. * Durante muitos anos a UNESP foi uma das únicas instituições com essa vinculação formal ao Worldwide LHC Computing Grid (WLCG) BRASIL MEMBRO ASSOCIADO - reunião do Conselho (~15/dez), possibilidade do Brasil ser convidado para se tornar um membro associado do CERN – decisão em nível de Estado – criada comissão pelo Ministro de MCT (Sérgio Rezende) a decisão final passará pelo congresso (recursos viriam diretamente do orçamento da União) – o Brasil usou/usa uma <i>facility</i> que tem um custo enorme... e com isso espera-se uma colaboração....</p>	<p>>> em uma última etapa USP recebe os dados já construídos e reduzidos para análise >> cada conjunto/tomada de dados possui uma 'assinatura' dos processos que antecederam a sua geração</p> <p>Agenda/Objetivos de pesquisa >> cada pesquisador tem uma certa liberdade para fazer a análise que bem entender >> mas deve demonstrar para a colaboração que a análise é consistente, fundamentada e, sendo um tópico de interesse, possivelmente atrairá a atenção de outros grupos</p> <p>Gestão da colaboração >> <i>Boards/Executive board/Collaboration Board/Institution Board</i> >> Cada um discute sobre diferentes assuntos, p. ex., recursos (conferências e alocação a longo prazo), preservação dos dados >> para apresentar uma análise/resultado – um comitê analisa o que será apresentado (em conferências e artigos etc.)</p> <p>>> ATLAS >> para acesso e uso dos dados/informações – tem que ser membro qualificado ou em processo de qualificação (ou estudante), qualificação >> nome aparece nos artigos/lista da colaboração, dados em si tem acesso como membro, mas só o membro qualificado tem a informação do 'como' usar os dados para análise</p> <p>>> para o acesso aos dados brutos é necessária uma requisição, pois o processo de reconstrução dos dados exige uma capacidade de processamento muito alta, gerando um impacto muito grande ao sistema</p>	<p>>> um grupo de 15 a 20 pessoas que se reúne com maior frequência, por videoconferência</p> <p>>> os membros do grupo devem ir ao menos uma vez por ano para o contato presencial e troca de ideias (momento café)</p> <p>Rateio de custo >> professores pagam >> estudantes não pagam</p> <p>O protótipo de um dos componentes (subdetector) do LHCb foi feito no laboratório da UFRJ A construção em série não foi feita no Brasil É um tipo de detector que hoje funciona em aeroportos etc. [transferência de tecnologia]</p> <p>Infraestrutura >> UFRJ desenvolve instrumentação</p> <p>Veio uma comissão do CERN ao Brasil O Brasil mandou uma carta de intenção Reunião no dia 15/12... A partir da associação, o Brasil pode participar de qualquer tipo de licitações com o CERN (cadeira, papel... etc.)</p> <p>*Membro do CERN >> parte das bolsas de formação de recursos humanos para brasileiros - contratos de 2 anos (além dos contratos fixos)</p> <p>Para criar uma Colaboração >> cientista faz uma proposta de experimento, avalia-se: - se ela é cientificamente relevante - se é tecnicamente viável - se ela se sustenta financeiramente >> monta um estatuto ('constituição') >> cada colaboração tem um porta voz/<i>spokesman</i> – eleito por voto dos</p>	<p>dados gravados imediatamente após as colisões e depois transmitidos para computadores, onde é feita uma primeira seleção (<i>trigger</i>)</p> <p>>> no caso do CERN – o LHC já é um acelerador de última geração >> o dimensionamento se dá pela quantidade de energia (de 1 TeV, no primeiro acelerador, até 14 TeV no LHC)</p> <p>>> até determinado momento, os dados ficavam '<i>on site</i>', havia um receio em se distribuir os dados, mas com o crescimento do volume, viu-se a NECESSIDADE de distribuição</p> <p>TIER 0 – Grava TIER 1 – Faz reconstrução TIER 2 – Fazem as análises</p> <p>>> "você tem maior poder de decisão à medida que você influencia o seu investimento" Acesso aos dados >> não havia como utilizar/tratar os dados fora do espaço em que eles eram gerados – a participação/colaboração era <i>in loco</i> >> com o avanço tecnológico houve uma 'revolução' – à medida que se passou a produzir mais dados, mostrou-se impossível tratar os dados em apenas um laboratório/local</p> <p>existem regras de participação nos <i>grids</i> >> certo número de computadores... >> capacidade - <i>storage</i> >> confiabilidade/estabilidade – 24/7</p> <p>Colaboração >> partilha 100% dos recursos e dos custos proporcionalmente ao investimento: >> intelectual, qual tipo de física é melhor de se fazer, ideias >> computação >> trabalho</p>
