

## **A CIÊNCIA ABERTA NA GESTÃO DE DADOS DE PESQUISA: UMA ABORDAGEM PARA ABRIR A CAIXA PRETA DA PRODUÇÃO CIENTÍFICA**

**JANETE SALDANHA BACH ESTEVAO**

Universidade Tecnológica Federal do Paraná, Programa de Pós-Graduação em Tecnologia e Sociedade, Brasil,  
janete.estevaso@ufpr.br

**ELAINE MANDELLI ARNS**

Universidade Tecnológica Federal do Paraná, Programa de Pós-Graduação em Tecnologia e Sociedade, Brasil,  
elaine.arns@ifpr.edu.br

**FAIMARA DO ROCIO STRAUHS**

Universidade Tecnológica Federal do Paraná, Programa de Pós-Graduação em Tecnologia e Sociedade, Brasil,  
faimara@utfpr.edu.br

### **RESUMO**

Discute-se, neste artigo, eminentemente bibliográfico, como a Ciência Aberta tem constituído uma nova abordagem para o processo de geração de conhecimento científico, com base na forma colaborativa que a produção científica tem sido criada e comunicada. Considera-se que os tradicionais resultados de pesquisa como artigos, dissertações e teses, ainda que disponíveis em acesso aberto sejam comparados a uma caixa-preta, de acordo com a Teoria Ator-Rede (TAR), constituindo o resultado acabado da análise de um determinado pesquisador ou grupo de pesquisadores. Este artigo busca evidenciar como a disponibilidade de dados de pesquisa interfere no processo de geração e de comunicação científica. Para tal utiliza-se das concepções da TAR na identificação do papel dos dados de pesquisa dentro da rede, composta por elementos humanos e não-humanos, os pesquisadores, as universidades, as agências de fomento e os sistemas de busca e de recuperação de informação. Como resultados, demonstra-se como a disponibilidade dos dados da pesquisa produz novos recursos de informação, viabilizada pela retroalimentação da atividade científica. Essa disponibilidade altera a dinâmica da rede na medida em que o pesquisador pode valer-se de novos insumos de informação, ou seja, os conjuntos de dados de pesquisa disponibilizados por outros pesquisadores e também precisará, ele mesmo, disponibilizar os dados que produziu. O acesso a estes recursos democratiza a produção científica e beneficia toda a sociedade.

**Palavras chave:** Ciência Aberta. Acesso Aberto. Gestão dos Dados da Pesquisa. Comunicação Científica. Teoria Ator-Rede.

## 1. INTRODUÇÃO: CIÊNCIA ABERTA E A GESTÃO DOS DADOS DA PESQUISA

Em dezembro de 2016, oito importantes agências de financiamento, dentre as quais a Bill & Melinda Gates Foundation, anunciaram o lançamento do Open Research Funders Group (ORFG), uma parceria concebida para aumentar o acesso aberto aos resultados de pesquisa (Open Research Funders Group, 2016). Com um investimento de quase US \$5 Bilhões de dólares anuais, o ORFG visa aumentar o compartilhamento aberto de artigos e dos dados de pesquisa, acelerando o ritmo e reduzindo as lacunas na partilha de informação, incentivando a inovação e promovendo a reprodutibilidade (Open Research Funders Group, 2016). Diante de tal iniciativa, depreende-se a notória relevância da ciência aberta, não apenas para dar visibilidade aos resultados das pesquisas, mas também aos dados que fundamentam estes resultados em artigos e textos científicos financiados por estas instituições.

Diversas agências reconhecem a importância dos dados de pesquisa e exigem uma adequada gestão destes dados, evidenciando a necessidade de assegurar meios e mecanismos para preservação e reutilização dos mesmos. Exemplificando, em 2005, o United States National Institute of Health (NIH) criou sua política de acesso e em 2007 foi aprovada uma legislação nacional que exigia da NIH o acesso aberto de todas as pesquisas financiadas pelo instituto, sendo que todos os artigos deveriam se depositados no repositório institucional PubMedCentral (Grigg, 2015).

A National Science Foundation (NSF) dos Estados Unidos requer, igualmente, desde 2011 que as pesquisas financiadas sejam acompanhadas de um plano de gestão de dados e a disponibilidade desses dados seja em acesso aberto (Grigg, 2015). O programa Europeu Horizon 2020 (H2020), desde 2013, também faz essa exigência. Este é o maior programa de pesquisa científica já criada no bloco europeu, prevendo um investimento de €80 bilhões entre 2014 a 2020 (European Commission, 2014). A garantia de publicidade em acesso aberto para os resultados das pesquisas do H2020 é um dos princípios âncoras do programa, sendo uma exigência prevista no regulamento do programa e um critério para a participação e para o pagamento de bolsas aos candidatos (European Commission, 2016a, p. 2016-2019; 2016b).

A ciência aberta, portanto, representa uma nova abordagem para o processo científico com base no trabalho cooperativo e em novas formas de difusão de conhecimentos por meio de tecnologias digitais e das novas ferramentas colaborativas (European Commission, 2016c, p. 33). A Organização para Cooperação e Desenvolvimento Econômico - OCDE (2015, p.7) conceitua ciência aberta como uma forma de "permitir o acesso aos resultados primários de pesquisa financiada com recursos públicos - publicações e os dados da pesquisa - disponíveis ao público em formato digital sem ou com restrição mínima". Porém, é mais do que isso, ciência aberta é frequentemente conceituada como um termo abrangente que envolve vários movimentos tendentes a suprimir as barreiras para compartilhar qualquer tipo de resultados, recursos ou dados da pesquisa, métodos ou ferramentas, em qualquer fase do processo de investigação científica (Organização para Cooperação e Desenvolvimento Econômico 2015).

Dada a complexidade de implicações no compartilhamento de dados abertos e no estabelecimento de uma cultura de ciência aberta, o objetivo deste artigo está delimitado em evidenciar se a disponibilidade de dados da pesquisa, um aspecto da Ciência Aberta, afeta a dinâmica do ciclo de pesquisa para geração de conhecimento científico. Para tal será necessário identificar os atores envolvidos na produção científica, a partir dos principais conceitos do corpus da Teoria Ator-Rede (TAR), privilegiando a descrição dos atores, que revelam a rede e sua

dinâmica.

