

LETRAMENTO INFORMACIONAL PARA REUSO DE DADOS NAS CIÊNCIAS SOCIAIS: REQUISITOS E COMPETÊNCIAS

INFORMATION LITERACY FOR DATA REUSE IN THE SOCIAL SCIENCES: REQUIREMENTS AND COMPETENCES

Janete Saldanha Bach Estevão^a

Faimara do Rocio Strauhs^b

RESUMO

Introdução: O ambiente da pesquisa científica vem sendo impactado pelo aumento da disponibilidade de dados em acesso aberto. De outro lado, no entanto, emergem dificuldades aos pesquisadores na gestão dos seus próprios dados e de terceiros, principalmente, em relação às questões ligadas ao reuso dos mesmos. Esta relação de disponibilidade *versus* gestão de dados é particularmente complexa nas Ciências Sociais, cujos dados são heterogêneos e variados. **Objetivo:** Propor requisitos e competências mínimas para o letramento informacional em reuso de dados. **Metodologia:** Partiu-se de um levantamento bibliográfico, empregando-se técnicas de análise de conteúdo com apoio da ferramenta NVivo para análise em três instâncias de investigação em que: (1) compara-se os quadros conceituais (*frameworks*) de letramento em dados; (2) reutiliza-se dados secundários abertos de pesquisa (*datasets*) e (3) identifica-se em Ambientes Virtuais de Aprendizagem quais os requisitos e as competências para o reuso de dados. **Resultados:** Propõe-se os requisitos e as competências para o letramento em reuso de dados e testa-se a sua validade por meio da aderência com as melhores práticas recomendadas do *World Wide Web Consortium* (W3C). **Conclusões:** Consolida-se a proposta de 16 requisitos de um ambiente virtual para o letramento em reuso de dados contextualizados na área das Ciências Sociais e de 37 competências para os pesquisadores. Destacam-se como requisitos principais no letramento para o reuso de dados o conhecimento ao acesso, às licenças e ao uso destas, e como competências essenciais as regras de acesso, os tipos de licenças, bem como, o usar conjuntos de dados disponíveis, principalmente os critérios de proveniência, de qualidade e de confiabilidade.

^a Doutora pelo Programa de Pós-Graduação em Tecnologia e Sociedade (PPGTE-UTFPR). Bibliotecária-Documentalista da Universidade Federal do Paraná (UTFPR). E-mail: janeteest@gmail.com

^b Doutora em Engenharia de Produção pela Universidade Federal de Santa Catarina (UFSC). Professora Titular da Universidade Tecnológica Federal do Paraná (UTFPR). E-mail: faimara@utfpr.edu.br

Descritores: Ciência Aberta. Dados de Pesquisa – reuso. Letramento em dados.

1 INTRODUÇÃO

A Ciência Aberta é uma prática que pressupõe maior participação cidadã no desenvolvimento de processos científicos e tecnológicos, ampliando sua ação na geração de conhecimento científico e no uso de uma maior variedade de canais de comunicação da informação científica, a partir de linguagens mais simples e que visem aproximar o público em geral daquilo que resulta dos trabalhos de pesquisa (FECHER; FRIESIKE, 2014).

A colaboração, e a contribuição da Ciência Aberta, pode ser considerada, por exemplo, em práticas como a disponibilização gratuita dos dados de pesquisa, das notas de laboratório e de processos que permitam a reutilização, a redistribuição e a reprodução da pesquisa, de seus dados e métodos (FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH, 2015, *online*).

Desde o início da década de 2000, surgiram instrumentos para fomento e regulação da gestão de dados de pesquisa como, por exemplo, a Declaração de Berlin (OPEN ACCESS, 2003), que promovia o compartilhamento do conhecimento a partir da ampliação da disponibilização de dados e as políticas criadas por órgãos de fomento à pesquisa, que visam intensificar o acesso aberto como meio de comunicação dos resultados da pesquisa científica, especialmente aquelas financiadas com recursos públicos (NATIONAL SCIENCE FOUNDATION, 2010).

Exemplos de entidades que advogam o acesso aberto aos dados de pesquisa são a *American National Science Foundation* (NSF), a *United States National Institutes of Health* (NIH), o *National Institute of Standards and Technology* – instituições norte-americanas, a Comissão Europeia e o *Australian Department of Innovation*, que juntos formam o *Research Data Alliance* (RDA). O RDA é um esforço internacional que visa construir uma infraestrutura social e técnica para permitir o compartilhamento de dados. A aliança reunia, em novembro de 2017, 6.300 membros de 132 países (RESEARCH DATA ALLIANCE, 2017).

São as exigências de agências de fomento, por meio de suas políticas institucionais, no entanto, que tem efetivamente permitido a abertura dos dados de pesquisa, relevando-se, igualmente, o papel de políticas nacionais como as dos Estados Unidos e de alguns países da Europa, (CREASER, 2011; GRIGG, 2015). Como destaque de política de exigência de compartilhamento, tem-se o maior programa de pesquisa científica já criado na Europa, o *Horizon 2020* (H2020), cujo investimento nestas, previsto entre 2014 e 2020, é de 80 bilhões de Euros (EUROPEAN COMMISSION, 2016). Garantir a publicidade e o acesso aberto aos resultados das pesquisas é condição mandatória para integrar o programa, inclusive para o pagamento de bolsas aos candidatos (EUROPEAN COMMISSION, 2016, p. 2016-2019).

A concessão de recursos para pesquisa tem sido cada vez mais vinculada às exigências de compartilhamento dos seus dados, contudo, infere-se que iniciativas e serviços de acesso aos dados e suporte para o reuso sejam tão necessários quanto à disponibilidade dos mesmos. Ainda que dados de pesquisa estejam depositados em repositórios se não forem divulgados, se o acesso estiver restrito, ou ainda, se não estiverem formatados em condições que viabilizem o seu uso, perdem o potencial de gerar valor às novas pesquisas. Quando os dados não estão em condições de potencial reuso, ou seja, não se encontram em padrões pertinentes para propiciar novos estudos e abordagens, gerando conhecimento novo, contribuem apenas para a saturação dos repositórios.

Os padrões para reuso de dados obedecem a um conjunto de princípios estabelecidos para garantir a usabilidade dos mesmos no longo prazo, e são conhecidos como princípios FAIR – *Findable, Accessible, Interoperable and Reusable*, significando que os dados precisam ser localizáveis, acessíveis, interoperáveis e reutilizáveis (WILKINSON *et al.*, 2016). A possibilidade de reuso de dados, salienta-se, passa pela exigência de habilidades dos pesquisadores para lidar com vastas quantidades de dados, como a alfabetização e a fluência tecnológica e informática (CREASER, 2011). Essas habilidades enquadram-se no âmbito do letramento informacional em dados, que devem atender às novas exigências do processo de produção científica, que incorporam os dados de

pesquisa (CARLSON *et al.*, 2011; JEFFRYES; JOHNSTON, 2013; MAYBEE *et al.*, 2015).

O letramento informacional em dados é um dos serviços relacionados aos dados de pesquisa (*Research Data Services – RDSs*), segundo Tenopir, Birch e Allard (2012), e tem sido promovido por várias bibliotecas universitárias (ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES, 2018). Essas instituições promovem ações de letramento para auxiliar os pesquisadores no acesso e no reuso de dados de pesquisa com algumas iniciativas internacionais identificadas na literatura (CARLSON *et al.*, 2011; KOLTAY, 2015; 2016).