---	--	--	---	---

<p>toda a capacidade computacional necessária ao processamento dos dados gerados, eles não caberiam no espaço do laboratório >> as hierarquias pensadas são para a organização do trabalho, mas não para a exploração dos resultados práticos – não há tecnologias já desenvolvidas para isso – quando há, as vezes o custo é proibitivo, por isso são montadas frentes para desenvolver alternativas [open hardware]</p> <p>Outros: Fermilab</p> <p>[Mais ao final da entrevista...]</p> <p>Histórico do CERN e o pioneirismo da física de altas energias no sentido da colaboração entre cientistas</p> <p>- projeto Manhattan – os físicos tiveram que aprender a se organizar, a colaborar</p> <p>Com o final da guerra, como consequência do sucesso do esforço de guerra para a geração da bomba atômica, cientistas envolvidos conseguiram muito dinheiro do governo americano para montar os primeiros aceleradores de partículas em Berkeley/Chicago, os europeus se deram conta de que também precisavam de aceleradores, mas não tinham dinheiro/recursos – anos 1950, continente devastado pela guerra – embrião da Comunidade Europeia – 1954 (CERN) – e o CERN se materializa mesmo por volta de 1958 –</p> <p>- biólogos não aprenderam a se organizar como os físicos</p>	<p>DADOS e SFW</p> <p>- <i>sfw</i>: o <i>framework</i> do CERN é muito específico, muito voltado às necessidades da colaboração, não é comprado/comprável</p> <p>- dados:</p> <p>>> sendo membro da colaboração tem acesso irrestrito aos dados – e isso vale para qualquer instituição/membro/pesquisador oficial da colaboração</p> <p>>> saem dos detectores >> são selecionados >> guardados em <i>tape</i> (fita magnética) >> são distribuídos</p> <p>Se um pesquisador precisa de um determinado <i>dataset</i> ele/a pode importar/solicitar, ficando disponível para qualquer finalidade – não há discriminação entre os membros da colaboração</p> <p>14'16" – “[...] <i>mas os dados não tornados públicos. O que as pessoas precisam entender é seguinte: suponha que a gente torne público amanhã os dados, o que vai mudar pra humanidade? Absolutamente nada. Primeiro, vão ser pouquíssimas pessoas que vão conseguir fazer alguma coisa com aquele dado. Porque o dado em si é um monte de bit, pra você manipular aqueles dados você precisa ter as ferramentas corretas, certo? [...] E preciso ter uma estrutura computacional disponível enorme.</i>”</p> <p>[*sobre ferramentas corretas, o pesquisador refere-se ao <i>framework</i> usado pelo CERN para análise e manipulação dos dados]</p> <p>- problema que pode surgir na liberação dos dados – falta de controle de qualidade na produção / atividade científica que é feita usando esses dados</p>	<p>>> outras restrições de espaço utilizado, tempo gasto... etc. tudo é monitorado visando a preservação do sistema</p> <p>Investimento</p> <p>>> contribuição anual com base no número de pesquisadores qualificados – chamado manutenção e operação – calculado com base no PIB do país – pesq valor cheio, pós-doc 75%, estudantes não pagam >> há uma expectativa que os membros também colaborem para manter o sistema funcionando – <u>as análises não contam como investimento</u> – colaborando de outras formas, com infraestrutura/computação, desenvolvimento dos sistemas de detecção etc.