

Autorização concedida ao Repositório Institucional da Universidade de Brasília (RIUnB) pela Professora Michelli Pereira da Costa, em 22 de novembro de 2018, para disponibilizar o trabalho, gratuitamente, para fins de leitura, impressão e/ou download, a título de divulgação da obra.

REFERÊNCIA

COSTA, Michelli Pereira da; LEITE, Fernando César Lima. Fatores que exercem influência na comunicação dos dados de pesquisa: uma revisão sistematizada da literatura no campo da Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais eletrônicos...** Londrina: Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI/UEL), 2018. Disponível em:
<<http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/viewFile/1265/1869>>. Acesso em: 26 nov. 2018.

XIX encontro nacional
de pesquisa em
ENANCIB ciência da informação

// SUJEITO INFORMACIONAL E AS
PERSPECTIVAS ATUAIS EM CIÊNCIA
DA INFORMAÇÃO. //

22-26
OUTUBRO
2018
LONDRINA/PR



XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018

GT-7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

FATORES QUE EXERCEM INFLUÊNCIA NA COMUNICAÇÃO DOS DADOS DE PESQUISA: UMA REVISÃO SISTEMATIZADA DA LITERATURA NO CAMPO DA CIÊNCIA DA INFORMAÇÃO

Michelli Costa (Universidade de Brasília)

Fernando César Lima Leite (Universidade de Brasília)

FACTORS THAT INFLUENCE RESEARCH DATA COMMUNICATION: A SYSTEMATIC REVIEW OF THE LITERATURE IN THE INFORMATION SCIENCE FIELD

Modalidade da Apresentação: Comunicação Oral

Resumo: A noção da comunicação de dados de pesquisa, derivada da noção de comunicação científica na Ciência da Informação, engloba os macroprocessos de produção, compartilhamento e uso de dados de pesquisa produzidos ou utilizados para o desenvolvimento de investigação. Adicionalmente, a filosofia e as iniciativas relacionadas à ciência aberta produziram novos elementos que atualizam e tornam mais complexas análises acerca do fenômeno. Os objetivos do trabalho foram identificar e analisar as discussões sobre comunicação de dados de pesquisa no âmbito da Ciência da Informação, e a partir da ótica da área, identificar fatores que exercem influência na comunicação dos dados de pesquisa. O método adotado foi o da revisão sistematizada da literatura, de 13 artigos indexados na base de dados LISA e que tinha como objeto de análise central a discussão sobre a comunicação de dados de pesquisa. Como principais resultados foram obtidas quatro dimensões predominantes, a saber, imperativos da ciência moderna, infraestrutura, financiamento e contexto social. Tais dimensões se desdobram em categorias teóricas que representam fatores intervenientes do processo de comunicação de dados de pesquisa. A discussão das categorias permitiu explicitar elementos relevantes para a compreensão das práticas de comunicação de dados de pesquisa, à luz da ciência da informação.

Palavras-Chave: Dados de pesquisa; Comunicação científica; Ciência aberta; Ciência da informação

Abstract: The research data communication concept is composed by the process of production, sharing and use of research data produced or used during the research development. It is based in the scientific communication theory from the Information Science field. Also the open science concepts and initiatives brought new elements that widened and complicated the phenomenon analyses. The study aimed to identify and analyze the research

data communication topics related to the Information Science field, and based in the area perspective to determine the key factors that influence the research data communication. To achieve it the researches applied the systematic literature review in 13 papers indexed by the LISA database and that are focused in the research data communication. The main results achieved are the definition of four prevalent dimensions: modern science imperative; infrastructure; funding and social context. These dimensions develop to theoretical categories that represent and impact the research data communication process. The categories discussion brought relevant elements to understand the research data communication practices related to the Information Science field.

Keywords: Research data; Scientific communication; Open science; Information science

1 INTRODUÇÃO

A comunicação científica é fenômeno de interesse da Ciência da Informação (CI) desde a sua gênese, como é possível perceber em artigo seminal de Borko (1969). Nesse contexto da comunicação humana destaca-se o veículo que, há pelo menos cinco séculos, tem sido considerado central para o desenvolvimento da ciência, os periódicos científicos. Segundo Meadows (1999), os periódicos surgiram no século XVII e rapidamente foram se tornando essenciais ao desenvolvimento da ciência.

Em razão dos avanços tecnológicos e das articulações internacionais em prol do acesso aberto ao conhecimento científico e da ciência aberta, experimentados no limiar do século XXI, o modelo tradicional de comunicação científica e o próprio periódico científico passaram a sofrer questionamentos. Por acesso aberto entende-se a disponibilização livre e irrestrita dos resultados de pesquisas científicas em texto completo por meio da *Internet* (LYNCH, 2003). Dentre outras abordagens, o acesso aberto tem sido apontado como elemento básico para superar limitações relacionadas tanto ao acesso à literatura científica internacional por países em desenvolvimento, quanto à disseminação do conhecimento científico de pesquisadores de localidades que estão à margem dos periódicos científicos comerciais tradicionais (CHAN *et al.*, 2005).

A perspectiva do acesso aberto tem por base uma “nova filosofia” iniciada nos anos 1990, que foi denominada por Costa (2006) filosofia aberta. Associada a essa filosofia estão as iniciativas de *software* livres e dos arquivos abertos. A ampliação da noção de abertura e de democratização do acesso trazida por esta filosofia fez emergir um novo modo de pensar, produzir e comunicar a ciência, que recebeu o nome de ciência aberta. Nela o acesso público e irrestrito não requer apenas as publicações científicas, mas representa mudança paradigmática em todo o processo de produção da ciência, desde os sistemas utilizados para

a pesquisa, em que predominam os *software* livres, até as metodologias de avaliação do alcance da pesquisa, nas quais discute-se o uso de métricas alternativas para a ciência. Neste escopo ampliado, a comunicação dos dados de pesquisa é apontada como uma das condições para avaliação dos resultados da pesquisa e para o desenvolvimento acelerado, colaborativo e eficaz da ciência como um todo.

Dados de pesquisa são os dados produzidos ou utilizados para o desenvolvimento de uma pesquisa. Os processos de produção, o compartilhamento e o uso dos dados de pesquisa serão tratados nesse estudo como comunicação dos dados de pesquisa. A apropriação da noção de comunicação substancia-se na abordagem de Garvey (1979), que concebe a comunicação científica em modelo que contempla os três macroprocessos, nomeadamente, produção, compartilhamento e uso da informação.