Diante da crescente disponibilidade dos dados de pesquisa, da potencialidade de reuso dos mesmos e da oportunidade de letramento informacional para os pesquisadores, evidencia-se uma lacuna de estudos que envolve os desafios da comunicação científica em face de ambientes de pesquisa cada vez mais complexos, impulsionados pela geração e pelo uso de grandes volumes de dados, conhecido como o quarto paradigma da ciência (GRAY *et al.*, 2005). Neste sentido investigou-se quais seriam os requisitos necessários e as competências mínimas de letramento informacional para reuso de dados de pesquisa, principalmente em relação aos critérios que atestem a proveniência, a qualidade e a confiabilidade dos dados, reportando-se aqui os resultados.

Este artigo apresenta, como estrutura, a revisão de literatura com os conceitos fundantes – o letramento informacional em dados e os dados de pesquisa, seguida da metodologia utilizada para a coleta e a análise de dados, a apresentação e a discussão dos resultados, finalizando com a proposta e o resultado da validação dos requisitos.

2 REVISAO DE LITERATURA

O letramento informacional é um conceito consolidado e amplamente aceito na Ciência da Informação e na Biblioteconomia, sendo descrito como a habilidade de reconhecer a necessidade de buscar, de identificar, de localizar, de avaliar e de usar informações para resolver um determinado problema (AMERICAN LIBRARY ASSOCIATION, 1989). Para Calzada Prado e Marzal

(2013), há outra competência tão essencial quanto o letramento informacional, no entanto: o letramento em dados.

Segundo Carlson e Johnston (2015), o letramento em dados está na fase de surgimento e sua terminologia, para Koltay, Horstmann e Witt (2016) não é homogênea. O termo letramento em dados está referenciado na literatura de diversas formas, tais como: letramento em dados científicos (QIN; D'IGNAZIO, 2010), letramento em dados de pesquisa (SCHNEIDER, 2013) e letramento em informação de dados (CARLSON *et al.*, 2011).

Carlson *et al.* (2011) conceituam letramento em dados como a capacidade de coletar, usar e avaliar dados de forma a entender o que estes significam, incluindo habilidades como ler gráficos e tabelas adequadamente, inferir conclusões corretas a partir dos dados e reconhecer o uso equivocado ou inadequado destes (CARLSON; JOHNSTON, 2015).

O letramento em dados passou a ser considerado como uma parte do letramento informacional, mas Koltay (2016) destaca que o conceito é adotado de forma autônoma por vários outros autores. Koltay (2015; 2016) distingue entre os conceitos de letramento informacional e de letramento em dados, em que descreve as particularidades próprias da atividade científica sob a perspectiva dos dados. As fronteiras entre os dois conceitos não são rígidas.

Ainda que o letramento informacional inclua as habilidades de acesso, de apropriação e do uso da informação (AMERICAN LIBRARY ASSOCIATION, 1989), e o letramento em dados considere também habilidades informáticas de manipulação de registros e arquivos (CREASER, 2011), considera-se que o **letramento informacional em reuso de dados** seja o ponto de convergência das prerrogativas de ambas conceituações (CARLSON *et al.*, 2011; JEFFRYES; JOHNSTON, 2013; MAYBEE *et al.*, 2015). Sob a perspectiva do reuso de dados, são necessários requisitos que evoquem análises críticas próprias do letramento informacional (necessidade, disponibilidade, acesso, aquisição de dados), quanto habilidades em curadoria e gestão dos dados, próprias do letramento em dados (KOLTAY, 2016).

Quanto aos dados de pesquisa, a conceituação varia de acordo com as disciplinas e não há um consenso. Normalmente, referem-se ao “material de fato

registrado comumente aceito na comunidade científica como necessário para validar os resultados da pesquisa” (ENGINEERING AND PHYSICAL SCIENCES RESEARCH COUNCIL, 2018, *online*, tradução nossa). Os dados são formados por variados tipos de registros, podendo ser estruturados e armazenados em distintos formatos de arquivos.

Os dados de pesquisa também possuem finalidade diferente de outros tipos de conteúdo pois "são coletados, observados ou criados, para fins de análise para produzir resultados de pesquisa originais" (UNIVERSITY OF EDINBURGH, 2016, *online* – tradução nossa). São considerados uma “representação reinterpretabil da informação de forma formalizada, adequada à comunicação, interpretação ou processamento” (DIGITAL CURATION CENTER, 2007, *online* – tradução nossa). Os dados requerem contextos e significados conectados para que sejam passíveis de reuso (WOOD, 2013). Nesse sentido, não basta apenas a disponibilidade, o papel dos metadados é essencial para assegurar a identificação e a reutilização dos dados. Segundo Sayão e Sales (2015), a disponibilidade dos dados viabiliza a preservação digital dos mesmos, que são deslocados dos discos rígidos e computadores pessoais dos pesquisadores e passam a ser depositados em ambientes adequadamente configurados para serem acessíveis no futuro.

Anjos e Dias (2019) reforçam o papel dos dados como um instrumento de apoio em diversas etapas do processo de investigação científica, assim como Freitas e Leite (2019), que consideram os dados de pesquisa principais envolvidos na comunicação científica tanto quanto como os pesquisadores, as editoras, as bibliotecas, as agências de fomento e as universidades. Portanto, a disponibilidade e o reuso de dados se complementam.

Devido à crescente preocupação com a reprodutibilidade de pesquisa, essa disponibilidade de dados para reuso vem sendo estimulada em diversas áreas. O estudo de Curty *et al.* (2017), a partir de respostas de 595 pesquisadores de diferentes áreas e países sobre a prática de reuso de dados, evidenciou, contudo, que os novos estudos que usam dados gerados em pesquisas anteriores disseminam-se mais nas Ciências Exatas e Biológicas, mais ainda enfrentam muita resistência nas Ciências Sociais

Curty *et al.* (2017) sustentam que as barreiras e os comportamentos inibidores ao reuso de dados das Ciências Sociais estão ligados às questões éticas, aos riscos de má interpretação e ao uso incorreto de dados. Essas características impeditivas do reuso de dados estão relacionadas ao letramento (ou a falta deste) – habilidades e competências em manejo dos conjuntos de dados secundários. Desta forma, o letramento em dados e o letramento informacional continuam sendo mandatórios como evidenciado em Calzada Prado e Marzal (2013) principalmente devido à complexidade do ambiente informacional que cresce de forma contínua.

Dessa perspectiva, para Koltay (2015), o letramento em dados tem também um papel de estímulo ao pensamento crítico, levando em conta os desafios de privacidade e de confiabilidade decorrentes da disponibilidade dos dados (KOLTAY, 2015). De acordo com Calzada Prado e Marzal, (2013), a relação entre letramento e qualidade dos dados é vital e deve considerar a avaliação das fontes em que os dados são coletados para evitar erros ou problemas. A importância da avaliação das fontes excede a qualidade da coleta os dados, antes é uma análise crítica crucial para a interpretação de dados secundários a fim de se apresentarem conclusões corretas.

Estabelecido o viés conceitual, apresenta-se, sequencialmente, o processo metodológico da pesquisa, cobrindo os procedimentos de levantamento e de revisão sistemática da literatura, a metodologia e as técnicas para análise dos conteúdos.