A gestão dos dados está diretamente ligada à sustentabilidade da pesquisa financiada com recursos públicos. Na gestão de dados de pesquisa, o sucesso depende dos arquivos eletrônicos e seus formatos, da infraestrutura tecnológica para armazenamento, dos sistemas de recuperação da informação e visualização de dados, além das políticas públicas e institucionais que regulam todo o “desenvolvimento científico” (Callon, 1993, p. 92). Outros aspectos como preservação digital, processos de curadoria de dados, protocolos e interoperabilidade de sistemas, qualidade dos dados, governança institucional, barreiras comportamentais e culturais em relação ao compartilhamento dos dados, direito autoral e propriedade intelectual, serviços de apoio à gestão dos dados, são exemplos de elementos heterogêneos que, apesar de também constituírem um ambiente complexo para a gestão dos dados da pesquisa, não serão abordados neste trabalho.

Esta atuação atribuída aos dados revela-os como um elemento não humano presente na rede da comunicação científica, demonstrando a sua importância no processo de construção do conhecimento. É justamente a agência atribuída para além dos atores sociais tradicionais que sustenta todo o corpus da TAR (Sayes, 2014). Nesta direção, escolheu-se a Teoria Ator-Rede (TAR) como abordagem que permitirá seguir os atores de forma indistinta, somada aos demais procedimentos metodológicos que delineiam esta pesquisa, descritos a seguir.

## **2. METODOLOGIA**

Os procedimentos metodológicos utilizados para o desenvolvimento desta pesquisa seguiram a seguinte sequência de apresentação: (a) revisão da bibliografia de base da Teoria Ator-Rede; (b) levantamento bibliométrico e seleção de artigos referente à Gestão de Dados da Pesquisa; (c) análise do conteúdo dos artigos que tratam das implicações da Gestão de Dados no ciclo de geração do conhecimento científico; (d) descrição do papel dos dados na dinâmica do ciclo de pesquisa, enquanto elemento da rede de produção e comunicação científica (composta, dentre outros elementos, por pesquisadores, universidades, agências de fomento, entre outros detalhados na seção 5).

A partir da revisão bibliográfica, foram definidos os conceitos e os descritores utilizados no levantamento bibliométrico nas bases de dados disponíveis no Portal da Capes. Os termos de busca utilizados foram os mais recorrentes na literatura internacional para a gestão de dados de pesquisa, como Data Research Management (Gestão de dados da pesquisa), Data Curation (curadoria de Dados) e Research Data Lifecycle (Ciclo de vida dos dados).

Buscou-se identificar quais alterações os dados de pesquisa, quando disponíveis, podem promover no ciclo de pesquisa da produção científica contemporânea. Nesta perspectiva, para verificar se a disponibilidade dos dados em si atua e influencia na dinâmica da rede, seguiu-se a fundamentação da Teoria Ator-Rede. Para este trabalho, a identificação da rede de produção do conhecimento científico se dará a partir da descrição dos atores envolvidos no processo de comunicação científica, buscando evidenciar a justaposição entre elementos humanos e não-humanos. Para a contribuição da descrição, se faz necessário extrair o conhecimento da teoria base deste trabalho, a Teoria-Ator Rede enquanto ferramenta de análise social (Callon, 1993).

A proposta da TAR para a sociedade é de que elementos não-humanos se somam aos humanos para formar o ‘coletivo’, ideia que transcende a própria noção de ‘sociedade’. A ideia de coletivo ao invés de sociedade é um dos principais conceitos da TAR, dentre outros elementos descritos

na próxima seção.

### **3 REVISÃO DE LITERATURA: VISÃO GERAL DOS PRINCIPAIS ELEMENTOS DA TEORIA ATOR-REDE (TAR)**

Para Latour (2001) a ideia de construção da sociedade se associa à noção de coletivo, em que a “sociedade é a interação de trocas de propriedades de um coletivo de humanos e não-humanos” (Latour, 2001, p. 218). Para esse autor, a sociedade é construída, mas não socialmente, antes “os humanos estenderam suas relações sociais a outros atuantes com os quais trocaram inúmeras propriedades, formando coletivos” (Latour, 2001, p. 227). Latour (1998) pode ser considerado um construtivista, mas para ele, o conhecimento é construído com o envolvimento de uma rede heterogênea de materiais, representações, financiamentos, pressões econômicas, disputas políticas, em uma cadeia infundável de elementos.

A rede refere-se a fluxos, circulações, alianças, movimentos, não se remetendo a uma entidade fixa, mas às conexões ou às interações que transformam os recursos dispersos em uma rede (Latour, 1986, p.180). A materialidade relacional da TAR é a ideia de que tudo é constituído em interação e que nada tem existência definida fora destas interações (Latour, 2006). São as interações dos atores que formam um conjunto relativamente estável de entidades associadas que alcançam certa durabilidade. Unidas, essas entidades são capazes de desassociar a forças hostis (naturais e sociais) encontradas, sem serem elas mesmas desassociadas (Law, 1993, p. 128). A abordagem da TAR pressupõe que sempre há um grau de divergência entre os elementos da rede, em escalas diferentes. A associação destes elementos em forma de redes autossustentáveis é que fazem com que esses elementos sejam hábeis em resistir à dissociação.

As redes são construídas por atores que têm ação, interferência, ou na terminologia de Latour, tem “agência”. Isso os torna “actantes”, ou seja, só pode ser considerado um ator aquele (ou o que) exerce ação nas relações com outros elementos da rede. O ator é aquele que promove a aliança, ou seja, quem ou o quê realiza o aliciamento dos demais entes para um objetivo (Sismondo, 2010). O conceito de ator é central na TAR, pois é por meio da sua força em mobilizar outros que a rede funciona. É do ator a capacidade de arregimentar esses elementos heterogêneos, balancear as suas forças, subjugar as demais forças hostis e resistir às ameaças externas, sejam estas naturais ou de qualquer outra característica, sociais, políticas, econômicas, transformando-as e associando-as ao projeto, combinando as suas funções das entidades que permite a estabilidade da rede (Law, 1993).

Atores são as entidades que fazem coisas, que atuam, portanto, só são consideradas existentes se deixam ‘provas’ ou ‘rastros’ da sua agência (Latour, 2006). Se o elemento não faz a sua presença ser percebida pela sua influência na estrutura da rede, então se pressupõe que para aquela rede, este elemento não existe (Law, 1993, p. 131; Latour, 2006). Sendo assim, a extensão da rede é definida pela presença de atores que são capazes de se fazerem percebidos de uma maneira individual (Law, 1993, p. 132). Seguir os atores ou os cientistas significa descrever o que eles fazem, tanto quando eles tentam transformar a sociedade, quanto quando eles buscam construir o conhecimento científico ou o sistema tecnológico.