Apesar de a discussão sobre a comunicação dos dados de pesquisa ter obtido destaque na perspectiva da ciência aberta e do acesso aberto, ela os precede. O tema dos dados de pesquisa no âmbito da comunicação científica para a CI retoma uma discussão iniciada nos anos 1970 e que ainda não foi sistematicamente desenvolvida. Apesar do tópico ter sido introduzido na área há mais de quatro décadas, ainda são poucos os estudos que tratam dos dados de pesquisa no âmbito da comunicação científica, conforme constatou esse estudo¹. Este trabalho possui dois objetivos. O primeiro objetivo é identificar e analisar discussões sobre comunicação de dados de pesquisa no âmbito da CI. O segundo é identificar, a partir da ótica da área, fatores que exercem influência na comunicação dos dados de pesquisa.

2 PROCEDIMENTOS METODOLÓGICOS

A pesquisa buscou identificar as discussões existentes na CI sobre a comunicação dos dados de pesquisa. Para tanto, foi selecionada uma base de dados de ampla representatividade na CI, a *Library and Information Science Abstracts* (LISA). Segundo sua descrição, a LISA indexa mais de 440 periódicos de 68 países e 20 línguas diferentes no campo da biblioteconomia e da CI (LISA, 2016).

A busca foi realizada em julho e agosto de 2016. Por meio do argumento de busca [(ti(“research data”) OR ti(“scientific data”)) AND sharing], chegou-se a um conjunto de 15 documentos. O critério utilizado para a busca privilegiou os documentos que tratavam do tema do compartilhamento dos dados de pesquisa como assunto principal e que refletiam esse

¹ O tema foi explorado no capítulo 4 da tese de doutorado de Costa (2017).

tratamento no título do artigo. Complementarmente foi considerada a variação que o termo “dados de pesquisa” apresenta na literatura em língua inglesa, a saber “research data” ou “scientific data”. Não houve qualquer restrição temporal na busca realizada.

Os quinze documentos selecionados tiveram seus resumos analisados de forma a certificar a pertinência destes para o estudo proposto. A verificação sucedeu na exclusão de dois documentos, um por representar uma obra monográfica de natureza didática sobre o tema e o outro por se tratar de título. Portanto, Treze documentos resultaram da estratégia apresentada.

Para a análise dos 13 artigos, recorreu-se ao método de Revisão Sistematizada da Literatura (RSL). Este tipo de revisão de literatura, segundo Grant e Booth (2009), busca sintetizar evidências de outras pesquisas, podendo incluir ou não pesquisa abrangente e avaliação de qualidade. A síntese de seus resultados é tipicamente em formato de narrativa com acompanhamento tabular.

Para cada artigo selecionado, elaborou-se uma ficha de análise. O instrumento permitiu sistematizar aspectos sobre a publicação (tal como sua origem e sua data), informações sobre a autoria do artigo (sua origem e filiação institucional), aspectos sobre o conteúdo e sobre o objeto da pesquisa. A investigação sobre o conteúdo dos artigos teve como foco o objetivo do trabalho, temas abordados e bases teóricas. Adicionalmente, foram examinadas as metodologias utilizadas pelos estudos. Finalmente, investigou-se o tratamento acerca dos fatores que influenciam a comunicação dos dados de pesquisa e as iniciativas apontadas como facilitadoras do processo.

A análise sistematizada dos artigos permitiu a identificação de um conjunto de categorias teóricas indicativas dos fatores que influenciam a comunicação aberta dos dados de pesquisa nas discussões da CI. Tais categorias tem por objetivo descobrir conceitos e relação nos dados de forma a organiza-los em um esquema exploratório teórico, conforme descreveu Strauss e Corbin (2008) sobre o processo de pesquisa qualitativa com uso de categorias teóricas.

3 DISCUSSÃO E ANÁLISE

A revisão sistematizada da literatura permitiu a identificação de duas dimensões relevantes para a pesquisa. A primeira dimensão - a caracterização das discussões sobre a comunicação dos dados de pesquisa dentro do campo da CI - é relevante para retratar o tratamento do tema no campo de pesquisa da área. Conforme perspectiva de Borko (1969), a CI é a área do conhecimento que foi criada com o objetivo de estudar os processos relacionados

à comunicação científica de modo a aprimorá-los. Com isso, a CI é uma área do conhecimento que tem por compromisso, entre outros, a discussão de temas relevantes para a comunicação científica, onde a comunicação dos dados de pesquisa representa uma de suas formas contemporâneas. A segunda dimensão diz respeito aos fatores influenciadores na comunicação dos dados de pesquisa.

3.1 Caracterização da literatura da CI sobre a comunicação dos dados de pesquisa

As discussões acerca da comunicação de dados de pesquisa iniciaram-se na década de 70 com a publicação do Modelo United Nations International Scientific Information System (UNISIST). Nele foi previsto que a interação entre produtores e usuários de informação científica seria promovida por diferentes canais, no qual um deles seriam os canais tabulares, responsáveis ao que neste artigo é denominado de comunicação de dados de pesquisa (UNESCO; UNIONS, 1971). Embora o tema tenha surgido precocemente na CI em comparação com outras disciplinas, ele não se manteve com relevância entre as preocupações da área ao longo dos anos. A temática somente voltou a receber destaque a partir das iniciativas em prol da ciência aberta, no início dos anos 2000.

A tendência ascendente de aparição da temática a partir dos anos 2000 é percebida no conjunto de documentos analisados. Os dois documentos mais antigos datam de 2004 e a maior parte do conjunto de documentos concentram-se entre 2014 e 2015. Observa-se também que a discussão acerca da comunicação dos dados de pesquisa na CI não apresenta uma distribuição geográfica uniforme. As publicações são exclusivas de instituições editoras do norte do mundo, com ampla concentração no Reino Unido (54%) e nos Estados Unidos (31%). A centralização das publicações nas duas regiões deve-se à predominância de duas editoras, a Wiley e a Taylor & Francis, originárias dos Estados Unidos e Reino Unido, respectivamente.

Os dados sobre a concentração das publicações chamam atenção principalmente devido à natureza das editoras que predominam como publicadoras do tema. As editoras Wiley e Taylor & Francis estão entre as cinco editoras acadêmicas comerciais que mais lucram com a comercialização da comunicação científica no mundo, conforme denunciou Morrison (2012).

O antagonismo dos interesses postos demonstram a profundidade da problemática acerca da comunicação científica e da ciência aberta. Editoras comerciais auferem até mesmo de trabalhos que propõem ou defendem ações que visam conferir transparência, abertura e ampla possibilidade de acesso aos frutos da comunicação científica. Não se pretende aqui,

entretanto, elaborar uma crítica acerca dos fatores e dos interesses que imbricam tal relação. Apenas evidencia-se a predominância de editoras comerciais na discussão e que elas, aparentemente, representam as forças contraditórias das filosofias abertas.