3 METODOLOGIA

Esta foi uma pesquisa de análise qualitativa, com procedimentos de levantamento bibliográfico, revisão sistemática da literatura com a localização dos dados e a análise comparativa do conteúdo de acordo com os preceitos de Bardin (1995).

A pesquisa fundou-se sobre três pilares: (i) as fontes de informação, (ii) as estratégias de buscas e (iii) o uso de filtros e recursos avançados:

- 1) As fontes de informação foram bases selecionadas nos seguintes portais de pesquisa: (i) O Portal de Periódicos da Capes, (ii) o

Portal EBSCOHost¹ e a (iii) Base referencial Web of Science.

- 2) A estratégia de busca foi construída utilizando-se os seguintes descritores: ("*open data*" OR "*research data*" OR "*scientific data*") AND ("*Data literacy*" OR "*information data literacy*" OR "*data information literacy*"); Em português usou-se: (“dados abertos” OR “dados de pesquisa” OR “dados científicos”) AND (“letramento de dados” OR “letramento em dados” OR “letramento informacional em dados” OR “letramento informacional de dados”). Aplicou-se ainda recursos como o uso das aspas na combinação dos conceitos entre si, que garantem a exatidão na recuperação da informação.
- 3) A aplicação de filtros e recursos avançados contou com os seguintes parâmetros: no Portal de Periódicos da Capes, o campo utilizado para busca no texto foi o de ‘Assuntos’. No portal da EbscoHost, o campo padrão selecionado para a busca foi: ‘Assunto, títulos e resumo’. Já na base Web of Science, o campo de busca utilizado foi ‘Tópico’, que engloba os campos Assuntos, títulos e resumo, automaticamente. Em todas as fontes pesquisadas, foram aplicados os filtros de pesquisa por período de cobertura entre os anos de 2012 e 2019. Também se adotou o critério de periódicos revisados por pares.

Com o uso destas estratégias foram recuperadas 173 referências, como apresentado na Tabela 1.

Tabela 1 - Frequência de conceitos-chaves – busca combinada

Busca Combinada	Referências Capes	Referências EBSCOHost	Referências Web of Science	Total Referências
“Open Research Data” (AND) “Data information literacy”	19	116	38	173

Fonte: Autoria própria (2019).

¹ Dentre as bases de dados disponíveis no portal EBSCOHost, as mais relevantes para o tema desta pesquisa foram a Information Science & Technology Abstracts (ISTA) e a Library, Information Science & Technology Abstracts with FullText (LISTA), por cobrirem a Ciência da Informação.

Os registros recuperados no levantamento bibliográfico foram os insumos para a revisão sistemática da literatura (RSL), constituída por artigos que compuseram o *corpus* teórico de pesquisa (PAGANI; KOVALESKI; RESENDE, 2017). Adotou-se da metodologia de Guertin e Bernhard (2005) as seguintes etapas da RSL: especificação do assunto, definição das estratégias de localização da informação, seleção dos documentos, extração e tratamento das informações e elaboração do referencial teórico.

No tratamento dos dados para a análise dos quadros conceituais (*frameworks*) e dos dados de pesquisa secundários coletados, utilizou-se a análise de conteúdo, considerada por Bardin (1995) uma técnica empírica, que segue princípios que devem ser consistentes ao longo de toda análise, passando por quatro fases: (i) busca exaustiva de conteúdo representativo, homogêneo e pertinente; (ii) análise quali-quantitativa; (iii) interpretações das evidências levantadas; (iv) descrição das inferências alcançadas.

A proposta metodológica, de forma esquemática, é composta de cinco etapas representadas na **Erro! Fonte de referência não encontrada..**

Figura 1 - Proposta metodológica



Fonte: Autoria própria (2019).

Especificado o tema, as três primeiras etapas de pesquisa – 1, 2 e 3 – Figura 1 – foram de análise em triangulação de três instâncias de investigação, com vistas ao reuso de dados já aplicados, para as extrações de requisitos.

Na primeira instância identificou-se *frameworks* (quadros conceituais) na

literatura que apontassem competências essenciais para a gestão de dados de pesquisa. Os *frameworks* selecionados para esta análise foram os de Calzada Prado e Marzal (2013), Carlson e Johnston (2015), Carlson *et al.* (2011), Foster (2015), Jeffryes e Johnston (2013) e Vitae (2018). A análise desses permitiu levantar competências gerais de letramento em dados.

Em seguida reutilizou-se três conjuntos de dados secundários (*datasets*) disponíveis, compostos de pesquisas que tratavam da experiência empírica de pesquisadores no reuso de dados sendo: (1) cientistas de instituições norte-americanas que relatam o comportamento de pesquisadores em compartilhar dados (KIM, 2016); (2) respostas de pesquisadores sobre seu próprio comportamento em reusar dados (KIM, 2017), e (3) estudo com pesquisadores que examinou atitudes e comportamentos relacionados ao compartilhamento e a reutilização de dados em suas próprias pesquisas (BERGHMANS *et al.*, 2017).

A terceira instância de análise investigou em onze Ambientes Virtuais de Aprendizagem – AVAS, quais as competências para o letramento especificamente no reuso de dados que estão sendo abordadas – Quadro 1. As competências foram comparadas com os requisitos identificados nas análises dos *frameworks*.

A partir dos resultados comparados das análises, a etapa seguinte (4) foi a de identificar os requisitos necessários para a proposta de requisitos e de competências para o letramento em reuso de dados contextualizadas nas Ciências Sociais. Finalizando, a etapa (5) validou a proposta pelo grau de aderência das competências às 35 recomendações de boas práticas do World Wide Web Consortium (W3C) (LÓSCIO; BURLE; CALEGARI, 2017).

Para apoiar a fase de análise dos resultados utilizou-se o *software* NVivo nas etapas de codificação e de classificação das categorias. Os dados foram codificados recebendo atributos que os qualificam em uma determinada categoria. A organização e a ordenação de dados já qualificados, mas situados em diferentes classes – categorias de contexto, categorias de análise e unidades de registro – a partir de uma ou mais variáveis, é o que se denomina etapa da classificação, segundo os conceitos de Bardin (1995).

A categoria de contexto é o elemento que permite criar o significado maior

para determinar as unidades de análise (intermediárias) e as unidades de registro (finais) (BARDIN, 1995). As unidades de registro, no entanto, determinam o tema-eixo, podendo orientar todo o processo de classificação.

Os onze recursos analisados encontram-se no Quadro 1.

Quadro 1 - Recursos de Aprendizagem online analisados

Item	Título	Referência
(1)	Bodleian Data Library – University of Oxford	University of Oxford (2018-2019)
(2)	Consortium of European Social Science Data Archives (CESSDA) e European Research Infrastructure Consortium (ERIC)	Consortium Of European Social Science Data Archives (2018)
(3)	Data Management Course	Johnston e Jeffryes (2012)
(4)	Data Observation Network for Earth (DataONE)	DATAONE (2012);
(5)	Digital Curation Center - DC 101 materials: Access and Reuse	Digital Curation Center (2014-2019)
(6)	Essentials 4 Data Support	Research Data Netherlands (<i>online</i>)
(7)	Facilitate Open Science Training for European Research	FOSTER (2015)
(8)	Managing and Sharing Research Data handbook	Corti <i>et al.</i> (2014)
(9)	Research Data and Management Training – MANTRA	University of Edinburgh (2017)
(10)	New England Collaborative Data Management Curriculum (NECDMC)	LAMAR SOUTTER LIBRARY (2012)
(11)	Research Data Management and Sharing – Coursera	Coursera (2016)

Fonte: Autoria própria (2019).