Na designação de Latour (2001, p. 201), o termo sociedade é expresso como um ‘coletivo’: “agora que os não humanos já não se confundem com objetos, talvez seja possível imaginar um coletivo no qual os humanos estejam mesclados com eles”. Em diversos exemplos da agência

atribuída aos elementos não humanos (Callon, 1993; Callon, 1986; Latour, 1995; Law, 1993), demonstram-se como estes elementos tiveram uma “atuação” no processo e nos resultados finais dos objetivos de cada projeto. A teoria implica no aliciamento de aliados para o estabelecimento de uma rede estável em que os resultados dependem no gerenciamento de todas as entidades que corroboram para o mesmo objetivo.

Para a TAR, os não-humanos são elementos necessários para estabilizar o coletivo de humanos (Sayes, 2014, p. 137). Segundo Sayes (2014), a controvérsia em torno da agência atribuída aos não-humanos está no fato da compreensão parcial do real significado desta agência. A partir das concepções que o termo não humano representa na TAR, remete uma crítica à tradição filosófica da relação sujeito-objeto sendo tratada de forma radicalmente distinta ou polarizada. Para entender a forma de agência dos não-humanos, Latour (2001) apresenta o conceito de mediação técnica.

### **3.1 Agência de elementos não-humanos: o conceito de mediação técnica**

Latour denominou a atuação não humana como uma mediação técnica (Latour, 2001, p. 205-218). A primeira forma da mediação técnica são os pressupostos da agência dos não-humanos na medida em que estes interferem na ação humana, atuando como mediadores e realizando a translação de objetivos. No caso do cidadão que pega uma arma, um terceiro agente surge da fusão dos dois e nasce um vínculo que não existia. A translação modifica os dois originais e o objetivo da ação. Não são nem as pessoas nem as armas que matam. A responsabilidade deve ser dividida entre os vários atuantes. A arma interfere sobre o humano, que interfere sobre a arma (Latour, 2001, p. 205).

A mediação técnica também acontece como partes de uma composição, em que a ação é uma das propriedades das entidades associadas (Latour, 2001, p. 208-209). Quem pratica a ação não é uma propriedade de humanos, mas de uma associação de atuantes. Os objetivos são redefinidos por associações com atuantes não-humanos e que a ação é uma propriedade da associação inteira, não apenas dos atuantes chamados humanos (Latour, 2001, p. 210, 211). Por exemplo, voar é uma propriedade de toda a associação de entidades, que inclui aeroportos e aviões, rampas de lançamento e balcões de venda de passagens.

O entrelaçamento de tempo e espaço, terceira forma da mediação técnica, são agrupamentos de atores de diferentes ordens temporais e espaciais que reúne uma grande variação de número de atuantes e composição de objetos. Como quarto significado de mediação técnica está à transposição da fronteira entre signos e coisas (Latour, 2001, p. 212-216), que provoca o deslocamento de significado. Um motorista altera seu comportamento em função de uma lombada, que deveria diminuir a velocidade para não atropelar uma pessoa (que era o objetivo original do objeto). Porém o motorista desloca seu significado para “vá devagar para não quebrar a suspensão do seu carro”. A ação “diminua a velocidade” está articulado com o concreto, ou delegado ao concreto. Além do descolamento do significado (redução pela força do concreto e não pela moralidade da lei de limite de velocidade) há a alteração da própria substância. A lombada tem o “papel” de atuar como um guarda de trânsito, que não está ali. Foi feita por um engenheiro que não está ali. Ela ocupa aquele espaço e tempo – está presente dia e noite. É uma coisa que age confiantemente como legado (Latour, 2001, p. 216). Em última análise, o quebramolas não é feito apenas de matéria: está repleto de engenheiros e legisladores que misturam suas vontades e perfis históricos aos do cascalho, concreto, tinta e cálculos matemáticos. A translação

técnica reside no ponto cego onde sociedade e matéria trocam propriedades (Latour, 2001, p. 218).

Sismondo (2010) exemplifica o resultado de pesquisa de Pasteur no laboratório, em como essa estrutura se torna primordial para os resultados que posteriormente lhe conferiram influência em relação às medidas de saúde pública, prestígio e poder, passando pelo aliciamento do levedo (elemento não humano) na construção de uma aliança com ele. O laboratório representa para o cientista o que o armamento representa para um exército.

### **3.1.1 Laboratórios**

Como teoria materialista, a TAR pressupõe que a Ciência e Tecnologia - C&T - traduzem matéria e ações em uma e em outra. Uma forma de entender é o exemplo dos laboratórios, em que a natureza é manipulada pelas suas ferramentas. As ferramentas nos laboratórios são o que conferem o poder aos cientistas de atuar sobre a natureza pelas funcionalidades de manipulação que estas ferramentas proporcionam: aumento de escala, separação de partes de objetos e entidades não humanas, sujeição a testes e controles para que as alianças sejam estabelecidas, posteriormente sendo transformadas em documentos escritos como relatórios e artigos. Na TAR, o poder da C&T reside no arranjo de atores e no poder dos laboratórios e de suas funcionalidades, que estão atreladas a observações, a formulações e manipulações. Entretanto, para as Ciências Sociais, o texto científico é o equivalente funcional do laboratório: “[...] (o texto) é o local dos testes, experimentos e simulações. Dependendo do que se passa nele, há ou não há um ator, há ou não há uma rede sendo traçada. E isso depende inteiramente da maneira precisa como ele é escrito – e cada novo tópico exige uma nova maneira de ser tratado por um texto” (Latour, 2006, p.347).

O trabalho da tecnociência está em compreender os interesses de variados atores, traduzi-los para que todos trabalhem em conjunto e em acordo (Sismondo, 2010, p.90), de forma simétrica, em um processo denominado sociologia da tradução (Callon, 1986). A noção de tradução compreende o processo de contínua mudança de objetivos e interesses que ocorrem nos diversos estágios das relações entre natureza e sociedade (Callon, 1986, p. 15-18). Callon (1986) faz uma crítica à forma com os sociólogos analisam o conteúdo da C&T e da sociedade, com parcialidades e entre processos naturais e processos sociais. É este tratamento imparcial aos atores humanos e não-humanos, que leva a outra ideia importante da TAR: a da simetria generalizada.