O domínio dos países do norte do mundo em termos de procedência das publicações repetiu-se também entre a origem dos autores. Em cada um dos artigos selecionados buscou-se identificar o lugar de origem da instituição do primeiro autor. Como resultado constatou-se que eles são predominantemente dos Estados Unidos (38%) e do Reino Unido (46%). Entre outros elementos, o resultado aponta para a influência do contexto das duas regiões para a discussão do tema.

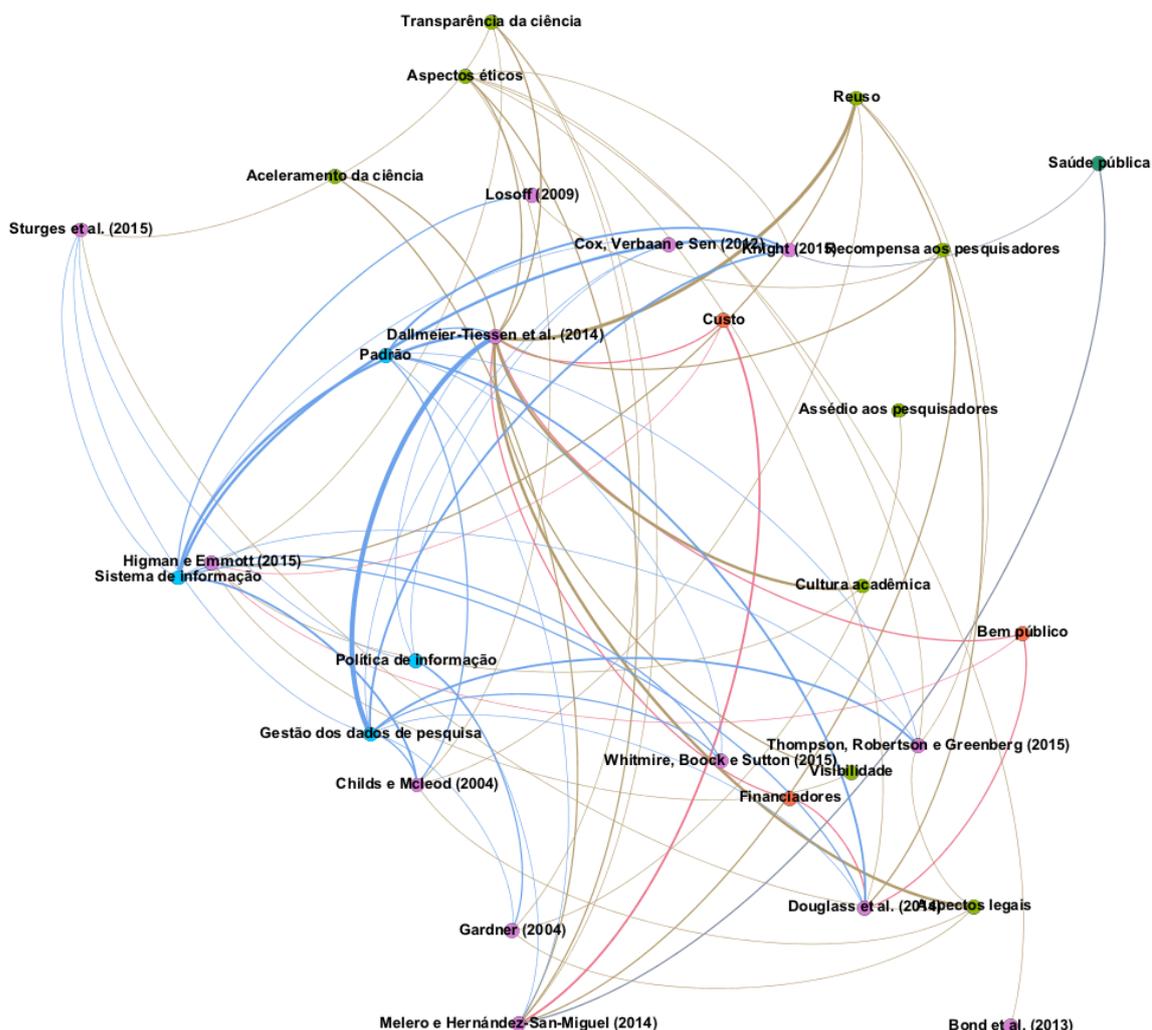
Além da procedência das publicações e de seus autores levantou-se, de forma sistematizada, informações sobre a metodologia dos estudos selecionados. Majoritariamente, as pesquisas são de natureza qualitativa. Embora tenha-se identificado pesquisas com técnicas relacionadas a abordagem quantitativa (23%), todas as análises foram fundamentadas em princípios qualitativos, mesmo nos estudos de natureza mista.

No que concerne às técnicas para a coleta dos dados adotadas nos artigos que fizeram parte do estudo, observou-se uma distribuição uniforme. As principais técnicas identificadas foram pesquisas bibliográficas (31%), entrevistas (23%), questionários (23%) e pesquisas documentais (23%). Ressalta-se que nenhum dos artigos classificados na categoria pesquisas bibliográficas se declaram como dessa natureza. No entanto, eles foram assim categorizados por não apresentarem dados que amparassem suas discussões e resultados, senão a sustentação teórica de revisões bibliográficas.

3.2. Fatores que exercem influência na comunicação dos dados de pesquisa

A revisão sistematizada da literatura permitiu a identificação de 17 fatores que influenciam os processos de comunicação dos dados de pesquisa. O resultado foi obtido por meio da execução de três passos. O primeiro consistiu na rotulagem livre dos elementos. Todos elementos citados foram tratados como fator interveniente, independente de sua ocorrência. No segundo passo, todos os elementos foram traduzidos para termos semelhantes e agrupados. Por fim, foi elaborada uma rede de relações entre os fatores identificados e os autores dos textos, a partir do software Gephi (Figura 1).

Figura 1: Categorias preliminares e dimensões de análise



Fonte: Elaboração própria a partir do software Gephi

Conforme explica Jacomy et al. (2014), a figura gerada é resultante de uma metodologia de análise de redes. Tal rede representa a relação dos autores e dos fatores intervenientes citadas por eles. Cada artigo, representado por suas autorias, e cada fator representam um nó na rede. A relação entre eles é expressada pelas linhas, dotadas de força. Portanto, nessa representação, quanto mais grossa a linha de ligação entre os nós, mais relevante é a relação evidenciada.