Utilizaram-se diversos recursos de codificação e de visualização da informação, incluindo as técnicas de codificação automatizada peculiares ao *software* NVivo, mas também codificação manual, quando necessário. A codificação utilizou-se dos arquivos dos *frameworks* para identificar as categorias de contexto e as categorias de análise do estudo. A classificação das unidades de registro embasou a identificação das Categorias de Análise, camada mais ampla do que as Unidades de Registro (BARDIN, 1995).

A identificação das Categorias de Análise partiu do agrupamento semântico e léxico da terminologia usada nas Unidades de Registro, recuperadas da codificação manual dos *frameworks*. Exemplificando, a partir da recuperação, da descrição e da análise dos termos *Access*, *Critically assess*, *Data Access* e *Discovery*, optou-se por manter a nomenclatura única *Access*, como categoria de análise conforme apresentado no Quadro 2, que traz os

exemplos *Access* e *Acquisition*.

Quadro 2 - Exemplo de agrupamento de categorias de análise de *frameworks*

CATEGORIAS DE ANÁLISE	UNIDADES DE REGISTRO	CATEGORIAS DE ANÁLISE	UNIDADES DE REGISTRO
<i>Access</i>	<i>Access</i>	<i>Acquisition</i>	<i>Acquisition</i>
	<i>Critically assess</i>		<i>Acquisition of Data</i>
	<i>Data Access</i>		<i>Collecting</i>
	<i>Discovery</i>		<i>Collecting</i>
	<i>Discovery</i>		<i>Extract</i>
	<i>Information seeking</i>		<i>Finding and/or obtaining data</i>
	<i>Locate</i>		

Fonte: Autoria Própria a partir de análises do NVivo (2019).

Após a análise e a extração dos conceitos combinados dos *frameworks*, foram identificadas 17 categorias: (1) Access (Acesso), (2) Acquisition (Aquisição), (3) Analysis (Análise), (4) Archiving (Arquivamento), (5) Curation (Curadoria), (6) Data management (Gerenciamento de Dados), (7) Documentation (Documentação), (8) Licences (Licenças), (9) Metadata (Metadados), (10) Open Science (Ciência Aberta), (11) Policy (Política), (12) Quality (Qualidade), (13) Reuse (Reutilização), (14) Sharing (Compartilhamento), (15) Storage (Armazenamento), (16) Using (Uso), (17) Visualization (Visualização). Este conjunto inicial de requisitos foi comparado com as demais instâncias de análises, mantendo-se apenas os requisitos relativos ao reuso, cujos resultados são demonstrados na sequência.

4 RESULTADOS DAS ANÁLISES DOS DADOS SECUNDÁRIOS (DATASETS) E DOS RECURSOS DE APRENDIZAGEM ONLINE

Comparou-se os requisitos identificados nos *frameworks* com as perspectivas e as dificuldades dos pesquisadores na prática do reuso de dados secundários utilizados nas próprias pesquisas (Seção 4.1).

A partir da análise dos Recursos de Aprendizagem *Online* – Seção 4.2, obteve-se uma base para a proposta conceitual dos requisitos e das competências essenciais para o reuso de dados (Seção 4.3). Sequencialmente, esses requisitos foram validados pelos critérios de melhores práticas da W3C, consolidando a proposta de pesquisa (Seção 4.4).

4.1 RESULTADOS DA ANÁLISE DOS *DATASETS*

A análise dos *datasets* de Kim (2016), de Kim (2017) e de Berghmans *et al.* (2017), demonstrou a percepção dos pesquisadores e também as oportunidades de letramento que corroboraram para a elaboração da proposta desta pesquisa. Salienta-se que estes *datasets* são dados secundários disponíveis em repositórios de dados abertos.

A pesquisa de Kim (2016), investigou a percepção dos pesquisadores sobre as práticas de compartilhamento de dados, identificando entre outros elementos: padrões de metadados; impacto nas métricas de citação; elaboração e localização da documentação adjunta aos conjuntos de dados, minimizando risco de mau uso; identificação e localização de recursos disponíveis para auxílio, como treinamentos *online*.

A pesquisa de Kim (2017), avaliou o comportamento do pesquisador quanto ao reuso de dados de pesquisa em várias áreas do conhecimento e levantou, entre outros elementos: receio da má interpretação dos dados, seguido do medo de causar algum tipo de infração no reuso de destes; maior compreensão das regras de licenças de uso e reuso; melhor entendimento da documentação de apoio; capacitação nas atividades de coleta e curadoria dos dados; lacuna de oportunidade para que as bibliotecas acadêmicas possam suprir a ausência de pessoas preparados para capacitar os pesquisadores em gerenciar os seus próprios dados.

Berghmans *et al.* (2017), avaliaram um amplo conjunto de dados de pesquisadores de múltiplas áreas e em diferentes países sobre os hábitos e as percepções nas práticas diárias de pesquisa detectando: necessidade de maior familiaridade com conceitos de Ciência, acesso e dados abertos; 40% da confiança dos pesquisadores nos dados está na documentação completa e nos metadados adequados; 60% de todos os respondentes não sabem qual o tipo de licença escolher; percepção dos pesquisadores de que não há padrões para a citação de dados de pesquisa nas suas respectivas áreas.

4.2 ANÁLISE DOS RECURSOS DE APRENDIZAGEM ONLINE – RESULTADOS

Os resultados das análises foram elaborados a partir das buscas textuais de 685 arquivos gerados da análise dos 11 recursos de aprendizagem. Para cada competência identificada, foi elaborada uma estratégia de pesquisa personalizada. O *Software* NVivo permite a extração de trechos localizados na busca, agrupando-os por ordem dos documentos pesquisados. Todos os conteúdos recuperados foram analisados, totalizando 560 páginas com trechos das estratégias de busca executadas.

Os trechos selecionados foram a base para as descrições para cada competência. Os mais recorrentes entre os avaliados foram, em ordem de representatividade entre os recursos analisados, os requisitos identificados como 08. *Licences*, 01. *Access* e 16. *Using* – Quadro 3.

Quadro 3 – Requisitos mais recorrentes - recursos de aprendizagem

Códigos	Número de itens codificados	Número de referências de codificação
01. Access	207	355
08. Licences	243	635
16. Using	190	587

Fonte: Autoria própria (2019), com base no Resultado do *Software* NVivo.

A partir da análise em triangulação dos resultados obtidos com os *frameworks*, os *datasets* e os recursos de aprendizagem *online*, propõe-se o conjunto de requisitos com vistas ao reuso de dados nas Ciências Sociais.

4.3 PROPOSTA DO CONJUNTO DE REQUISITOS E COMPETÊNCIAS ESSENCIAIS AO LETRAMENTO EM DADOS

Identificou-se 16 categorias de análises – requisitos – e 37 unidades de registro – competências, para formar um conjunto de requisitos e competências essenciais ao letramento informacional em reuso de dados. O conjunto de requisitos e de competências pode indicar aos pesquisadores, de forma antecipada, as exigências necessárias para a viabilizar a reutilização de dados secundários – Quadro 4.

Para aqueles sem experiência nesta prática, as necessidades de

letramento na reutilização podem se configurar desafios a serem enfrentados, principalmente por não serem conhecidas antecipadamente ao reuso efetivo dos dados.