</p>	<p>institutos que participam e não pelo valor investido</p> <p>>> tem um conselho</p> <p>>> o CERN ‘fiscaliza’ a colaboração, mas ela não delibera na colaboração</p> <p>Conselho do CERN – cada país tem um membro e o conselho elege o presidente</p> <p>Acesso aos dados</p> <p>>> todos os membros tem acesso a todos os dados</p> <p>>> com a preocupação do final das colaborações, pensa-se na ideia de se deixar os dados abertos para que qualquer um possa utilizar e fazer o que bem entender (não em nome da colaboração), mas há a preocupação em como deixar os dados legíveis para outros pesquisadores</p>	<p>>> pessoal</p> <p>>> material/instrumentação</p> <p>*material/instrumentação</p> <p>Isso já conta como investimento (e o governo financia), nesse sentido, por que investir os R\$ 15 mi?</p> <p>Manutenção & Operação</p> <p>>> no CERN, dá direito a um <i>bureau</i> – uma sala, telefone, papel, computador – material de consumo</p>
---	--	---	---	---

	<p>AGENDA DE PESQUISA</p> <p>- o cientista não decide o montante ou a direção dos investimentos (isso é papel do congresso que também leva em consideração os interesses da sociedade), mas o primeiro trabalha sim numa função de convencimento da importância dos investimentos – e nesse meio, está também a discussão de sobre qual área é mais importante ou 'merece investimento' – não há uma fórmula pra isso – para o Brasil, que tem menos experiência em ciência (desde a década de 1950...) copiar outros países não é um mau negócio</p>	<p>- A validação de qualquer estudo derivado da análise dos dados gerados nas colaborações do CERN depende de uma avaliação prévia dos outros membros da colaboração</p> <p>- quem vai garantir que você manipulou corretamente os dados?</p> <p>- a revisão por pares só vai até certo ponto...</p> <p>- pesquisador explica todo o processo de "controle de qualidade" das publicações da colaboração e como isso fica armazenado em um documento chamado de "TWiki da colaboração"</p>			
(c)	<p>Propriedade intelectual/Licenciamento</p> <p>- CERN não patenteia - resultados são primordialmente científicos, a aplicação prática praticamente não ocorre</p> <p>- com relação ao <i>software</i> – não há ferramentas ou aplicações comerciais disponíveis, por isso o CERN precisa desenvolvê-las [está expandindo os limites, ou <i>pushing the boundaries...</i>] – para essas aplicações não há uso direto – a experiência adquirida nesse processo é uma grande vantagem (10'26")</p> <p>Justificação/relevância [*governança]</p> <p>11'40" - "[...] eu sempre digo, eu sempre enfatizo o seguinte: o que a gente faz não é útil pra sociedade. Um pouco pra chocar a pessoa né? É óbvio que é útil, mas não é minha função torná-lo útil."</p> <p>- publicações vs. avaliação</p> <p>>> Brasil publicou mais <i>papers</i>, mas caiu na relevância científica (p. 11 da</p>	<p>Propriedade intelectual/Licenciamento</p> <p>- no caso do CERN, é mais comum na parte tecnológica, exemplos: invenção de um novo mecanismo para a detecção de uma determinada partícula; uma ressonância magnética, feita com mais eficiência, de forma que foi inventada para detectar determinada partícula no acelerador...</p> <p>- há um setor específico no CERN voltado para a transferência de tecnologia</p> <p>Medida da colaboração/resultados</p> <p>- Artigo não é um bom indicativo da contribuição individual, mas é uma boa medida da contribuição efetiva para o conhecimento</p> <p>- há mecanismos para medir a colaboração individual -</p> <p>Física do Brasil começa com o estudo de física de partículas/de altas energias – não há como construir dois aceleradores do porte do LHC, por isso a colaboração</p> <p>O sucesso de um é o sucesso de todos</p>	<p>Publicações</p> <p>>> para apresentar uma análise/resultado – um comitê analisa o que será apresentado (em conferências/artigos etc.)