A Figura 1 aponta a centralidade da pesquisa de Dallmeier-Tiessen et al. (2014) e a relevância de alguns fatores entre o conjunto de textos analisados. Em especial, destacam-se as

categorias sistemas de informação, padrões, gestão de dados de pesquisa e reuso como as mais recorrentes. Ademais foi possível organizar os fatores a partir de dimensões de análise. A organização se deu a partir de aproximações conceituais entre os fatores e resultou na proposição de quatro dimensões: imperativos da ciência moderna, infraestrutura, financiamento e contexto social. A partir das dimensões de análise, serão discutidos os fatores intervenientes e suas propriedades.

3.2.1 Fatores relacionados aos imperativos da ciência moderna

O fenômeno investigado nesse estudo relaciona-se com as iniciativas da ciência aberta. Os princípios que sustentam tal proposta estão fundamentados nas mesmas ideias que foram base para as revoluções científicas e que caracterizam a ciência moderna, conforme explicou Boulton (2013). Na análise resultante da revisão sistematizada da literatura foram identificadas nove categorias que se relacionam conceitualmente com as prerrogativas da ciência moderna e serão apresentadas a seguir.

3.2.1.1 Aceleração da ciência

A construção histórica do conhecimento é uma das características da ciência moderna, destacadas pelo filósofo Comte-Sponville (2011). Ela pressupõe que um novo conhecimento se baseia em conhecimentos já registrados e comunicados. Com isso, a comunicação dos conhecimentos é entendida como um elemento que acelera o desenvolvimento da ciência, permitindo assim que novos resultados sejam descobertos de forma mais rápida e eficiente. A noção da comunicação dos dados de pesquisa como um fator que promove o desenvolvimento científico com maior rapidez foi discutido por Melero; Hernández-San-Miguel (2014) e Dallmeier-Tiessen et al. (2014). De acordo com os autores, iniciativas que favorecem a ampla divulgação dos dados de pesquisa promovem a maximização do uso dos dados para a produção de novos conhecimentos de forma otimizada e acelerada.

3.2.1.2 Reuso dos dados de pesquisa

O aceleração da ciência é possibilitado por meio do reuso dos dados, aspecto defendido por cinco artigos do conjunto analisado. No entanto, além dos benefícios relacionados ao favorecimento da acumulação de conhecimentos e construção histórica da ciência, os autores destacaram aspectos que inviabilizam ou dificultam o processo de comunicação dos dados de pesquisa e que estão diretamente relacionados ao reuso dos dados.

Dallmeier-Tiessen et al. (2014), Thompson, Robertson e Greenberg (2015) e Douglass et al. (2014) enfatizaram os custos relacionados aos processos para a comunicação de dados. De acordo com os autores, a preparação dos dados para sua publicação demanda mais tempo e mais esforços na produção dos dados. Isso porque os dados de pesquisa que são divulgados precisam estar estruturados em alguns padrões e com maior nível de detalhamento do que a pesquisa que os produziu demandaria. Além do esforço adicional para a preparação dos dados, Douglass et al. (2014) e Higman e Emmott (2015) pontuam que, em muitos casos, dados produzidos por uma pesquisa não terão aplicabilidade para outros contextos.

Além dos questionamentos quanto à utilidade dos dados para outras pesquisas, foram destacados receios dos pesquisadores com tal processo. Childs e Mcleod (2004) consideraram que pesquisadores se apresentam temerosos com respeito aos usos que podem ser feitos de seus dados. Apesar de os autores não especificarem a origem do temor dos pesquisadores, Dallmeier-Tiessen et al. (2014) apresentam pistas sobre a questão. Estes últimos afirmaram que o receio dos pesquisadores com relação aos possíveis reusos de seus dados relacionam-se com duas possibilidades. A primeira delas refere-se à desqualificação dos seus dados e, por consequência, dos resultados obtidos em sua pesquisa. A segunda diz respeito às possíveis divergências de análises resultantes dos dados. Com isso, dados comunicados abertamente seriam utilizados para desqualificar ou evidenciar contradições dos resultados apresentados pelos pesquisadores que os produziram.

Em todos os casos, os argumentos estão relacionados a uma das características essenciais da ciência moderna, conforme fora explicitado por Comte-Sponville (2011). Tal característica diz respeito a refutabilidade do conhecimento científico, que demanda transparência nos processos de produção do conhecimento científico.

3.2.1.3 Transparência da ciência

Cinco artigos do conjunto analisado argumentaram que a comunicação dos dados de pesquisa favorece, em alguma medida, o exercício da refutabilidade, característica da ciência moderna amplamente discutida por teóricos como Karl Popper (2013) e Thomas Kuhn (2005). Nesse sentido, Dallmeier-Tiessen et al. (2014) defenderam que a divulgação dos dados é importante para a validação dos resultados apresentados em um artigo ou outra forma de comunicação científica. Além disso, os autores destacaram que esses dados são úteis para testar teorias propostas no futuro e embasar novas teorias. Portanto, os autores consideram que a

comunicação dos dados de pesquisa é necessária para conferir integridade à ciência, uma vez que erros são mais facilmente detectados e a ciência é produzida com maior transparência.

O argumento do favorecimento da transparência da ciência por meio da comunicação dos dados de pesquisa também foi identificado nas discussões de Sturges et al. (2015) e Higman e Emmott (2015). Similarmente, Melero; Hernández-San-Miguel (2014) apontam que processos de comunicação de dados tem condições de proporcionar mais qualidade para a ciência produzida, uma vez que a possibilidade de diálogo com outras pesquisas é ampliada, bem como a capacidade de identificar erros e corrigi-los. Por fim, as considerações de Gardner (2004) sugerem que a ampla divulgação dos dados permite uma maior compreensão da pesquisa e que, portanto, confere maior transparência.

3.2.1.4 Assédio aos pesquisadores

No limiar da discussão entre a promoção da transparência da produção do conhecimento científico e os receios dos pesquisadores sobre os propósitos do reuso de seus dados, equilibra-se o aspecto apresentado por Gardner (2004). De acordo com o autor, a legislação estadunidense que determina que todos os pesquisadores devem divulgar seus dados de pesquisa sempre que for solicitado é um fator que atrasa as pesquisas por deixar pesquisadores a mercê dessas demandas. Além disso, Gardner (2004) denuncia que tais solicitações têm representado um instrumento para assediar pesquisadores que produzem conhecimento que contrariam o interesse de grandes corporações comerciais, como, por exemplo, corporações da indústria farmacêutica.