Quadro 4 - Requisitos e competências para o letramento em reuso de dados

Categorias de Análise	Unidades de Registro	Categorias de Análise	Unidades de Registro
1. Access	1.1 Access data sources	7. Licences	7.3 Licenses types and copyright
	1.2 Discovery		7.4 Using data ethically
	1.3 Access data	8. Metadata	8.1 Metadata Description
2. Analysis	2.1 Reading		8.2 Metadata Standards
	2.2 Interpreting/ Understanding data	9. Open Science	9.1 Concept of Open Science
	2.3 Evaluating / 2.4 Critical thinking		9.2 Open Access
3. Acquisition	9.3 Open Data Journals		
3. Acquisition	3.1 Obtaining data		9.5 Open Data use and reuse
	3.2 Collecting Data	10. Policy	10.1 Funding policies
4. Curation	4.1 Conversion	11. Quality	11.1 Quality Assurance
	4.2 Formats	12. Reuse	12.1 Reproducible Research
	4.3 Merging		12.2 Legal requirements
	4.4 Processing		12.3 Re-use policies
	4.5 Synthesize4	13. Sharing	13.1 Re-sharing
5. Data Management	5.1 Data Management Plan	14. Using	14.1 Databases
	5.2 Naming conventions		14.2 Data collections
6. Documentation	6.1 Data sets description	15. Visualization	15.1 Visualization tools
7. Licences	7.1 Citation	16. Support and formal training	16.1 Support and formal training
	7.2 Intellectual Property Rights		

Fonte: Autoria própria (2019).

Os requisitos (CA) e as competências essenciais (UR) foram descritas e explicadas nas unidades de contexto (UC) de forma hierárquica, com níveis de recuo, para facilitar a visualização, sendo: Nível 1 - CA (Categorias de Análise), Nível 2 - UR (Unidades de Registro) e Nível 3 - UC (Unidades de Contexto).

Detalha-se a título de exemplo a Categoria de Análise Acesso (Access) e as Unidades de Registro (Access Data Sources), (Discovery) e (Access Data). As Unidades de Contexto foram indicadas pela referência numérica para cada plano conceitual de origem – Quadro 01 (Framework 01 - F01, Framework 02 - F02, Framework 03 - F03, Framework 04 - F04, Framework 05 - F05 e Framework 06 – F06).

CA 1. Acesso. (Access)

- UR 1.1 *Acessar Fonte de Dados (Access Data Sources)*
- UC Localizar fontes de informação adequadas a essa necessidade; Capacidade de acessar fontes de dados apropriadas às informações necessárias (F02). Estar ciente das possíveis fontes de dados, ser capaz de avaliá-las. (F02). Usar diferentes portais de pesquisa do Acesso Aberto (F03). Diferenciar nível de direitos de acesso (F04). Avaliar a confiabilidade, reputação, moeda, autoridade e relevância das fontes. (F04).
- UR 1.2 *Descobrir (Discovery)*
- UC Localizar e utilizar repositórios de dados disciplinares. (F01). Definir com precisão a necessidade informacional (F02). Identificar e acessar recursos bibliográficos apropriados, arquivos e outras fontes de informações relevantes, incluindo recursos baseados na *Web*, fontes primárias e repositórios. (F04)
- UR 1.3 *Acessar os Dados (Access Data)*
- UC Identificar os possíveis controles de acesso. Planejar os controles de acesso apropriados. (F06). Relatar o acesso potencial (F06).

Os requisitos e as competências foram submetidos a validação bibliográfica, com verificação da aderência às recomendações das melhores práticas da W3C (LÓSCIO; BURLE; CALEGARI, 2017).

4.4 VALIDAÇÃO DO GRAU DE ADERÊNCIA ÀS RECOMENDAÇÕES W3C

A partir da comparação da descrição das melhores práticas da W3C com os requisitos e as competências propostos, foi possível verificar se as mesmas tinham algum grau de aderência às recomendações para gestão de dados na *internet* (LÓSCIO; BURLE; CALEGARI, 2017). Todas as 35 melhores práticas da W3C foram identificadas como aplicáveis aos requisitos de reuso de dados. O grau de aderência foi determinado pela quantidade de recomendações propostas pela W3C que eram correspondentes aos requisitos e competências. Quatro competências foram aderentes a mais de três recomendações da W3C, demonstrando forte conexão com estas. São elas: *1.3 Access data*, *7.1 Data sets description*, *8.2 Citation* e *12.1 Quality Assurance*.

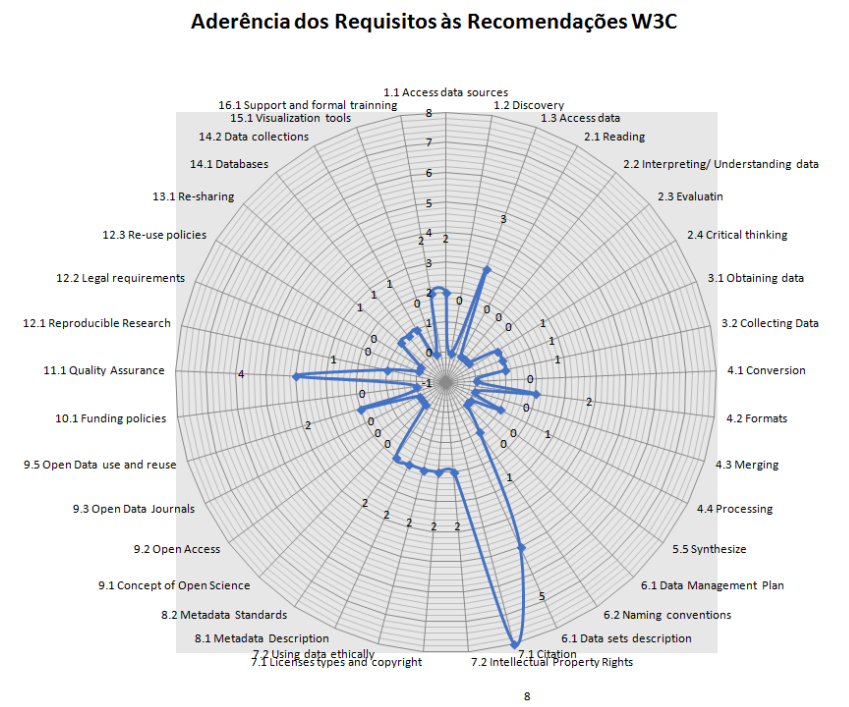
Os requisitos *1.2 Discovery*, *2.1 Reading*, *2.2 Interpreting/ Understanding data* não tiveram aderência, no entanto. Porém, destaca-se que são competências mais relacionadas com a capacidade analítica dos pesquisadores do que com capacidade técnica. Por ser esta última a ênfase do conjunto de recomendações da W3C, a não aderência não compromete a validade dos

resultados.

Outros requisitos sem aderência: 6.1 Data Management Plan, 10.1 Concept of Open Science, 10.2 Open Access, 10.3 Open Data Journals, 11.1 Funding policies, 13.2 Legal requirements, 13.3 Re-use policies. Essas competências estão mais próximas dos conceitos acadêmicos e do reuso de dados de pesquisa, considera-se, portanto, o fato de as recomendações da W3C serem mais direcionadas para o uso de dados de forma ampla na internet, o que pode justificar a não aderência à W3C. As competências 5.1 Conversion, 5.3 Merging e 17.1 Visualization tools apresentam características de funcionalidades de softwares e não encontraram aderência real às recomendações da W3C.