</p> <p>Propriedade intelectual</p> <p>>> caso chegue a algum processo ou produto que possa ser patenteados – seria como criar uma propriedade intelectual em cima de dados/recursos que estão em domínio público</p> <p>>> deve haver o compartilhamento da propriedade entre a instituição de origem e o CERN</p> <p>Colaboração num horizonte de 10 a 20 anos</p> <p>>> haverá um número muito maior de colisões – vai aumentar o tempo para chegar a resultados</p> <p>>> quantidade de energia – colisões operarão com energia entre 7 a 8 TeV, a partir de de 2015 13 TeV – as análises posteriores vão usar os dados das colisões com as primeiras faixas de energia, por isso deve haver uma preocupação com a preservação dos dados nesse estado, além da possibilidade de criação de um algoritmo</p>	<p>Publicações</p> <p>>> qualquer publicação só pode ser feita se tiver um acordo de toda a colaboração</p> <p><i>Editorial board</i> – 13 pessoas – aprovam cada uma das publicações da colaboração</p> <p>Propriedade intelectual</p> <p>Aluno de doutorado em sanduiche no CERN – solução de problema que pode gerar uma patente – registro internacional em comum acordo</p> <p>Quem assina é a instituição</p> <p>A formação do <i>editorial board</i> NÃO é a afetada pelo volume de investimentos</p> <p>Instituições que pagam muito dinheiro, que colocam mais recurso fazem uma pressão para ter certos cargos de direção, exemplo de Oxford que queria a direção de um detector.... etc..</p>	<p>Publicações</p> <p>>> há regras para participar como membro da colaboração – em sendo membro, você já conta como autor</p> <p>Colaboração num horizonte de 10 a 20 anos</p> <p>>> CMS deve durar mais uns 20 anos</p> <p>>> depende muito da evolução das tecnologias, dos materiais</p> <p>>> discute-se o tipo de aceleração – se mudar para laser daria uma enorme diferença – mas não há tecnologia que sustente</p>

	<p>Rev. Fapesp agosto) – caiu em todas, mas subiu na Física</p> <p>[retorno de investimento]</p> <p>>> em uma visita em que um representante do CERN veio verificar as condições da ciência brasileira >> ao contrário do que se pensa, o CERN não é uma instituição de Fisicos, mas de engenheiros. Há muito mais engenheiros no quadro de contratados no quadro do CERN, são altamente especializados, supercapacitados e para o Brasil é muito importante treinar essa classe de engenheiros</p>	<p>O fracasso de um é o fracasso de todos [**ponte governança]</p>	<p>que permita análises nesses diferentes contextos</p> <p>>> futuramente há também a possibilidade de se utilizar a hospedagem de dados na nuvem (Amazon, Google etc.)</p>		
(d)	<p>Cenário Brasil</p> <p>- Porque é importante para o Brasil ser associado: as empresas brasileiras de tecnologia podem participar de concorrências em licitações e editais do CERN, juntamente com empresas de países europeus</p> <p><i>12'21"</i> – “No Brasil, por exemplo, você olha agora o ministério de C&T está muito interessado em inovação e o que é inovação? Pra que serve a inovação? Não é? Pensa dois segundos. Claro que inovação, qualquer indústria, [por exemplo], se eu sou um industrial eu tô interessado ali, [mas], tem uma graduação brutal do que seria inovação, eu, por exemplo, se eu sou um pequeno industrial eu quero melhorar um pouquinho a eficiência, como é que eu uso um operário a menos... tem esses pequenos ganhos que é a grande parte da inovação no dia-a-dia das empresas. Aí depois você tem inovações que são mudanças um pouco de paradigma nos mecanismos de produção, que é mais de engenharia sofisticada. Tem aí, vamos dizer, as grandes mudanças, a</p>	<p>Cenário Brasil</p> <p>>> há um aprendizado que vai muito além dos conhecimentos a respeito da física de altas energias...</p> <p>- Há uma capacitação de gerenciamento / definição de projetos / metas – computação, eletrônica, tratamento de dados</p> <p>- para o físico experimental há a necessidade de diferentes habilidades e expertise</p> <p><i>26'47"</i> – “A universidade hoje em dia tem uma estrutura computacional respeitável, para qualquer área fazer ciência devido ao fato dela ter tido um grupo participando em física de altas energias que se dispôs a implementar essa estrutura para toda a universidade”</p> <p><i>27'17"</i> – “Tinha um professor, quando eu implementei o GridUNESP, ele veio falar pra mim que ele que tinha uma molécula,</p>	<p>Cenário Brasil</p> <p>>> o processamento distribuído é feito exclusivamente com o framework do Atlas – nesse caso o acesso/contato é quase que compulsório</p> <p>>> o próprio ambiente [colaborativo], você precisa do suporte do contato com outros pesquisadores, o que é muito bom para os estudantes, pois faz com que os mesmos desenvolvam uma independência, uma pro-atividade</p> <p>>> ganho de conhecimento/experiência</p> <p>BARREIRAS</p> <p>>> carência de um projeto de desenvolvimento científico de longo prazo / todos as chamadas, editais são muito sazonais</p> <p>Brasil não teria condições de atuar em física experimental de altas energias fora da colaboração – mesmo países europeus ou os Estados Unidos – também porque o tempo de formação para poder atuar nos experimentos leva em torno de 15 anos.