### **3.2 Simetria generalizada**

Na TAR, há a ideia de uma teoria geral das relações entre elementos heterogêneos. Em sua proposta de análise de redes heterogêneas (ou Engenharia heterogênea), Law (1993) a partir do exemplo das navegações portuguesas, destaca a simetria generalizada, atribuindo aos elementos humanos ou não o mesmo tipo de análise, e a definição recíproca em que atores são as entidades que exercem influência sobre outros elementos (Law, 1993, p. 132). Correntes marítimas e ventos contrários contribuíram para a expansão portuguesa além de caravelas e marinheiros (Law, 1993, p. 131).

A simetria generalizada consiste em não categorizar os elementos heterogêneos que compõem a

rede, sejam processos, entidades, objetos, pessoas, instituições ou regras ou máquinas, força naturais ou grupos sociais. É necessário estar a lidar com a heterogeneidade em toda a sua complexidade (Law, 1993, p. 117). Humanos e não-humanos (objetos e entes naturais) devem ser integrados no mesmo plano, sem privilégios para um ou outro grupo, já que ambos são interdependentes (Latour, 1986). Para Law (1993), é a função desses componentes associadas que conferem durabilidade a rede e a torna menos suscetível a dissociação (Law, 1993, p. 130). É como dar voz a esses elementos para que eles façam suas próprias explicações. Em última instância, a ideia de simetria generalizada é de que a que a sociedade e a natureza seriam coproduzidas (Latour, 1992, p. 287).

#### **4 OS DADOS DA PESQUISA SOB A PERSPECTIVA DA TEORIA ATOR-REDE**

Tanto a ciência como a tecnologia (C&T) são forças poderosas na sociedade moderna industrializada, entretanto, a forma como sua força é criada e desenvolvida é obscura (Callon; Law; Rip, 1986, p.3). Para os autores, “a ideia que há um método científico especial, um domínio onde a verdade prospera em detrimento do poder, é um mito” (Callon; Law; Rip, 1986, p.4). Ao buscar uma contribuição para o estudo das dinâmicas da ciência, Callon, Law e Rip (1986, p. 3) defendem que “um entendimento da dinâmica da ciência somente é possível quando a força da ciência no dia a dia da sociedade é considerada”.

Para Feenberg (2010, p. 69), a C&T exerce uma grande influência na vida diária das pessoas, porém, seu desenvolvimento ainda está sob o controle dos senhores dos sistemas técnicos, obscurecendo a própria democracia política. Feenberg (2010) argumenta que a tecnologia moderna, se administrada em outro contexto social, poderia ser um instrumento de democratização. Esse entendimento correto da mudança social e científica requer o abandono da dicotomia entre ciência e política, sendo que a compreensão dessas transformações pressupõe que os atores sejam seguidos, não somente nos palácios (quando fazem política), mas também nos laboratórios (quando fazem ciência) (Callon; Law; Rip, 1986, p. 4).

De acordo com Grigg (2015), alguns princípios que estão atrelados à abertura dos dados da pesquisa e corroboram com estas perspectivas de democratização da ciência são: a publicidade, conferindo maior visibilidade aos pesquisadores; a possibilidade de reuso de dados em novas conexões; a indução da colaboração em rede e a transparência dos dados utilizados para os resultados; a aceleração da produção de novas pesquisas; o atendimento às regras de financiadoras de pesquisa; aperfeiçoar os investimentos em pesquisa e reduzir a duplicação; promover a reprodutividade, a verificabilidade para garantir a boa prática científica; o acesso à pesquisa de importância social e maior consciência dos desafios da sociedade.

Para viabilizar este acesso aberto à produção científica, duas mudanças contribuíram com a democratização do acesso aos textos produzidos nas pesquisas. A primeira diz respeito ao movimento pelo *software* livre de código fonte aberto, que viabilizou a infraestrutura tecnológica necessária ao livre acesso. Iniciado ainda na década de 1970, após diversos desdobramentos do movimento do *software* livre foi o que favoreceu o desenvolvimento das ferramentas para publicação e disponibilização da produção científica de forma aberta e gratuita (European Commission, 2006). A segunda, mais recentemente, refere-se às mudanças políticas das agências de fomento que tem promovido um ambiente social favorável à democratização da ciência, na medida em que reduz ou anula os custos de acesso à pesquisa financiada com recursos públicos

e/ou privados. Segundo o Open Science and Research Initiative, a principal premissa da pesquisa aberta é a democratização da ciência (Open Science And Research Initiative, 2014, p.2).

A perspectiva da ciência aberta transcende então o acesso aos resultados de pesquisas. Além do acesso aberto às publicações, os dados de investigação abertos, o *software* de fonte aberta, a colaboração aberta, a revisão por pares aberta, as anotações de pesquisa abertos, os recursos educacionais abertos, as monografias, as dissertações e teses abertas, a ciência do cidadão e o *crowdfunding* (financiamento coletivo) de pesquisa também estão nas fronteiras da ciência aberta (Bueno De La Fuente, 2016). Diante dessa concepção mais ampla, a ciência aberta pode ser considerada um importante mecanismo para a viabilidade da pesquisa científica contemporânea, que utiliza intensivamente dados digitais de pesquisa (Sayão; Sales, 2016). A prática da ciência é realizada de tal forma que outros possam colaborar e contribuir, sempre que os dados da pesquisa, notas de laboratório e outros processos de pesquisa estejam disponíveis gratuitamente e de forma facilitada, em termos que permitam a reutilização, a redistribuição e reprodução da pesquisa e seus dados. Com isso, a gestão de dados se torna relevância no contexto das instituições de investigação científica e financiadores da ciência.

Latour (2000) ao abordar o processo de construção científica, denuncia a vulgarização da ciência perante a maioria de leigos na sociedade, mantendo a atividade científica em uma caixa preta em que apenas os estudos sociais de Ciência, Tecnologia e Sociedade (CTS), vêm tentando abrir (Latour, 2000, p. 33-34). “A caixa preta contém aquilo que já não precisa ser repensado, coisas cujo conteúdo tornou-se uma questão de indiferença” (Latour; Callon, 1981, p. 284). A proposta metodológica de Latour para o estudo da construção dos fatos se dá, de forma simples, como ele denomina, a partir do melhor de todos os guias: seguir os próprios cientistas. Ao invés de analisar uma caixa preta fechada, ou o resultado final, como fatos científicos ou artefatos técnicos, ele prefere seguir os passos dos cientistas no momento e nos lugares em que os seus elementos constituintes ainda estão instáveis, sendo construídos.