Sendo assim, têm-se um percentual de aderência de 92% entre as demais 37 competências identificadas e as recomendações da W3C, o que se infere ser um percentual significativo para validar a proposta. O nível de aderência de cada um dos requisitos às melhores práticas é apresentado no **Erro! Autoreferência de indicador não válida..**

Gráfico 1 - Aderência dos Requisitos às Recomendações W3C



Fonte: Autoria própria (2019).

Apresentada a proposta de requisitos e competências, apresentam-se as

conclusões finais.

5 CONCLUSÕES

Os dados de pesquisa passaram a ser mais evidenciados, exigidos, monitorados e até mesmo usados como critérios de aceite para publicações. Isso culminou em uma maior disponibilidade desses conjuntos de dados, em acesso aberto, em diversos repositórios no mundo. O indexador re3data², por exemplo, listava em janeiro de 2020, 2.454 repositórios de dados de pesquisa, em todas as áreas do conhecimento, hospedando milhares de conjuntos de dados em cada repositório. Contudo, evidenciou-se a necessidade de letramento informacional para o reuso de dados por parte dos pesquisadores, nem sempre acostumados e, portanto, não preparados para esta prática.

Os pesquisadores evidenciaram nas análises: receio de se equivocarem no reuso de dados; dificuldades para a localização destes; dificuldades com a documentação dos dados; limitações na compreensão das regras e formas de citação, de direitos autorais e de tipos de licenças dos dados, além de não serem capaz de distinguir os critérios de qualidade dos dados para o reuso. Esta falta de habilidades relacionadas ao reuso constituem barreiras para que o pesquisador se beneficie da disponibilidade de dados e, conseqüentemente, da apropriação dos benefícios potenciais estimulados por um ambiente cada vez mais orientado a abertura e a colaboração, como o da Ciência Aberta.

As competências mais recorrentes nos achados indicam relações com regras de acesso, tipos de licença e como usar os conjuntos de dados disponíveis. Apesar dos dados de pesquisa estarem sob a égide da Ciência Aberta, que implica em poucas, ou nenhuma barreira de acesso, nota-se que um dos resultados mais recorrentes refere-se às questões de licenças e de direitos autorais, demonstrando que o modelo de recompensa da Academia, que exige publicação e a constatação desta, ainda se sobressai sobre os princípios da Ciência e do Acesso Abertos.

Daí a relevância de uma proposta de requisitos para o letramento

² <http://re3data.org/>

informacional no reuso de dados, visando a melhoria qualitativa da prática de produção da Ciência. Essa apresenta competências que vão auxiliar pesquisadores não habituados a reutilizar dados de pesquisa no contexto da Ciência Aberta, em que os principais requisitos como o Acesso, as Licenças e o Uso, cobrem desde a primeira etapa do ciclo final do reuso de dados, permitindo identificar critérios de qualidade e da promoção da confiança no reuso dos conjuntos de dados. Sendo assim, o letramento para o reuso de dados inicia com o letramento informacional, como o acesso, a descoberta, a avaliação crítica e os direitos autorais de conteúdo dos *datasets*, chegando ao letramento em dados em si, que inclui atividades de curadoria, de padronização dos dados e do manejo de arquivos eletrônicos e *softwares* relacionados.

Dessa forma estabelece-se um ciclo virtuoso da comunicação dos resultados acadêmicos, em que mais recursos são disponibilizados, como os dados de pesquisa, com menos restrições possíveis, podendo ser reaproveitados, promovendo a retroalimentação, a colaboração e a integração dos recursos científicos. O fato dos dados estarem disponíveis não garante, por si só, sua reutilização, exigem-se competências neste processo. Sem o letramento para o reuso pesquisadores não tem acesso, não obtêm e não reusam dados, limitando suas fronteiras. Não usufruem dos valores inerentes à Ciência Aberta, dos seus recursos científicos que lhe conferem caráter mais democrático, com disponibilização acessível de dados de alto valor.

Como proposta para trabalhos futuros, os conjuntos de requisitos e das competências podem ser ampliados ou aprofundados, considerando áreas específicas. Promover o letramento de outros aspectos da gestão de dados, como a curadoria, a preservação a longo prazo, as implicações sob a ótica da proteção de dados pessoais, pode ser outro viés, principalmente, na atualidade em que legislações que regulam a privacidade se consolidam em vários países, como por exemplo o Regulamento Geral sobre a Proteção de Dados (GDPR) no âmbito Europeu e a Lei Geral de Proteção de Dados (LGPD), prevista para vigorar no Brasil em agosto de 2020.

REFERÊNCIAS

ANJOS, R. L. dos; DIAS, G. A. Atuação dos profissionais da informação no ciclo de vida dos dados - Dataone: um estudo comparado. **Inf. Inf.**, Londrina, v. 24, n. 1, p. 80 – 101, jan./abr. 2019.

AMERICAN LIBRARY ASSOCIATION (ALA). **Presidential Commission on Information Literacy: Final Report**. American Library Association: Chicago, IL, 1989. 6 p. Disponível em: <http://www.ala.org/acrl/publications/whitepapers/presidential>. Acesso em: 12 jun. 2020

ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES (ACRL). **Guidelines, standards and frameworks**. Association of College and Research Libraries: Chicago, IL. Disponível em <http://www.ala.org/acrl/standards/ilframework>. Acesso em: 21 jan. 2018.

BARDIN, L. **Análise de conteúdo**. Lisboa: edições 70, 1995.

BERGHMANS, S.; COUSIJN, H.; DEAKIN, G.; MEIJER, I.; MULLIGAN, A.; PLUME, A.; RIJCKE, S.; RUSHFORTH, A.; TATUM, C.; VAN LEEUWEN, T.; WALTMAN, L. **Open Data: the researcher perspective - survey and case studies, 2017**. Mendeley Data, versão 1, 4 abr. 2017. Disponível em: <https://data.mendeley.com/datasets/bwrnfb4bvh/1>. Acesso em: 25 abr. 2017. doi: 10.17632/bwrnfb4bvh.1

CALZADA PRADO, J.; MARZAL, M. Á. Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. **Libri: International Journal of Libraries & Information Services**, Berlin, v. 63, n.2, p. 123–34. 2013. Disponível em: <https://doi.org/10.1515/libri-2013-0010>. Acesso em: 02 jul. 2017.

CARLSON, J. R.; FOSMIRE, M.; MILLER, C.; NELSON, M. R. S. Determining Data Information Literacy Needs: A Study of Students and Research Faculty. **Libraries Faculty and Staff Scholarship and Research**, West Lafayette, v. 23, p. 1-30, 2011. Disponível em: http://docs.lib.purdue.edu/lib_fsdocs/23. Acesso em: 12 dez. 2017.

CARLSON, J. R.; JOHNSTON, L. R. **Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers**. Purdue Information Literacy Handbooks. West Lafayette, Indiana: Purdue University Press, 2015. 271 p. Disponível em: <https://goo.gl/riC4Es>. Acesso em: 12 dez. 2017.

CONSORTIUM OF EUROPEAN SOCIAL SCIENCE DATA ARCHIVES (CESSDA). Data in the social sciences. *In*: CESSDA DATA MANAGEMENT EXPERT GUIDE. Bergen, Norway: CESSDA ERIC, 2018. Disponível em: <https://www.cessda.eu/DMGuide>. Acesso em: 23 mar. 2019.