</p>	<p>Cenário Brasil</p> <p><i>37'39"</i> – é um laboratório de pesquisa – só tem pesquisadores – independente da qualificação mestre/doutor – esse trânsito/essa interação é enriquecedora – falando dos benefícios de formação dos alunos...</p> <p>Não há guerras nacionalistas na colaboração</p> <p>>> até há disputas, mas de quais ideias são melhores, qual a melhor forma de se fazer determinada tarefa...</p>	<p>Cenário Brasil</p> <p>No Brasil há um dificuldade muito grande para participação</p> <p>>> REDE – RNP não mais atende</p> <p>TIMING (física de altas energias)</p> <p>>> há um tempo oportuno para atuação/ação – e nesse timing os recursos devem estar à mão / disponíveis – a administração científica brasileira não tem essa noção – e comunidade brasileira de física de altas energias é um muito pequena comparada com outras áreas</p> <p>BRASIL MEMBRO ASSOCIADO</p> <p>*somente proposta*</p> <p>>> Santoro considera ruim – não há materialização do retorno de investimentos para o Brasil – não há garantias seguras que os R\$ 15 mi investidos vão retornar ao Brasil</p> <p>Benefícios às indústrias??</p> <p>>> o benefício deve ser mesmo às indústrias??</p> <p>Benefícios para o Brasil</p> <p>>> no cenário atual da medicina brasileira os equipamentos para detecção de tumores de câncer de</p>

<p>gente trabalha muito mais no investimento que vai esse tipo de ciência, são as grandes mudanças de paradigma. Essas [inovações] não tem como você prever [...] Esses equipamentos que a gente usa pra detectar partículas, que foram inventados lá na década de 1960, hoje em dia estão na..., você vai fazer uma tomografia e ali há um detector de partículas, tem físicos de altas energias envolvidos? Não, hoje em dia não tem, não é? Mas são decorrentes dessa, vamos dizer, desse investimento”</p> <ul style="list-style-type: none"> - a necessidade de comunicação entre os físicos de altas energias no Brasil fomentou a ampliação das redes lógicas (via RNP) - a participação na colaboração é essencial para o avanço no campo da física de altas energias - o fluxo de tecnologia acontece de maneira natural – há reciprocidade entre os resultados para os participantes na colaboração - a relação entre universidade / indústria se dá, em maior parte, no desenvolvimento em instrumentação científica 	<p><i>uma proteína na gaveta dele há alguns anos, que ele não tirava da gaveta porque era uma coisa extremamente interessante.... ele não tirava essa proteína da gaveta pra fazer estudo de enovelamento de proteína porque ele não tinha capacidade computacional pra fazer isso e agora ele iria tirar a proteína da gaveta. Então isto mostra um pouco qual era o interesse, é gente se lançar em áreas... tem áreas [em que] o poder computacional é impeditivo, certo, você não pode fazer nada naquela área se você não tiver uma excelente capacidade computacional.”</i></p>			<p>mama não permitem um diagnóstico seguro / decisivo para o médico – contudo já existem equipamentos que permitem um melhor diagnóstico sem necessidade de cirurgia</p> <p>>> esse objeto é uma aplicação usada em detectores</p> <p>>> consequências: materiais que derivam da necessidade de uso na física de altas energias – chapas de aço com maior número de propriedades interessantes... maior flexibilidade etc. – evolução dos materiais</p> <p>Dificuldades para o Brasil</p> <p>>> física de altas energias – são poucos laboratórios e todos concentrados – com isso a necessidade de viagem – tem \$\$ mas o esforço para conseguir recurso é monstruoso – formulários e burocracia... [*provisão de recursos a longo prazo]</p>