3.2.1.5 Visibilidade da pesquisa

A ampliação da visibilidade da pesquisa é um fator que favorece o interesse por parte dos pesquisadores e suas instituições para promoverem a comunicação dos dados de pesquisa, conforme ressaltaram Dallmeier-Tiessen et al. (2014) e Higman e Emmott (2015). Para os autores, a ampliação da divulgação dos dados aumenta também a possibilidade de reconhecimento do trabalho e atribuição da autoria àqueles resultados.

Uma das formas de expressão da visibilidade é a taxa de citação atribuídas às publicações. No contexto da ciência aberta, argumenta-se que quanto mais abertamente um trabalho for comunicado, maior será seu alcance e tão maior será sua taxa de citação. Até o momento não foi localizado nenhum estudo que confirmou ou refutou essa relação para o

contexto dos dados de pesquisa. Entretanto o argumento é recorrente nas discussões acerca do acesso aberto às publicações científicas.

3.2.1.6 Recompensa aos pesquisadores

Diferentemente da comunicação via artigos de periódicos, dados de pesquisa ainda não são valorizados como produção científica para os pesquisadores. Segundo Dallmeier-Tiessen et al. (2014), pesquisadores alegaram que o compartilhamento dos dados de pesquisa não gera *status* e nem promoção em suas carreiras, portanto, não compensava seus esforços. Essa discussão foi pautada por Melero, Hernández-San-Miguel (2014), Dallmeier-Tiessen et al. (2014) e Losoff (2009). De forma geral, os autores destacaram que o baixo valor acadêmico que é atribuído à divulgação dos dados somada às faltas de incentivos para a prática são aspectos que inibem o processo de divulgação dos dados de pesquisa.

É necessário ressaltar, no entanto, que a baixa valorização acadêmica refere-se apenas ao processo de divulgação, que compõem um dos três macroprocessos que envolve a noção de comunicação de dados de pesquisa adotada nesse estudo. Portanto, admite-se que a produção e o uso dos dados de pesquisa são processos valorizados e até mesmo requeridos para a promoção da carreira acadêmica dos pesquisadores.

3.2.1.7 Cultura acadêmica

Outro aspecto relacionado diretamente com os imperativos da ciência moderna é a cultura acadêmica vigente. O artigo de Dallmeier-Tiessen et al. (2014) foi o único a problematizar essa questão explicitamente. Os autores consideraram dois aspectos como influenciadores nas práticas da comunicação dos dados de pesquisa, a cultura predominante na área e a idade dos pesquisadores.

Segundo Dallmeier-Tiessen et al. (2014), a cultura refere-se às normas de comportamento manifestadas nas ações dos pesquisadores. A depender da área do conhecimento, existirão normas de comportamento mais ou menos favoráveis à ampla comunicação dos dados de pesquisa. No entanto, os autores ponderam que, em muitos casos, são os interesses comerciais envolvidos na pesquisa que determinarão o comportamento dos pesquisadores em relação a divulgação dos dados.

Somadas às normas de comportamento e aos interesses comerciais, Dallmeier-Tiessen et al. (2014) enfatizaram que o uso de tecnologias de informação e comunicação pode ser indutor de práticas mais colaborativas para a ciência, facilitando a interdisciplinaridade entre as

áreas e a abertura na divulgação dos dados de pesquisa. De acordo com os autores, o uso das tecnologias com tal finalidade é mais recorrente entre pesquisadores jovens.

3.1.2.8 Aspectos legais

Os aspectos legais envolvem as aplicações jurídicas nos três macroprocessos da comunicação dos dados de pesquisa. As legislações envolvidas nas discussões são diversas e por vezes conflituosas, conforme pontuou Dallmeier-Tiessen et al. (2014). De acordo com os autores, as divergências legais são decorrentes das dissonâncias entre as políticas de diferentes agências de fomento, países e regiões.

Childs e Mcleod (2004) e Dallmeier-Tiessen et al. (2014) ressaltaram alguns aspectos acerca das implicações dos direitos de propriedade intelectual na comunicação dos dados de pesquisa. Segundo Childs e Mcleod (2004) tais direitos podem regulamentar as práticas de armazenamento, acesso e reuso dos dados de pesquisa. Os autores não exploraram as nuances da relação, apenas apontaram-na. No mesmo sentido Sturges et al. (2015), Gardner (2004) e Thompson, Robertson e Greenberg (2015) apresentaram a temática como uma das barreiras para a ampla comunicação dos dados de pesquisa.

3.2.1.9 Aspectos éticos

Os aspectos éticos foram recorrentes nas discussões dos artigos analisados, eles estavam declaradamente presentes em oito dos treze artigos analisados. No entanto, observou-se que tais aspectos estão centrados em preocupações como a privacidade dos participantes (CHILDS; MCLEOD, 2004; DALLMEIER-TIESSEN ET AL., 2014; MELERO; HERNÁNDEZ-SAN-MIGUEL, 2014; THOMPSON, ROBERTSON; GREENBERG, 2015). Duas subcategorias foram identificadas a partir da discussão, a anonimização dos dados e a obtenção da permissão dos participantes, discutidas a seguir. Ambas as subcategorias foram apresentadas como alternativas para permitir a comunicação aberta dos dados e manter o direito à privacidade dos participantes.

3.2.2 Fatores relacionados ao contexto social das ciências da saúde

As pesquisas nas ciências da saúde partilham das prerrogativas da ciência moderna, caracterizadas no tópico anterior. No entanto, a área possui algumas especificidades que merecem destaque neste artigo. Suas singularidades favorecem a comunicação aberta dos dados de pesquisa, conforme sugeriu Knight (2015), Melero e Hernández-San-Miguel (2014) e Bond et al. (2013). Adicionalmente, os argumentos em defesa da ciência aberta no campo das

ciências da saúde ganham mais legitimidade em alguns temas de amplo interesse social, como é o caso da saúde pública. Portanto, os fatores para fundamentar a discussão sobre o contexto social trataram diretamente do prisma da saúde pública.

3.2.2.1 Saúde pública

Dois artigos ressaltaram o contexto das pesquisas em saúde pública como impulsionador da comunicação aberta dos dados de pesquisa. O mais recente foi publicado por Knight (2015). Apesar da relação entre a saúde pública e a comunicação dos dados de pesquisa não ser explorada diretamente, ela é evidenciada pela experiência relatada no artigo. A mesma relação encontra-se também presente de forma indireta no artigo de Melero e Hernández-San-Miguel (2014). A partir dos resultados obtidos em dois diretórios internacionais de políticas dessa natureza, os autores concluíram que as áreas relacionadas à saúde pública apresentam mais políticas de informação para o acesso aberto do que as outras áreas do conhecimento.