CORTI, L.; EYNDEN, V. V. D.; BISHOP, L.; WOOLLARD, M. **Managing and Sharing Research Data: a Guide to Good Practice**. 2 ed. Essex: UK Data

Service, 2014. Disponível em: <https://www.ukdataservice.ac.uk/manage-data/handbook>. Acesso em: 20 fev. 2019.

COURSERA. **Research Data Management and Sharing**. 2016. Disponível em: <https://www.coursera.org/learn/data-management>. Acesso em: 20 fev. 2019.

CREASER, C. Scholarly communication and access to research output. *In*: EVANS, W.; BAKER, D. **Libraries and Society**: role, responsibility and future in an age of change. Boston: EBSCO Publishing. 2011. p. 53-66.

CURTY, R. G.; CROWSTON, K.; SPECHT, A.; GRANT, B. W.; DALTON, E. D. Attitudes and norms affecting scientists' data reuse. **PloS one**, v. 12, n. 12, 2017. Disponível em: <https://doi.org/10.1371/journal.pone.0189288>. Acesso em: 02 jan. 2018.

DATAONE. Education Modules. 2012. Disponível em: <https://www.dataone.org/education-modules>. Acesso em: 02 jul. 2018.

DIGITAL CURATION CENTER. **DC 101 materials**: Access and Reuse. 2004-2019. Disponível em: <http://www.dcc.ac.uk/training/train-the-trainer/dc-101-training-materials>. Acesso em: 31 jan. 2019.

ENGINEERING AND PHYSICAL SCIENCES RESEARCH COUNCIL. **EPSRC policy framework on research data**. Disponível em: <https://epsrc.ukri.org/about/standards/researchdata/>. Acesso em: 12 jun. 2020.

EUROPEAN COMMISSION. **H2020 Programme. H2020 AGA – Annotated Model Grant Agreement**. Version 2.1.1. 1 jul. 2016. p. 216-219. Disponível em: <http://goo.gl/SXbKxE>. Acesso em: 20 nov. 2016.

FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH. Open Science Learning Objectives, 2015. FOSTER Consortium Members. Disponível em: <https://goo.gl/Lkm1XT>. Acesso em: 20 jul.2016. p. 536-552. doi: 10.5281/zenodo.15603.

FECHER, B.; FRIESIKE, S. Open Science: one term five schools of thoughts. *In*: BARTLING, S.; FRIESIKE, S. (Org). **Opening Science**. New York: Springer International Publishing, 2014, p. 17–47. Disponível em: <https://goo.gl/3qrg2o>. Acesso em 28 jan. 2017.

Foster (2015),

FREITAS, M. A. de; LEITE, F. C. L. Atores do sistema de comunicação científica: apontamentos para discussão de suas funções. **Inf. Inf.**, Londrina, v. 24, n. 1, p. 273 – 299, jan./abr. 2019.

GRAY, J.; LIU, D. T.; NIETO-SANTISTEBAN, M.; SZALAY, A.; DEWITT, D. J.; HEBER, G. Scientific data management in the coming decade. **ACM SIGMOD**

Record, [s.l.], v. 34, n. 4, p. 34-41, jan. 2005. Disponível em:
<https://arxiv.org/ftp/cs/papers/0502/0502008.pdf>. Acesso em: 23 jan. 2018.

GRIGG, K. S. Data in the Sciences. *In*: KELLAM, L. M.; THOMPSON, K. **Datalibrarianship: The Academic Data Librarian in Theory and Practice**. Association of College and Research Libraries. Chicago: American Library Association, 2016, p. 179-192.

GUERTIN, H.; BERNHARD, P. **Les 6 étapes d'un projet de recherche d'information**. École de bibliothéconomie et des sciences de l'information (EBSI), Université de Montréal, Québec, 2005. Disponível em:
<http://www.ebsi.umontreal.ca/jetrouve/projet/index.htm>. Acesso em: 02 jan. 2018.

JEFFRYES, J. N.; JOHNSTON, L. An E-Learning Approach to Data Information Literacy Education. *In*: Annual Conference & Exposition - ASEE, Atlanta, Georgia, **Anais [...]**, 2013. Disponível em: <https://peer.asee.org/19170>. Acesso em: 12 jun. 2020.

JOHNSTON, L. R., JEFFRYES, J. **Data Management Course**. University of Minnesota Libraries: Minnesota, 2012. Disponível em:
https://sites.google.com/a/umn.edu/data-management-course_structures/home-1. Acesso em: 02 jan. 2018.

KIM, Y. **Scientific Data Reuse Survey**. Inter-university Consortium for Political and Social Research, 2017. Disponível em: <http://doi.org/10.3886/E100404V1>. Acesso em: 25 abr. 2017.

KIM, Y. **Scientists' Data Sharing Behaviors**. Inter-university Consortium for Political and Social Research, 2016. Disponível em:
<http://doi.org/10.3886/E100087V7>. Acesso em: 25 abr. 2017.

KOLTAY, T. Are you ready? Tasks and roles for academic libraries in supporting Research 2.0. **New Library World**, [s.l.], v. 117, n. 1-2, p. 94-104, 2016. Disponível em: <https://doi.org/10.1108/NLW-09-2015-0062>. Acesso em: 12 jun. 2020

KOLTAY, T. Data literacy: in search of a name and identity. **Journal of Documentation**, [s.l.], v. 71, n. 2, p. 401-415, 2015. Disponível em:
<https://doi.org/10.1108/JD-02-2014-0026>. Acesso em: 12 jun. 2020.

KOLTAY, Tibor. HORSTMANN, Wolfram; WITT, Michael. Data governance, data literacy and the management of data quality. **IFLA Journal**, v.42, n. 4, p. 303-312, 2016. Disponível em: <https://doi.org/10.1177/0340035216672238>. Acesso em: 02 jan. 2018.

LAMAR SOUTTER LIBRARY. New England Collaborative Data Management Curriculum - Necdmc., University of Massachusetts Medical School. 2012.

Disponível em: <http://library.umassmed.edu/necdmc/index>. Acesso em: 02 jan. 2018.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. **Data on the Web best practices**. W3C Data on the Web Best Practices Working Group, World Wide Web Consortium (W3C), 31 jan. 2017. Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em: 30 jun. 2018.

MAYBEE, C.; CARLSON, J.; SLEBODNIKC, M.; CHAPMAN, B. "It's in the Syllabus": Identifying Information Literacy and Data Information Literacy Opportunities Using a Grounded Theory Approach. **The Journal of Academic Librarianship**, [s.l.], v. 41, n. 4, p. 369-376, Jul. 2015. Disponível em: <http://dx.doi.org/10.1016/j.acalib.2015.05.009>. Acesso em: 12 jun. 2020

NATIONAL SCIENCE FOUNDATION. **Dissemination and sharing of research results**. 2010. Disponível em: <https://goo.gl/E6nsXV>. Acesso em: 10 fev. 2018.

NVivo (2019).

OPEN ACCESS. Berlin Declaration on open access to knowledge in the sciences and humanities. Berlin, 2003. Disponível em: https://openaccess.mpg.de/67605/berlin_declaration_engl.pdf. Acesso em: 08 set. 2017.

PAGANI, R. N.; KOVALESKI, J. L.; RESENDE, L. M. Tics na composição da methodi ordinatio: construção de portfólio bibliográfico sobre modelos de transferência de tecnologia. **Ciência da Informação**, Brasília, v. 46, n. 2, 2017. Disponível em: <https://brapci.inf.br/index.php/res/v/18841>. Acesso em: 15 ago. 2019.