--	---	--	--	---

(1)	Jesús Mena-Chalco (UFABC) – 12/11/2013	Daniel Tavares (LNLS) – 26/11/2013	Roberto Marcondes (USP) – 02/12/2013
(a)	<p>Grupo Computação e Matemática (Nap-USP)</p> <ul style="list-style-type: none"> - Biologia e Medicina – integração de bancos / genoma - Estudo de redes brasileiras de coautoria - estudo sobre prospecção de dados acadêmicos (bibliometria), ciëntometria - Desenvolvimento do ScriptLattes 	<p>Fez intercâmbio em uma instituição da França durante a graduação e com isso conheceu o CERN (mas não tem relação com o trabalho atual)</p> <ul style="list-style-type: none"> - No LNLS – atua no Projeto Sirius – novo acelerador com proposta de entrega para 2016 – com parte significativa da instrumentação desenvolvida no Brasil - Grupo de instrumentação científica – eletrônica para instalações/laboratórios científicos 	<p>Grupo Computação e Matemática</p> <ul style="list-style-type: none"> - Biologia e Medicina – integração de bancos / genoma + Ciências Sociais e Humanas (mais recente) Núcleo de pesquisa em parceria com outros projetos de outras áreas

<p>(b)</p>	<p>A e-Science está baseada nos avanços da tecnologia. >> a tecnologia permitiu a colaboração</p> <p>Havia uma preocupação maior com a obtenção de dados – havia esforços maiores para a geração de dados</p> <p>Atualmente, existem os dados, e os esforços estão direcionados à análise e exploração dos dados</p> <p>Dados são divididos em blocos e distribuídos para diferentes espaços de análise</p> <p>Uso dos dados</p> <p>25'50" – “[...] <i>armazenar os dados processados e anotados é fácil, seria uma coisa rotineira, uma coisa que é possível desde o ponto de vista computacional, descobrir conhecimento nesses dados, aí é a parte complicada.</i>”</p> <p>ACESSO AOS RECURSOS DE GRID - não há atualmente o receio quanto a distribuir o processamento em computadores pessoais – isso era uma preocupação antiga</p> <p>- atualmente há empresas especializadas na oferta desse serviço – em <i>cloud/Amazon</i> – você pode comprar a capacidade de processamento – ‘democratização’ da ciência – não só mais países ou instituições que têm equipamentos de ponta é que possuem a capacidade de avançar</p>	<p>A partir de abril/2012 Acelerador Sirius será todo desenvolvido no Brasil</p> <p>Opções - cooperação - compra/transferência de tecnologia com empresas</p> <p>>> aí houve contato com o CERN na iniciativa do Open Hardware (OH) – não há iniciativa de colaboração formal >> usa licença Open Hardware do CERN >> faz uso do framework do CERN para desenvolvimento de OH >> usam o repositório do CERN (nesse repositório qualquer empresa/grupo pode participar/compartilhar) >> ganham com a troca de experiência com as equipes de desenvolvimento do CERN</p> <p>Já desenvolveram três placas na iniciativa do OH</p> <p>Faz a especificação (das placas/do hardware) no Brasil e outros parceiros (o próprio CERN) pode produzir</p> <p>Vantagens do OH - customização - transferência de tecnologia - proteção pela Propriedade Intelectual</p> <p>O LNLS faz a especificação da placa e com isso tem o direito de licenciar/garantir os direitos – e com isso optaram pelas licenças open</p> <p>A abertura/novidade com relação às atividades do grupo, em seu início, deu maior liberdade ao uso das licenças abertas – mas aos poucos estão surgindo regulamentações específicas. >> a abordagem open garantiu sucesso e um avanço muito mais rápido – não trabalharam por conta própria, puderam contar com a colaboração de outras equipes para o desenvolvimento e as especificações >> durante o desenvolvimento das placas teve bastante input do pessoal de outros grupos – inclusive do CERN >> não pode garantir que é a melhor abordagem, sem ter também atuado junto a uma empresa</p> <p>Brasil como membro associado CERN</p> <p>>> com as especificações open hardware – pode haver um cenário em que duas empresas, uma brasileira e outra europeia, se propõe ao desenvolvimento/fabricação daquele hardware, a empresa que entregar um melhor produto, mais ‘robusto’, de melhor qualidade fica com a licitação...