O processo de investigação que gera os dados é muito mais difícil de ser replicado, já que envolve diversos recursos (Latour, 2000, p. 115). Latour (2000) descreve as implicações e as circunstâncias da construção de fatos científicos e de artefatos ao ponto de se tornarem caixas pretas, ou seja, o momento e que esses recursos científicos estão acabados, em que toda a sua complexidade está encapsulada em um conceito ou em um artefato de forma que se torna inquestionavelmente aceito. As caixas-pretas são o último estágio da construção científica. Quaisquer questionamentos ao objeto estabilizado custam muito, discordar e argumentar, abrir a caixa-preta é, para Latour, uma tarefa cara, quiçá inexecutável (Latour, 2000, p. 116). Esses questionamentos exigem perfazer todo o caminho, desde o acesso aos textos, aos pesquisadores, aos institutos de pesquisa que representam os governos e políticas públicas em que estes institutos são regulados, também aos laboratórios, às matérias primas, às pesquisas de campo, aos dados levantados, aos ambientes naturais ou sociais de coleta desses dados, às habilidades dos cientistas na condução dos seus experimentos.

Grosso modo, os resultados da pesquisa, publicados em forma de artigos científicos, teses e dissertações, podem ser comparados a uma caixa preta. São consolidados a partir de teorias, dados e métodos. Uma forma de abrir essas caixas-pretas seria a partir dos dados de pesquisa disponíveis, gerados pelos trabalhos dos cientistas antes que se tornem apenas os resultados publicados de suas análises. Diante da disponibilidade, refazer a análise dos dados por outros enfoques pode maximizar os recursos empregados na coleta desses dados, seja por um nova análise dos mesmos ou por ratificar estudos já elaborados. Pode ser feita a verificação da caixa-

preta a um custo aceitável. A importância da disponibilidade dos dados da pesquisa se dá principalmente pela possibilidade de compreender o processo de geração daquele conhecimento, sem obscuridades.

Para Latour (2000), uma quantidade ínfima de leitores faria a reexecução do que foi feito pelos autores, pois é uma opção cara e inviabilizada pela falta de acesso aos recursos que estão fora dos textos: os dados da pesquisa, ou “recursos que provêm de instrumentos, animais, figuras [...], do lugar de onde saíram: a natureza ou do laboratório” (Latour, 2000, p. 101). Refazer esse percurso implicaria na disponibilidade de máquinas e laboratórios, teorias e métodos utilizados. Nessa perspectiva, a disponibilidade de dados da pesquisa ampliaria o reuso dos mesmos não apenas para validações, mas para novas inferências a partir de abordagens distintas, tornando o processo de construção científica transparente, aberto e democrático.

Não é intenção aqui reduzir as atividades científicas e tecnológicas à transmissão da inscrição literária, no entanto, consideramos que as inscrições são uma característica central da Ciência e Tecnologia (C&T), pois o texto revela a estratégia de construção de mundo dos seus autores (Callon; Law; Rip, 1986, p. 10). A comunicação da produção científica está centrada na divulgação dos resultados das pesquisas. É a partir da descrição dos atores que é possível identificarem a constituição e a dinâmica das redes.

## **5. RESULTADOS: DESCRIÇÕES DOS ATORES E A DINÂMICA DA REDE**

A TAR não parte de uma rede pronta para ser descrita, mas de uma rede que se constrói a partir da descrição de atores que exercem ação (Latour, 2006, p.340). A TAR não é, portanto, uma configuração fechada pré-estruturada de uma rede, antes é uma forma de descrever a agência de atores que constitui uma rede. As estruturas são engendradas pelos movimentos e relações entre os atores, e não os atores se encaixam em um sistema previamente estruturado. A TAR também não considera uma explicação dos dados, não é um quadro explicativo (Latour, 2006, p.343-344). Sob estas diretrizes, buscou-se evidenciar o papel dos dados de pesquisa no ciclo de geração do conhecimento científico. Os dados de pesquisa constituem o elemento central que subsidia os resultados da pesquisa, pois são os dados coletados e utilizados para análise que serão posteriormente tratados e condensados em forma de resultados divulgados nos textos científicos. Se as inscrições são os alicerces dos textos científicos, como afirma Latour (2000), logo, os dados da pesquisa fazem parte da sustentação dos resultados da própria pesquisa.

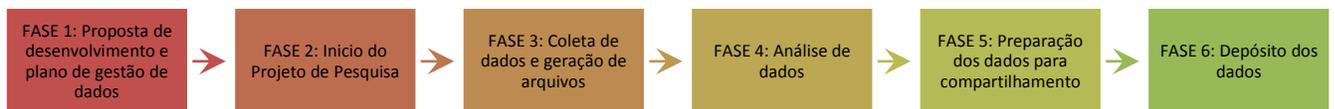
Desta forma, sob a perspectiva da TAR (Latour, 2001), em que elementos humanos e não-humanos tecem uma determinada rede, pode-se inferir que os pesquisadores, as agências de fomento, as instituições de ensino, os laboratórios, os bits, os arquivos eletrônicos em quaisquer formatos, os sistemas de armazenamento da produção acadêmica, as bibliotecas acadêmicas, os mecanismos de buscas online para recuperação de informação e pesquisas, além dos próprios dados de pesquisa em si, constituem, dentre outros elementos, a rede de processo de investigação e geração de conhecimento científico.

Considerando este escopo, na figura 2 são identificados alguns atores da rede de geração de conhecimento científico e como a disponibilidade dos dados altera a dinâmica desta rede. Os dados produzidos como parte das pesquisas tem uma ampla variedade de formatos, desde estatísticas e resultados experimentais até registros e transcrições de entrevistas (Borgman, 2012). Para que a gestão desses dados seja executada, diversos modelos de ciclo de vida dos dados têm

sido propostos. Os modelos analisados nesta pesquisa foram propostos por instituições proeminentes que advogam pelo acesso aos dados dentro da própria comunidade científica, se tornando referência para pesquisadores e gestores de dados (BALL, 2012), dentre eles, destacam-se: Digital Curation Centre (DCC) Curation Lifecycle Model, Dataone Data Lifecycle, DDI Combined Lifecycle Model, Uk Data Archive Data Lifecycle (Ball, 2012). Para este estudo selecionou-se o modelo de ciclo de vida ‘Data Life Cycle’ do Inter-university Consortium for Political and Social Research (ICPSR), que enfatiza as melhores práticas da gestão de dados por meio do ciclo de vida dos dados, trabalhando muito próximo aos pesquisadores que submetem seus conjuntos de dados para o uso da comunidade de pesquisa em Ciências Sociais (Inter-university Consortium for Political and Social Research, 2012, p.5).