QIN, J.; D'IGNAZIO, J. Lessons learned from a two-year experience in science data literacy education. *In*: ANNUAL IATUL CONFERENCE, 31., **Anais** [...], 2010. Disponível em: <https://goo.gl/jpiJqN>. Acesso em: 19 jan. 2017.

RESEARCH DATA ALLIANCE. **About RDA**. Disponível em: <https://www.rd-alliance.org/node/51727>. Acesso em 03 dez. 2017.

RESEARCH DATA NETHERLANDS. **Essentials 4 Data Support**, [s.d.]. Disponível em: <http://datasupport.researchdata.nl/en/>. Acesso em: 20 fev. 2019.

SAYÃO, L. F.; SALES, L. F. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN, 2015. 90 p. Disponível em: <https://goo.gl/6JBkKV>. Acesso em 3 ago. 2017.

SCHNEIDER, R. Research data literacy. In KURBANOGLU, S.; GRASSIAN, E.; MIZRACHI, D.; CATTS, R.; SPIRANEC, S. (Eds.), **Worldwide commonalities and challenges in information literacy research and practice**. Cham: Springer International, v. 397, p. 134-140, 2013.

TENOPIR, C.; BIRCH, B.; ALLARD, S. **Academic libraries and research data services: current practices and plans for futur - An ACRL White paper**. ACRL, 2012. Disponível em: <https://goo.gl/eoQsXH>. Acesso em 20 out. 2017.

UNIVERSITY OF EDINBURGH. **Our definitions: Research data 2016**. Disponível em: <https://goo.gl/wznsyF>. Acesso em: 01 mar. 2017.

UNIVERSITY OF EDINBURGH. Research Data Service staff in Information Services. **Research Data and Management Training (MANTRA)** [online course]. Out. 2017. Disponível em: <https://mantra.edina.ac.uk/>. Acesso em: 20 fev. 2019.

UNIVERSITY OF OXFORD. **Workshops and Classes: Research Data Management handouts. 2018-2019**. Disponível em: <https://libguides.bodleian.ox.ac.uk/workshops/rdm>. Acesso em: 20 fev. 2019.

VITAE. **The Vitae Researcher Development Framework**. 2018, *online*. Disponível em: <https://goo.gl/XiXfVb>. Acesso em: 05 jan. 2018.

WILKINSON, M. D.; DUMONTIER, M.; AALBERSBERG, I. J.; APPLETON, G.; AXTON, M.; BAAK, A.; BLOMBERG, N.; BOITEN, J-W.; SANTOS, L. B. S.; BOURNE, P. E.; BOUWMAN, J.; BROOKES, A. J.; CLARK, T.; CROSAS, M.; DILLO, I.; DUMON, O.; EDMUNDS, S.; EVELO, C. T.; FINKERS, R.; GONZALEZ-BELTRAN, A.; GRAY, A. J. G.; GROTH, P.; GOBLE, C.; GRETHE, J. S.; HERINGA, J.; HOEN, P. A. C.; HOOFT, R.; KUHN, T.; KOK, R.; KOK, J.; LUSHER, S. J.; MARTONE, M. E.; MONS, A.; PACKER, A. L.; PERSSON, B.; ROCCA-SERRA, P.; ROOS, M.; VAN SCHAİK, R.; SANSONE, S-A.; SCHULTES, E.; SENGSTAG, T.; SLATER, T.; STRAWN, G.; SWEERTZ, M. A.; THOMPSON, M.; VAN DER LEI, J.; VAN MULLIGEN, E.; VELTEROP, J.; WAAGMEESTER, A.; WITTENBURG, P.; WOLSTENCROFT, K.; ZHAO, J.; MONS, B. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, New York, v. 3, n. 1, p. 1-9, 2016. Disponível em: <https://doi.org/10.1038/sdata.2016.18>. Acesso em: 12 jun. 2020.

WOOD, J. Coping with the Data Deluge. *In: JUBB, M.; SHORLEY, D. The Future of Scholarly Communication*. London: Facet, 2013. p. 75-87.

INFORMATION LITERACY FOR DATA REUSE IN THE SOCIAL SCIENCES: REQUIREMENTS AND COMPETENCES

ABSTRACT

Introduction: The scientific research environment has been impacted by the increase availability of research data in open access, however, researchers face difficulties in managing their own data as well as third parties'; especially in data reuse related matters. The relation of availability *versus* data management is more complex in the Social Sciences, whose data are heterogeneous and variable. **Objective:** Propose the

necessary requirements and minimum data literacy skills for research data reuse. **Methodology:** Based on a bibliographical background, it uses the NVivo as an analysis tool to support the content analysis techniques into three research instances. It compares (1) the selected conceptual frameworks of data literacy, it reuses (2) secondary datasets about researchers' behavior in reusing data, it identifies (3) in the Virtual Learning Environments what are the competencies for reuse purposes. **Results:** It presents the necessary requirements and competences set for research data reuse and test its validity by adherence checking with the best practices recommended by the World Wide Web Consortium (W3C). **Conclusions:** The proposal for 16 requirements for a virtual environment for literacy in the reuse of contextualized data in the area of Social Sciences and 37 competencies for researchers is consolidated. The main requirements are access, license and use, and essential competencies are the access rules, licenses types and how to use the available datasets, especially considering provenance, quality and reliability criteria.

Descriptors: Open Science. Research Data – reuse. Data Literacy.

CARTA DE INFORMACIÓN PARA LA REUTILIZACIÓN DE DATOS EN CIENCIAS SOCIALES: REQUISITOS Y COMPETENCIAS

RESUMEN

Introducción: El entorno de investigación científica se ha visto afectado por la mayor disponibilidad de datos de acceso abierto, sin embargo, surgen las dificultades de los investigadores para gestionar sus propios datos y los de terceros, especialmente en relación con los problemas de reutilización. Sin embargo, esta relación de disponibilidad versus gestión de datos es más compleja en las ciencias sociales, cuyos datos son heterogéneos y variables. **Objetivo:** Proponer los requisitos necesarios y las habilidades mínimas de alfabetización de datos para la reutilización de los datos. **Metodología:** Se basa en una encuesta bibliográfica, utilizando técnicas de análisis de contenido compatibles con la herramienta NVivo para el análisis en las tres instancias de investigación, (1) comparamos los marcos de alfabetización de datos, (2) reutilizamos conjuntos de datos de investigación secundaria e (3) identificamos en los entornos de aprendizaje virtual que son las habilidades para la reutilización de datos. **Resultados:** Se proponen los requisitos para la alfabetización de reutilización de datos y se prueba su validez de acuerdo con las mejores prácticas recomendadas por el Consorcio World Wide Web (W3C). **Conclusiones:** Se consolida la propuesta de 16 requisitos para un entorno virtual para la alfabetización en la reutilización de datos contextualizados en el área de Ciencias Sociales y 37 competencias para investigadores. Los requisitos principales son el acceso, la licencia y el uso, y las competencias esenciales son las reglas de acceso, los tipos de licencias y cómo usar los conjuntos de datos disponibles, especialmente los criterios de procedencia, calidad y confiabilidad.

Descriptor: Ciencia abierta. Datos de investigación - Reutilización. Alfabetización en datos.

Recebido em: 27.01.2020

Aceito em: 10.06.2020