Melero e Hernández-San-Miguel (2014) atribuem a influência da política da agência Wellcome Trust (2011) como um dos elementos que configuram tal cenário. Portanto, faz-se necessário destacar os argumentos apresentados pela iniciativa e pelas discussões que a circundam. Segundo a política da Wellcome Trust (2011) a ampla comunicação dos dados de pesquisa acelera o progresso da saúde pública, melhorando assim a qualidade de vida das populações. O efeito, por si só, já é de grande interesse público e cumpre uma importante função social. No mesmo ano de publicação da política, Walport e Brest (2011) subscreveram-na e acrescentaram que a comunicação aberta dos dados de pesquisa maximiza o uso dos dados coletados e por consequência, a criação de novos conhecimentos a partir deles. De acordo com os autores, a maximização do uso dos dados, especialmente daqueles coletados das populações, é um princípio ético para as pesquisas na área da saúde pública. Isso porque, as populações não devem ser submetidas novamente à coleta de dados que já poderiam estar disponíveis.

3.2.3 Fatores relacionados ao financiamento

A experiência da agência Wellcome Trust (2011) e sua influência para o contexto das ciências da saúde ilustram a relevância das agências de fomento à pesquisa para a condução da abertura da comunicação dos dados. Nos casos em que o financiamento é de natureza pública ressalta-se a noção de que resultados constituem um bem público, ou seja, pertencente a toda sociedade. Tal perspectiva extrapola os pressupostos da ciência moderna de publicização dos resultados e busca se apresentar como um direito social.

3.2.3.1 Instituições financiadoras de pesquisa

As agências financiadoras de pesquisa, bem como as universidades, são consideradas agentes estratégicos para garantir a comunicação dos dados de pesquisa, conforme discutiu Dallmeier-Tiessen et al. (2014) e Douglass et al. (2014).

De forma geral, os autores defenderam que as instituições têm autoridade para requisitar parâmetros para coleta, divulgação e uso dos dados das pesquisas que financiam. Douglass et al. (2014) indicam as políticas de informação mandatórias como instrumentos estratégicos para estabelecer os requisitos do financiamento à promoção da comunicação dos dados de pesquisa de forma aberta. Segundo Dallmeier-Tiessen et al. (2014), isso permitirá que as instituições financiadoras de pesquisa otimizem os recursos investidos. A perspectiva otimista defendida por Dallmeier-Tiessen et al. (2014) baseia-se na expectativa que a comunicação dos dados de pesquisa tem potencial para promover uma economia nos recursos investidos na pesquisa em longo prazo.

3.2.3.2 Bem público

Além dos elementos discutidos na categoria aspectos legais sobre o direito de acesso à informação alguns autores pontuaram que pesquisas financiadas com recursos públicos, geram por excelência, bens públicos. Uma vez compreendidos os dados resultantes dessas pesquisas como bens públicos eles deveriam estar amplamente disponíveis para quem deles necessitassem. Para tanto, deveriam ser objetos de uma comunicação aberta. O argumento pode ser observado no artigo de Dallmeier-Tiessen et al. (2014) e Higman e Emmott (2015) de forma superficial e no artigo de Douglass et al. (2014), de forma mais aprofundada.

Douglass et al. (2014) discutiram a questão do bem público a partir das políticas de comunicação de dados do governo estadunidense. Os autores alegaram que, especialmente os dados de pesquisas e de agências governamentais obtidos com recursos públicos, devem ser livremente disponibilizados para o público. A defesa da alegação, entre outros aspectos, baseou-se no pronunciamento do então presidente dos Estados Unidos, Barack Obama, que afirmou que a prática seria um elemento que fortaleceria a democracia do país.

3.2.3.3 Custos

A discussão acerca do custo da comunicação dos dados de pesquisa centrou-se nos aspectos relativos aos investimentos para infraestrutura necessária, que é objeto ou demanda financiamento para estruturá-la. Dallmeier-Tiessen et al. (2014) e Higman e Emmott (2015) apontaram que os custos para a manutenção da infraestrutura podem ser elevados, mas que é altamente relevante para os objetivos da comunicação. Portanto, os custos envolvidos nos processos e na infraestrutura para suportar a comunicação dos dados de pesquisa precisam ser levados em consideração no financiamento da pesquisa. Nesse sentido, Dallmeier-Tiessen et al. (2014) enfatizam que a requisição para a divulgação dos dados pode ser uma forma de controlar os recursos investidos. No entanto, os autores pontuam que os custos devem ser considerados também para o longo prazo, de forma a garantir a preservação dos recursos.

Além das considerações acerca do custo para infraestrutura do sistema de informação propriamente dito, Melero e Hernández-San-Miguel (2014) e autores relacionados enfatizaram dois aspectos igualmente relevantes para a análise aqui proposta. O primeiro relaciona-se com a economia dos recursos e o segundo com os custos da coleta e análise dos dados.

3.2.4 Fatores relacionados com a infraestrutura

Um dos elementos centrais relacionados ao financiamento da pesquisa e da comunicação dos seus dados diz respeito aos custos da infraestrutura para esta finalidade. Nessa dimensão foram reunidas categorias que permitem a expressão e operacionalização do fenômeno investigado. Em decorrência do universo da discussão, a CI, sistemas de informação, padrões e políticas de foram as categorias mais presentes nos artigos analisados.

3.2.4.1 Sistemas de informação

A análise aqui empregada permitiu identificar que os aspectos referentes aos sistemas de informação foi o fator mais citado na literatura selecionada. Sob a categoria sistemas de informação foram classificados os temas relacionados com repositórios de dados de pesquisa, periódicos científicos, sistemas de descoberta e segurança, tecnologia e interoperabilidade dos sistemas de divulgação dos dados de pesquisa.

Compreendeu-se da análise que os repositórios de dados de pesquisa constituem alternativas eficientes para o depósito e armazenamento dos dados e, além disso, são capazes de absorver a diversidade do volume e os tipos de arquivos que os dados de pesquisa podem apresentar (LOSOFF, 2009; STURGES ET AL., 2015). No entanto, foram detectadas a

preocupação com a garantia da preservação e segurança dos dados, a capacidade de integração desses sistemas com outros, como, por exemplo, periódicos científicos (LOSOFF, 2009), a tecnologia envolvida (CHILDS; MCLEOD, 2004) e os sistemas de busca (DALLMEIER-TIESSEN ET AL., 2014; DOUGLASS ET AL., 2014).