O arquivamento de dados é um processo que acompanha todo o ciclo de vida do estudo que estará sendo realizado, devendo ser um elemento incorporado à própria metodologia da pesquisa (Inter-university Consortium for Political and Social Research, 2012, p.8). A figura 1 apresenta as etapas do ciclo de vida proposto pela ICPSR.

*Figura 1: Etapas da gestão dos dados propostas pela Inter-university Consortium for Political and Social Research (ICPSR)*



Fonte: Adaptado de ICPSR (2012, p.9).

A sequência da gestão dos dados que constam na figura 1 representam as etapas que o pesquisador precisará executar para que os seus dados de pesquisa sejam disponibilizados. Cada uma dessas etapas carrega inúmeros desdobramentos como regras de conformidade, práticas de gestão, protocolos, formatos e requisitos legais para que os dados estejam disponíveis não apenas de forma aberta, mas também com segurança, com qualidade atestada e preservado frente à obsolescência tecnológica. Desta forma, os dados de pesquisa deixam de figurar como subprodutos das atividades de pesquisa e se tornam um foco de grande interesse para todo o mundo científico (Sayão; Sales, 2016). Fazem parte dos elementos não humanos.

### **5.1 A rede de comunicação científica a partir da disponibilidade dos dados da pesquisa**

Os modelos tradicionais de comunicação científica abrangiam a disponibilidade dos resultados de pesquisas, como artigos, teses e dissertações, após o pesquisador percorrer todos os processos envolvidos desde a pesquisa para elaboração do documento, fases de avaliação por pares, publicação, tratamento, disseminação, acesso até a assimilação e o uso desses resultados para criação de novas pesquisas e, conseqüentemente, novos documentos (Sayão; Sales, 2012, p.128). Ao seguir os atores dos modelos tradicionais de comunicação científica, os pesquisadores, as universidades, os pares da comunidade científica, as agências de financiamento, as revistas e os editores científicos, seus sistemas de busca *online* e a própria sociedade, é possível ver como exercem agência na dinâmica da rede de comunicação científica.

A partir do movimento pelo acesso aberto, às políticas de acesso, os repositórios institucionais alteraram a forma e principalmente o custo com que os resultados de pesquisa são disponibilizados (Rodrigues, 2008 apud Sayão; Sales, 2012, p.128-129). Com a constante intensidade em geração e uso dos dados, os repositórios de dados de pesquisa e as políticas de gestão dos dados tem dado o suporte para a que os arquivos digitais de dados tenham um papel relevante, alterando a própria forma de construção da ciência. A figura a seguir apresenta o processo de comunicação científica e as alterações provocadas no processo a partir da presença de novos elementos na rede, como o acesso aberto aos resultados de pesquisa e a disponibilidade dos dados em repositórios apropriados (Figura 2).

Nota-se que o pesquisador é quem absorverá mais uma atividade, que é a publicação dos dados da pesquisa, que antes ficam subutilizados no modelo tradicional. Esta atividade compreende todas as etapas do ciclo de vida dos dados, conforme a figura 1. Entretanto, as implicações desta mudança são diretamente benéficas à própria atividade de pesquisa. Além dos sistemas tradicionais de buscas, que requerem pagamento para acesso, os recursos abertos, viabilizados pela retroalimentação da atividade científica, ficam disponíveis ao pesquisador (área com fundo amarelo), ampliando seu escopo de fontes de informação de forma gratuita. O acesso aos repositórios institucionais que trazem os resultados de outras pesquisas e aos repositórios de dados de outros pesquisadores permite a democratização da produção científica e em última instância, beneficia a sociedade como um todo.

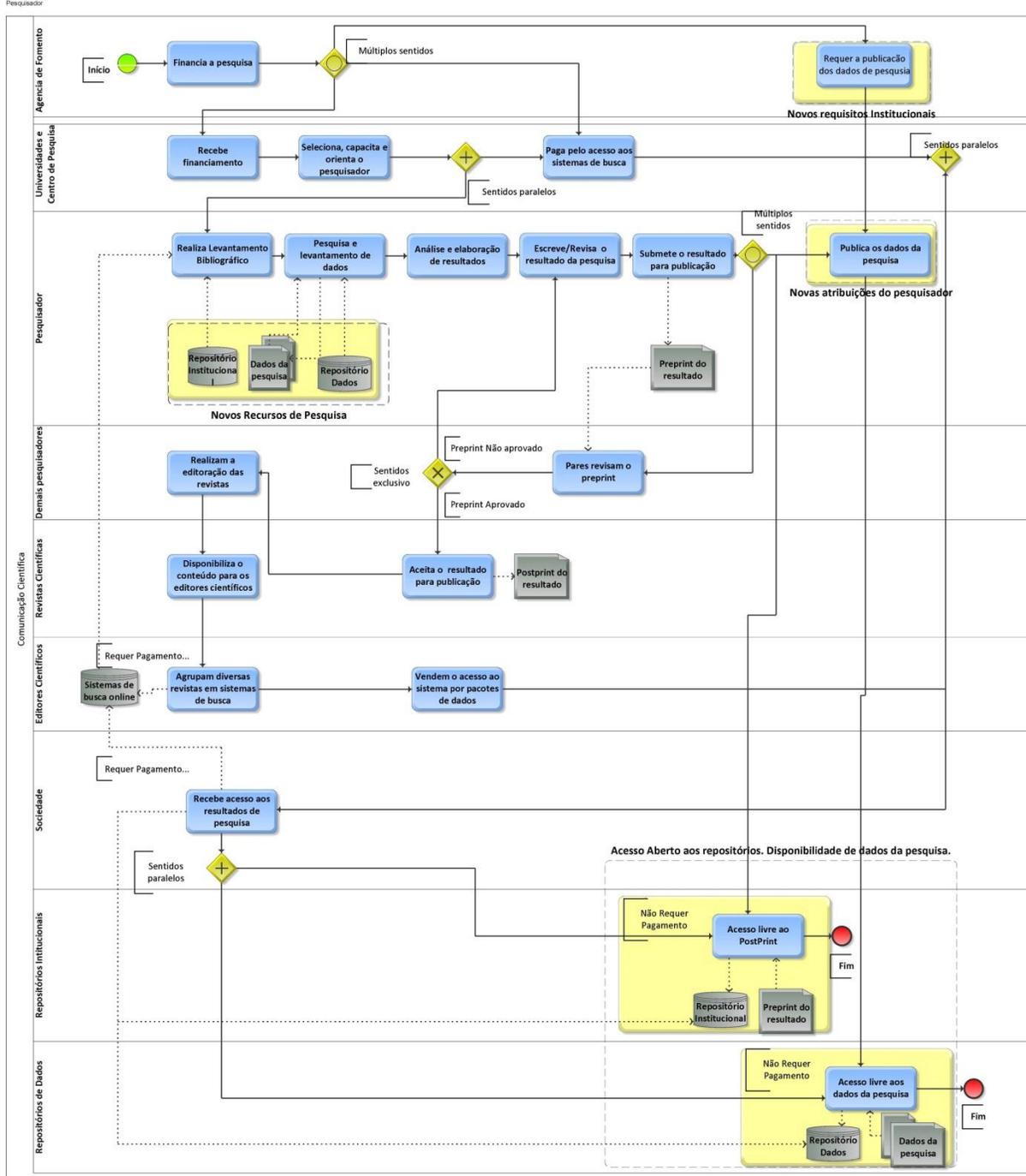
Utilizar a perspectiva da TAR, neste artigo, permite que os elementos não humanos, como os dados da pesquisa, tenham a mesma valoração dentro da rede que os elementos humanos, e para isso resgata-se aqui a noção de simetria generalizada. A simetria generalizada é o pressuposto que atribui o mesmo valor para elementos não humanos e os humanos, em uma dada rede. Considera-se, portanto, que os dados da pesquisa são um ator-rede, pois sua presença é marcada por uma série de alterações desdobradas e percebidas na rede, pela mudança que provoca na dinâmica de produção do conhecimento. Os dados, quando disponíveis, alteram a forma como os pesquisadores podem produzir análises, inferências, comparações, resultando no próprio conhecimento científico.