</p>	<p>Todos já utilizavam mesmo antes de nominar</p> <p>Agenda/Objetivos de pesquisa Quando há uma colaboração – quais são os fenômenos em que o parceiro (biólogo etc.) está preocupado, qual o tipo de dado, como é armazenado, qual o tipo de ruído, quais outras fontes de dados.... Modelagem >> até a entrega dos resultados / interpretação do pesquisador</p> <p>Uso dos dados/diretrizes >> é bem definido em conjunto com o parceiro/de acordo com a forma de trabalho das diferentes áreas/no caso da biologia, um dos primeiros passos é publicar nos bancos abertos / no caso de dados de mudanças climáticas</p> <p>CERN – a discussão sobre o compartilhamento de dados já ocorreu (inclusive no início da web), mas atualmente já é bem organizado, hoje todas as diretrizes já estão bem definidas, todos já tem tudo bem claro na forma de trabalhar</p> <p>9'40" – “[...] <i>o que acontece muito com o CERN, é que é uma comunidade muito específica né, tem uma outra característica que também é verdade nesse caso que é o seguinte: o CERN em si, por ser um grande experimento mundial, eles precisam se organizar de uma certa maneira pra poder financiar esse experimento e é, normalmente, um projeto que envolve um prazo muito longo, então eles têm que planejar o negócio pra o ano que vem, daqui a cinco anos, daqui a dez anos, daqui a quinze anos... e assim por diante, e que custa muito dinheiro, por isso você tem que ter uma colaboração internacional [...]</i>”</p> <p>No caso do genoma, você consegue pensar em pequenos projetos/iniciativas... há a possibilidade de compra de dados para análises</p> <p>Investimentos nas colaborações do núcleo >> FAPESP + USP – financia equipamentos infra computacional necessária as questões colocadas em pauta, recursos humanos/bolsas / não mantém as pesquisas dos parceiros</p>
------------	---	--	--

<p>(c)</p>	<p>Mensuração e avaliação das colaborações</p> <p>O conhecimento (decorrente das colaborações) é finalmente registrado em artigos científicos – e os artigos são em coautoria</p> <p>Como avaliar a participação nas coautorias – artigos com muitas coautorias – e com essas sendo internacionais – são os mais bem avaliados?</p> <p>A qualidade dos artigos de algumas áreas já está ‘saturada’, na Física por exemplo, ‘todos’ os artigos são internacionais, são publicados em revistas A1 [em coautoria também] – número de citações não serve também, pois os artigos recebem um grande número de citações</p> <p>- o objetivo do trabalho colaborativo não são as publicações, estes são produtos, >> a criação de metodologias, formas de operacionalização do trabalho colaborativo que podem ser aproveitados em projetos da mesma natureza, mas esse aproveitamento ainda é muito restrito às questões/problemas específicos (não exclusivos das áreas do conhecimento), por exemplo, as metodologias geradas para o mapeamento do genoma humano não se aplicam à Física >> há a criação de ferramentas muito específicas para cada problema ou trabalho >> não há quem organize essa colaboração interdisciplinar / intergrupos</p> <p>Bancos especializados - NCBI - de astronomia - ArXiv - PubMed</p> <p>Embora funcionem de forma parecida, esses bancos não são interoperáveis, não se falam</p> <p>Há uma reinvenção da roda diversas vezes, para atender a cada problema de pesquisa que surja</p> <p>- Dificuldade na preservação digital a longo prazo – recursos – há uma afluência maior de recursos enquanto a resolução do problema está andamento, mas após isso, os recursos para a guarda e preservação dos dados são escassos</p>	<p>Interesse nos resultados</p> <p>>> Alemanha potencialmente usará a eletrônica do LNLS em seus sistemas >> não há encomenda prévia – aos poucos vão surgindo interessados...