Outros autores, como é o caso de Thompson, Robertson e Greenberg (2015) e Knight (2015), indicaram concordância em relação à necessidade de sistemas de informação destinados a operacionalização da comunicação dos dados de pesquisa. No entanto, os autores não envergaram nem para os repositórios de dados tão pouco para os periódicos. Os apontamentos dos autores foram apenas no sentido de evidenciar a necessidade e destacar, de maneira geral, que os sistemas deveriam ser seguros e interoperáveis. Nenhuma característica a respeito dos repositórios ou periódicos foram exploradas pela literatura analisada.

3.2.4.2 Padrões

A segunda categoria mais presente entre os artigos refere-se aos padrões necessários aos processos, procedimentos e sistemas que envolvem a comunicação dos dados de pesquisa. Childs e Mcleod (2004) e Whitmire, Boock e Sutton (2015) discutiram duas naturezas de padronização, uma relacionada aos procedimentos e outra ligada ao formato dos dados. De acordo com o levantamento realizado por Childs e Mcleod (2004), a falta de padronização nos procedimentos relacionados com o registro é uma problemática para a comunicação dos dados. Vale ressaltar que a investigação realizada pelos autores apontou tal barreira a partir da percepção dos próprios pesquisadores. Com isso, os autores concluíram que prestação de serviços de orientações aos pesquisadores bem como indicação de padrões é fundamental para o sucesso do projeto analisado no trabalho. Whitmire, Boock e Sutton (2015) não problematizaram a questão da importância dos padrões, no entanto sua análise fundamenta-se na classificação do tipo de dados de pesquisa por área do conhecimento e, portanto, pressupõe a lógica de padrão para a tipologia dos dados.

Os demais autores que abordaram a temática dos padrões (Thompson, Robertson, Greenberg, 2015; Dallmeier-Tiessen et al., 2014; Douglass et al., 2014; Knight, 2015) focaram no aspecto dos metadados. Thompson, Robertson e Greenberg (2015) indicaram o uso do padrão de metadados como elemento necessário à adoção de estratégia de preservação dos dados de pesquisa. A mesma abordagem foi adotada por Dallmeier-Tiessen et al. (2014), que afirmaram que a falta de padrão para os metadados constituem uma barreira para a efetiva preservação

dos dados de pesquisa. Douglass et al. (2014) e Knight (2015), por seu turno, consideraram que a incerteza sobre os padrões de metadados a serem adotados para a comunicação dos dados de pesquisa é impeditiva para o processo e reverbera na capacidade de interoperabilidade na gestão dos dados de pesquisa.

A discussão acerca da gestão dos dados de pesquisa, por sua vez, foi demarcada por oito artigos analisados. De forma geral, todos eles apontaram para a necessidade da existência de um serviço de informação ou ressaltaram sua importância em contextos onde este já se faz presente, conforme relatado por Knight (2015) e Whitmire, Boock, Sutton (2015).

3.2.4.3 Políticas

As políticas para a comunicação dos dados de pesquisa foram citadas por quatro artigos, a saber, Dallmeier-Tiessen et al. (2014), Melero e Hernández-San-Miguel (2014), Gardner (2004) e Sturges et al. (2015). Para Dallmeier-Tiessen et al. (2014) as políticas que requerem, e não apenas sugerem, a divulgação dos dados de pesquisa são incentivos estratégicos para o sucesso da prática. Segundo os autores as políticas precisam ser determinadas pelas instituições dos pesquisadores ou agências que fomentam suas pesquisas.

No mesmo sentido foi apresentado o argumento de Melero e Hernández-San-Miguel (2014). No entanto, os autores enfatizaram o aspecto da orientação para os princípios do acesso aberto na divulgação dos dados de pesquisa. Como isso, Melero e Hernández-San-Miguel (2014) ressaltam que a forma de disponibilização dos dados deve seguir as prerrogativas do acesso aberto pressupondo que ela seja realizada por meio da internet e sem restrições de acesso.

A perspectiva da necessidade da política de informação também parece fundamentar a discussão proposta por Gardner (2004). O autor, que problematizou a proposta do *Freedom of Information Act* (FOIA), afirmou que uma política para divulgação dos dados de pesquisa deve prever o período entre a finalização da pesquisa e a publicação dos dados brutos. Adicionalmente, a política deverá indicar o nível de exaustividade em que os dados poderão ser demandados.

Em todos os três casos evidencia-se a preocupação com a política como um instrumento que promove a comunicação dos dados de pesquisa com vistas a garantir a qualidade da comunicação científica. No entanto, a relação só foi declaradamente estabelecida por Sturges et al. (2015). Os autores discutiram o papel das políticas editoriais dos periódicos científicos como requisitos para garantir a transparência dos resultados das pesquisas que publicam. Nesse

sentido, a transparência seria um elemento-chave para promover a qualidade dos periódicos científicos.

3 CONSIDERAÇÕES FINAIS

Entre os resultados, foi possível, ao mesmo tempo, delinear as discussões predominantes sobre a comunicação dos dados de pesquisa no campo da CI e a identificar de fatores influenciadores na comunicação dos dados de pesquisa. No que se refere a estes últimos, foram propostas quatro dimensões. A primeira diz respeito aos imperativos da ciência moderna, onde foram identificadas nove categorias que se relacionam conceitualmente com as prerrogativas da ciência moderna: aceleração da ciência, reuso dos dados de pesquisa, transparência da ciência, assédio aos pesquisadores, visibilidade, recompensa aos pesquisadores, cultura acadêmica, aspectos legais e aspectos éticos. A segunda dimensão infraestrutura, cujas categorias recorrentes foram sistemas de informação, padrões e políticas. A terceira dimensão foi o contexto social em que se insere a discussão sobre comunicação dos dados de pesquisa, representada recorrentemente pelo campo das ciências da saúde em que o tópico saúde pública é predominante como contexto. Por fim, a dimensão do financiamento do processo de comunicação de dados de pesquisa surgiu como proeminente, no qual destacaram-se as categorias instituições financiadoras de pesquisa, a pesquisa e os dados como bem público e, não menos relevante, os custos associados custo.

Muito embora as primeiras discussões sobre a comunicação dos dados de pesquisa datem da década de 1970, naquele momento prevista no modelo de comunicação científica idealizado pelo UNISIST (UNESCO, 1971), a intensificação do interesse pelo tópico é recente, emergindo no bojo do interesse pela ciência aberta. Nesse aspecto, é necessário reconhecer uma limitação da revisão desta revisão sistematizada da literatura que foi negligenciar, por estratégia de viabilização do estudo, livros, capítulos de livros e comunicações apresentadas em conferências, além de artigos de periódicos indexados em outras bases de dados. Com isso, é possível que tenha escapado ao olhar deste trabalho algum conhecimento pré-existente sobre o tópico, como é o caso de livro organizado por Hurd (1996).