Um ator, na perspectiva da TAR, precisa efetivamente arregimentar outros atores para que a sua ação seja ressonante na dinâmica da rede. A disponibilidade dos dados de pesquisa requer de outros actantes as suas atuações, como por exemplo, os próprios repositórios de dados, que tem requisitos distintos dos repositórios institucionais. Sem as características inerentes aos repositórios de dados, os dados não podem ser acessíveis e, conseqüentemente, reutilizáveis. Neste sentido, a ideia de mediação técnica da TAR pode ser entendida, pois a viabilidade técnica é o que mediará o acesso aos dados. Esta condição inerente ao elemento não humano condiciona a atividade humana para um determinado objetivo, o que explicado nas palavras de Latour (2001, p.218) é que eles estão “trocando propriedades”.

Diante da proeminência dos dados de pesquisa, este artigo busca evidenciar como a disponibilidade dos dados de pesquisa interfere no processo de geração e de comunicação científica. Nota-se na figura 2 a partir dos elementos destacados que se refere à disponibilidade dos dados, a alteração da dinâmica da rede. Na medida em que o pesquisador pode valer-se de novos insumos de informação, ou seja, os conjuntos de dados de pesquisa disponibilizados por outros pesquisadores para análise, comparação, aplicação e reuso desses dados na sua própria pesquisa, usando os repositórios de dados para a busca destes conjuntos de dados, ele também precisará, posteriormente, disponibilizar os dados que produziu nos mesmos repositórios. Em um

cenário tradicional, quando apenas a pesquisa bibliográfica foi realizada, no cenário viabilizado pela ciência aberta, a possibilidade da busca por dados altera a metodologia do pesquisador.

Figura 2 – A Comunicação Científica com Acesso Aberto e disponibilidade dos dados de pesquisa



Fonte: Elaboração própria com base em Sayao e Sales (2012, p. 129-131) e Setenareski (2013).

Na medida em que os dados constituem em si próprios os insumos para análises secundárias e novas inferências, eles exercem um papel de maior relevância, atuando sobre os demais atores da rede de geração de conhecimento científico, que agora reordenam sua forma de realizar a pesquisa. A geração intensiva e disponibilidade de dados projetou um elemento não-humano que até então era coadjuvante na rede de comunicação científica: os dados da pesquisa.

## 6 CONSIDERAÇÕES FINAIS

A produção científica e a comunicação dos seus resultados é um processo complexo, heterogêneo e interdependente. Entender a dinâmica dessas relações e os desdobramentos da atuação dos atores é essencial para que se compreendam os rumos que a produção e a comunicação da ciência têm tomado. Ainda que não seja a forma exclusiva, o acesso aberto tem sido uma prática cada vez mais estimulada pelos agentes de financiamento e pela própria comunidade científica. São iniciativas da própria comunidade, como o movimento pelo *software* livre de código aberto, posteriormente o movimento pelo acesso aberto às publicações científicas que, em conjunto com as ferramentas e protocolos técnicos desenvolvidos por essas mesmas comunidades, têm sustentado a execução de políticas de acesso aberto e compartilhamento de dados de pesquisa.

Os dados têm tido uma relevância ainda maior nos últimos anos na geração do conhecimento científico. Sua produção exponencial, alcançada pela disponibilidade de tecnologias de informação e comunicação, pela internet, e por recursos de geração de grande volume de dados tem lhe conferido o poder de mudar a dinâmica da própria rede. Os dados da pesquisa em si não são um elemento novo no processo de produção do conhecimento. Entretanto, a descrição do seu papel, como forma de simplificação da complexidade, revela a sua justaposição em relação aos outros elementos, demonstrando a sua agência que modifica a estrutura da rede. Em outras palavras, os dados da pesquisa que eram produzidos e esquecidos, desorganizados e não acessíveis nos arquivos pessoais dos pesquisadores ou das instituições, na medida em que são tratados, padronizados, preservados e disponíveis de forma gratuita, constituem um novo insumo para se fazer ciência.

## REFERÊNCIAS

Ball, A. (2012), *A review of data management lifecycle models*. Bath, UK: University of Bath. Extraído de <http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf>.

Borgman, C.L. (2012), The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, v. 63, n. 6, pp. 1059–1078.

Bueno De La Fuente, G. (2016), Challenges and strategies for the success of Open Science. In: Facilitate Open Science Training For European Research (FOSTER). *Open Science at the Core of Libraries*. Extraído de <https://www.fosteropenscience.eu/node/1424>.