</p> <p>Desenvolvimento de <i>hardware</i> num horizonte de 10 a 20 anos >> barreira – desenvolvem ferramentas abertas MAS não há <i>software</i> aberto para dar suporte a esse desenvolvimento >> o CERN tem uma iniciativa de desenvolvimento de <i>software</i> aberto para desenvolvimento de <i>hardware</i> aberto KiCad – uma espécie de Auto Cad</p>	<p>Resultados</p> <p>>> desenvolvimentos metodológicos são compartilhados na colaboração do núcleo com os parceiros</p> <p>>> se for avanço específico na área, a decisão é dos parceiros</p> <p>Colaboração num horizonte de 10 a 20 anos</p> <p>>> Pode haver a mesma forma de estrutura e armazenamento de dados >> na parte de desenvolvimento científico, deve nascer uma cultura que vai reconhecer o dado como um ativo importante de tudo que está acontecendo</p> <p>Sobre as ações de curadoria e necessidade de tratamento de dados - diretrizes</p> <p>Necessidade de mudança cultural - parte das comunidades - mistura de bottom>up e up>down</p> <p>17’48” – “<i>Tem uma parte que as comunidades vão criando, vão estabelecendo, mas tem uma parte que, por exemplo, o governo e as agências de fomento podem definir. Então se você pega iniciativas como a NSF, por exemplo, tem várias dessas iniciativas, que começam a exigir que um projeto de pesquisa ao ser apresentado ele deve prever o que vai acontecer com os dados, quais dados vão ser gerados e como esses dados vão ser preservados, aí [com] isso você começa a embutir uma cultura junto às diferentes comunidades, porque daí para o pesquisador ter acesso àquele auxílio ele tem que prever o que ele vai fazer com os dados, assim como ele prevê que ele tem que atualizar o currículo Lattes dele porque se não ninguém, simplesmente, vai financiar um novo projeto dele</i>”</p> <p>>> dados viram parte dos resultados intelectuais de um processo científico</p>
------------	--	--	--

	<p>Colaboração num horizonte de 10 a 20 anos</p> <ul style="list-style-type: none"> - Ampliação da colaboração virtual - Pode haver mudanças na forma / nas políticas como publicar os trabalhos - Mas a forma de se fazer o trabalho, a prática, será basicamente como a vigente – pesquisadores se reúnem, discutem o problema, dividem tarefas para resolução de problemas comuns - A autoria dos trabalhos – autores se reúnem e fazem um trabalho >> transição para um modo wiki >> não há um autor principal – quando se resolve um problema, será que precisa de um autor? - Não há artigos sem autores, pode haver artigos sem autores no futuro? – e nesse caso, quem assumiria a responsabilidade - As ferramentas deveriam ajudar a escolher os trabalhos que devem ser lidos – pois a capacidade de leitura da humanidade será a mesma de hoje, embora o número de artigos continue crescendo <p>O que vai mudar (ou está mudando):</p> <ul style="list-style-type: none"> - Colaboração - Forma de publicação - Forma de avaliação <p>E o que não vai?</p> <p>Licença, patente, aplicação comercial</p> <p>Um dos projetos / um dos modelos que deram certo para o estudo da e-Science foi a experiência da astronomia, pois não se vislumbram finalidades comerciais – questões políticas e de licença estão fora do horizonte</p>		
(d)			<p>>> para os grupos brasileiros que participam estão expostos ao que está acontecendo de pesquisa de primeira linha naquela área >> estão integrados na parte de e-Science também, infraestrutura</p> <p>Sobre a questão dos potenciais riscos envolvendo o compartilhamento, a cessão de dados</p> <p>23'15" <i>“Quando um consórcio brasileiro faz parte, quando um grupo brasileiro faz parte de um consórcio que ainda vai criar o sistema, é melhor ainda, porque daí nós fazemos parte, desde o começo, do design do experimento e tudo mais, o que, de certa forma, reconhece os direitos dos grupos brasileiros. Talvez uma limitação brasileira é que nós não temos é... – por várias razões, comunidade pequena, irregularidades do financiamento, falta de ambição de muitos dos pesquisadores da nossa comunidade – é difícil imaginar algum projeto desse que tenha sido criado no Brasil ou que já esteja explorando isso daí, liderando isso, seria ótimo se a gente tivesse grupos que procurassem liderar coisas desse tipo.”</i></p>