REFERÊNCIAS

- BOND, C. S. et al. The conceptual and practical ethical dilemmas of using health discussion board posts as research data. **Journal of Medical Internet Research**, v. 15, n. 6, p. e112, 2013.
- BORKO, H. Information science: What is it? **American Documentation**, v. 19, 1969.

BOULTON, G. Reinventing Open Science for the 21st Century. In: **Uma década de acesso aberto na UMinho e no Mundo**. Lisboa: Universidade do Minho, 2013. p. 239–250.

CHAN, L. et al. **Improving access to research literature in developing countries : challenges and opportunities provided by Open Access**. Trabalho / Paper. Disponível em: <<http://repositorio.unb.br/handle/10482/1009>>. Acesso em: 8 mar. 2016.

CHILDS, S.; MCLEOD, J. Sharing Research Records and Research Data: Findings from a Research Project in Higher Education. **New Review of Information Networking**, v. 10, n. 2, p. 131–145, 1 nov. 2004.

COMTE-SPONVILLE, A. **Dicionário Filosófico**. [s.l.] Martins Fontes, 2011.

COSTA, S. M. DE S. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. **Ciência da Informação**, v. 35, n. 2, 22 ago. 2006.

COX, A.; VERBAAN, E.; SEN, B. Upskilling Liaison Librarians for Research Data Management. **Ariadne**, n. 70, 2012.

DALLMEIER-TIESSEN, S. et al. Enabling Sharing and Reuse of Scientific Data. **New Review of Information Networking**, v. 19, n. 1, p. 16–43, 2 jan. 2014.

DOUGLASS, K. et al. Managing scientific data as public assets: Data sharing practices and policies among full-time government employees. **Journal of the Association for Information Science and Technology**, v. 65, n. 2, p. 251–262, 1 fev. 2014.

DROESCHER, F. D. As publicações em acesso livre e a avaliação de impacto. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 35, 5 dez. 2012.

FOSTER – Facilitate Open Science Training for European Research. Disponível em: <<https://www.fosteropenscience.eu>>. Acesso em: 7 dez. 2015.

GARDNER, W. Compelled Disclosure of Scientific Research Data. **The Information Society**, v. 20, n. 2, p. 141–146, 1 abr. 2004.

GARVEY, W. D. **Communication: The Essence of Science: Facilitating Information Exchange Among Librarians, Scientists, Engineers and Students**. [s.l.] Elsevier, 1979.

GRANT, M. J.; BOOTH, A. A typology of reviews: an analysis of 14 review types and associated methodologies. **Health Information and Libraries Journal**, v. 26, n. 2, p. 91–108, jun. 2009.

HAILEY, D. Scientific harassment by pharmaceutical companies: time to stop. **CMAJ: Canadian Medical Association Journal**, v. 162, n. 2, p. 212–213, 25 jan. 2000.

HIGMAN, R.; EMMOTT, S. Research data management and openness: The role of data sharing in developing institutional policies and practices. **Program**, v. 49, n. 4, p. 364–381, 1 set. 2015.

JACOMY, M. et al. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. **PLOS ONE**, v. 9, n. 6, p. e98679, 10 jun. 2014.

- HURD, J. **Models of scientific communications systems**. Information Today, v. 9, n. 33, 1996.
- KNIGHT, G. Building a research data management service for the London school of hygiene & tropical medicine. **Program**, v. 49, n. 4, p. 424–439, 1 set. 2015.
- LOSOFF, B. Electronic Scientific Data & Literature Aggregation: A Review for Librarians. **Issues in Science and Technology Librarianship**, v. 59, 2009.
- LYNCH, C. A. Institutional Repositories: Essential Infrastructure For Scholarship In The Digital Age. **ARL: A Bimonthly Report**, v. 226, p. 327–336, 2003.
- MEADOWS, A. **A comunicação científica**. Brasília: Briquet Lemos, 1999.
- MELERO, R.; HERNÁNDEZ-SAN-MIGUEL, J. Acceso abierto a los datos de investigación, una vía hacia la colaboración científica. **Revista española de Documentación Científica**, v. 37, n. 4, 2014.
- MORRISON, H. **Freedom for scholarship in the internet age**. Tese de doutorado—Canadá: Simon Fraser University, 2012.
- RUTTENBERG, S.; RISHBETH, H. Sir Granville Beynon: A Celebration of his Eightieth Birthday on 24 May 1994 World Data Centres — past, present and future. **Journal of Atmospheric and Terrestrial Physics**, v. 56, n. 7, p. 865–870, 1 maio 1994.
- STRAUSS, A.; CORBIN, J. **Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento de teoria fundamentada**. [s.l.] Artmed, 2008.
- STURGES, P. et al. Research data sharing: Developing a stakeholder-driven model for journal policies. **Journal of the Association for Information Science and Technology**, p. n/a–n/a, 1 maio 2015.
- THOMPSON, C. A.; ROBERTSON, W. D.; GREENBERG, J. Where Have All the Scientific Data Gone? LIS Perspective on the Data-At-Risk Predicament. **College & Research Libraries**, p. crl13–500, 2015.
- TRINE FJORDBACK SØNDERGAARD; JACK ANDERSEN; BIRGER HJØRLAND. Documents and the communication of scientific and scholarly information: Revising and updating the UNISIST model. **Journal of Documentation**, v. 59, n. 3, p. 278–320, 1 jun. 2003.
- UNESCO; UNIONS, I. C. OF S. **UNISIST: study report on the feasibility of a world science information system**. [s.l.] Unesco, 1971.
- WALPORT, M.; BREST, P. Sharing research data to improve public health. **The Lancet**, v. 377, n. 9765, p. 537–539, 18 fev. 2011.
- WELLCOME TRUST. **Sharing research data to improve public health: full joint statement by funders of health research**. Disponível em: <<https://wellcome.ac.uk/what-we-do/our-work/sharing-research-data-improve-public-health-full-joint-statement-funders-health#development-of-the-joint-statement-of-purpose>>. Acesso em: 15 out. 2016.

WHITMIRE, A. L.; BOOCK, M.; SUTTON, S. C. Variability in academic research data management practices: Implications for data services development from a faculty survey. **Program**, v. 49, n. 4, p. 382–407, 1 set. 2015.