Callon, M. (1993), Society in the Making: The Study of Technology as a Tool for Sociological Analysis. IN: In: Pinch, T. J. ; Bijker, W. E.; Hughes, T. P. *The social Construction of Technological Systems*. Cambridge (Mass.), MIT, pp. 77-97.

Callon, M. (1986), Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. In: Law, John. *Power, action and belief: a new sociology of knowledge?* London, Routledge, pp. 1-29.

Callon, M.; Law, J.; Rip, A. (1986), How to study the force of science. In: Callon, M.; Law, J.; Rip, A. (Eds.) *Mapping the dynamics of science and technology*. Houndmills: Macmillan Press, pp. 3-15.

European Commission. (2014), *Horizon 2020 em breves palavras: O programa-quadro de investigação e inovação da UE*. 40p. Extraído de <https://goo.gl/M72BgB>.

European Commission. (2016a), H2020 Programme. *H2020 AGA – Annotated Model Grant Agreement*. Version 2.1.1. 1st July 2016. pp. 216-219. Extraído de <http://goo.gl/SXbKxE>.

European Commission. (2016b), *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020*. 15 nov. 2016. Extraído de <https://goo.gl/RgkJpg>.

European Commission. (2016c), *Open innovation, Open Science, open to the world. A vision for Europe*. Brussels: European Commission. Directorate-General for Research and Innovation. Extraído de <https://goo.gl/M72BgB>.

European Commission. (2006), *Study on the economic and technical evolution of the scientific publication markets in Europe: final report*. Bruxelles: European Commission.

Feenberg, A. (2010), Racionalização Subversiva: Tecnologia, Poder e Democracia. In: Neder, Ricardo T. *A teoria crítica de Andrew Feenberg: racionalização democrática, poder e Tecnologia*. Brasília: Observatório do Movimento pela Tecnologia Social na América Latina / CDS / UnB / Capes, pp. 67-95. Extraído de <https://www.sfu.ca/~andrewf/coletanea.pdf>.

Grigg, K. S. (2015), Data in the Sciences. In: Kellam, L.; Thompson, K. *Datalibrarianship: The Academic Data Librarian in Theory and Practice*, pp. 179-192.

Inter-University Consortium For Political And Social Research (ICPSR). (2012), *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle*. 5. ed. Ann Arbor, MI, Extraído de <http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf>.

Latour, B. (2000), *Ciência em ação: como seguir cientistas e engenheiros sociedade afora*. São Paulo: UNESP, 438p.

Latour, B. (2006), Como terminar uma tese de sociologia, in: *Cadernos de campo*, São Paulo, n.14/15, pp. 339-352. . Extraído de <http://www.bruno-latour.fr/sites/default/files/downloads/90-DIALOGUE-POR.pdf>.

Latour, B. (1998), La Tecnología es la sociedad hecha para que dure. In: Domenech, Miquel; Tirado, Francisco Javier. *Sociologia Simétrica*. Ensayos sobre Ciencia, Tecnología y Sociedad. Barcelona: Gedisa Editorial, pp. 109-142.

Latour, B. (1992), One more turn after the social turn... In: McMullin, E. *The Social Dimension of Science*. Indiana: Indiana University of Notre Dame Press, pp. 272–294.

Latour, B. (1995), Os objetos tem história? Encontro de Pasteur com Withehead num banho de ácido láctico. *História, Ciências, Saúde: Manguinhos*, v. II, n.1, pp. 7-26, mar-jun.

Latour, B. (2001), Um coletivo de humanos e não-humanos: No labirinto de Dédalo. In: \_\_\_\_\_. *A Esperança de Pandora: Ensaios Sobre a Realidade dos Estudos Científicos*. Bauru, SP : EDUSC, pp. 201-246.

Latour, B. (1993), *We have never been modern*. Cambridge, MA: Harvard University Press.

Latour, B. Callon, M. (1981), Unscrewing the big Leviathan: how actors macro-structure reality and how sociologists help them to do so. In: In: Knorr-Cetina, K. D.; Cicourel, A.V. (eds.). *Advances in social theory and methodology: toward and integration of micro- and macro-Sociologies*. Boston: Routledge; Kegan Paul, pp. 277-303.

Law, J. (1993), Technology and Heterogeneous Engineering: The Case of Portuguese Expansion. IN: In: Pinch, Trevor J.; Bijker, Wiebe E.; Hughes, Thomas P. *The social Construction of Technological Systems*. Cambridge(Mass.), MIT, pp. 105-127.

Open Research Funders Group. (2016), *Proeminent funding organizations team up to launch Open Research Funders Group*. Washignton, DC, 15 Dez. Extraído de <http://www.orfg.org/news>.

Open Science And Research Initiative. (2014), *Open Science and Research Handbook*. Extraído de <https://avointiede.fi/documents/14273/0/Open+Science+and+Research+Handbook+v.1.0/50316d5d-440b-4496-b039-2997663afff8>

Organização Para Cooperação e Desenvolvimento Econômico. (2015), Making Open Science a Reality. (*OECD Science, Technology and Industry Policy Papers*, 25). Paris: OECD Publishing. Extraído de <http://dx.doi.org/10.1787/5jrs2f963zs1-en>.

Sayao, L. F.; Sales, L. F. (2016), Algumas considerações sobre os Repositórios digitais de dados de Pesquisa. *Inf. Inf.*, Londrina, v. 21, n. 2, pp. 90-115, maio/ago.

Sayao, L. F.; Sales, L. F. (2012), O impacto da curadoria digital dos dados de pesquisa na comunicação científica. *Encontros Bibli: revista eletrônica de Biblioteconomia e Ciência da Informação*, v. 17, n. esp. 2 – III SBCC, pp. 118-135.

Sayes, E. (2014), Actor-Network theory and methodology: just what does it means to say that nonhumans have agency? *Social Studies of Science*, v. 44, n. 1, pp.134-149.

Setenareski, L. E. (2013). *Repositórios digitais abertos: um movimento do livre acesso alternativo à estrutura oligopolizada das editoras*. 2013. 113 f. Dissertação (Mestrado) - Curso de Programa de Pós-graduação em Políticas Públicas, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, Curitiba. Extraído de <http://goo.gl/UB42yK>.

Sismondo, S. (2010), Actor Network Theory. In: Sismondo, Sergio. *An Introduction to Science and Technology Studies*. Wiley Blackwell: West Sussex, UK, pp. 81-92.