

Licenças de uso de dados de pesquisa: um estudo em políticas de repositórios

Jaqueline Insaurriaga Silveira¹; Samile Andréa de Souza Vanz²

RESUMO

Este estudo apresenta uma análise de políticas de gestão de repositórios para a atribuição de licenças de uso de dados de pesquisa. Metodologicamente, trata-se de um estudo descritivo realizado por meio de uma pesquisa documental em políticas de 18 repositórios universitários internacionais. A análise permitiu identificar nas políticas, a presença de diretrizes para a atribuição de licenças de uso de dados de pesquisa. Os resultados mostram a indicação de uso das licenças Creative Commons, Open Data Commons, AusGoal e licenças específicas dos repositórios. A análise revelou especificidades quanto a atribuição das licenças de uso e que esta, é uma prerrogativa dos depositantes em conformidade com as diretrizes indicadas nas políticas dos repositórios.

Palavras-chave: Dados de pesquisa; Licenças de uso; políticas de repositórios.

INTRODUÇÃO

O presente estudo é o recorte de um projeto de pesquisa maior, onde analisaram-se distintos aspectos presentes em políticas de gestão de 18 repositórios de dados de pesquisa de universidades internacionais. Este recorte aborda um dos seus objetivos específicos, as diretrizes relativas à atribuição de licenças de uso de dados de pesquisa, presentes nas políticas de repositórios.

A mudança de patamar dos dados de pesquisa, de meros subprodutos das atividades científicas, para o de recursos informacionais de valor, como protagonistas de uma ciência mais colaborativa, preocupada com a abertura, compartilhamento e reuso desses dados (MONTEIRO; LUCAS, 2019; SAYÃO; SALES, 2016), exige um olhar atento a todos os aspectos que envolvem seu acesso, compartilhamento e reuso.

¹Mestranda em Ciência da Informação pelo PPGCIN/UFRGS ; UFRGS; jaqjis@gmail.com

²Professor do Departamento de Ciência da Informação; FABICO/UFRGS; samilevanz@gmail.com

Embora o compartilhamento e reuso de dados de pesquisa esteja presente na agenda dos eventos científicos, de fato, essa é uma prática recente entre os pesquisadores brasileiros. Em estudo realizado sobre as práticas e percepções dos usuários nacionais de serviços de acesso aberto a dados de pesquisa, Vanz et al. (2018) constataram que, embora haja grande interesse pela temática, dentre os indivíduos pesquisados, ainda há equívocos relacionados ao entendimento do que significa compartilhamento e reuso de dados de pesquisa. O estudo também revela incipiência em relação à prática de compartilhamento e reuso de dados entre os pesquisadores brasileiros investigados.

Nas universidades e centros de pesquisa, de acordo com Santa Anna, Dias e Maculan (2019), os dados de pesquisa representam um dos principais produtos das investigações acadêmicas, sendo esta, uma das justificativas para a demanda existente entre as universidades brasileiras, pela implementação de seus repositórios de dados de pesquisa. Contudo, embora haja essa demanda das universidades brasileiras, as iniciativas de implantação de repositórios de dados de pesquisa, ainda são recentes. Fato que levou a busca por analisar as políticas de repositórios de universidades internacionais, visto já ser uma prática consolidada nessas instituições, desde o ano de 2002.

Em meio aos distintos aspectos que envolvem a abertura de dados de pesquisa, está o seu depósito e armazenamento em repositórios. Dentre os muitos desafios que o armazenamento e o processamento de dados de pesquisa apresentam para as instituições, estão: sua fragilidade, vulnerabilidade, obsolescência tecnológica, falhas de armazenamento, questões éticas e legais. Nesse contexto, os repositórios são os sistemas de informação adequados, visto contar com infraestruturas tecnológica e gerencial apropriadas ao armazenamento, gestão e preservação dos dados de pesquisa.

Relativo aos desafios legais, encontram-se também, os aspectos referentes ao uso de dados de pesquisa por terceiros e, aos direitos autorais do seu criador/pesquisador. Para disciplinar essa relação entre autor e usuário final, tem-se as licenças abertas de uso, as quais, segundo Monteiro (2017), proporcionam novas possibilidades de reuso dos dados de pesquisa, uma vez que possibilitam um equilíbrio jurídico entre os autores detentores dos direitos autorais e o usuário final.

As licenças abertas trazem clareza aos usuários para a reutilização de conjuntos de dados e, asseguram que os dados sejam publicados, sem ferir os direitos autorais.

Cabe ao criador dos dados de pesquisa a atribuição da licença de uso que considere adequada, considerando à finalidade do compartilhamento do seu trabalho e das expectativas que tem em relação ao seu uso por terceiros (OPEN KNOWLEDGE PORTUGAL, [20--]). Considera-se três abordagens possíveis para licenciamento aberto de dados: as licenças *copyleft*, que permitem atribuir e partilhar dados nos mesmos termos; as licenças *copyleftfree*, estipulam apenas a obrigação de atribuição da autoria original; e as licenças domínio público, em que os criadores abdicam de seus direitos autorais (OPEN KNOWLEDGE PORTUGAL, [20--]).

Dentre essas três abordagens possíveis para licença aberta de dados tem-se: as **Creative Commons (CC)**, que oferece um conjunto de seis tipos de diferentes licenças, além da disponibilização aos usuários de uma ferramenta pública de dedicação ao domínio público (CC0); as **Open Data Commons (ODC)**, que disponibiliza três licenças, dentre elas igualmente uma de domínio público (PDDL).

Há também, a licença Australian Governments Open Access and Licensing Framework (AusGOAL), lançada em 2011, a qual visando facilitar o acesso aberto a informações financiadas com recursos públicos, fornece apoio e orientação ao governo e setores relacionados. (WIKI AUSTRÁLIA AUSGOAL, 2015).

Diante do exposto, buscou-se conhecer práticas de acesso aberto a dados de pesquisa em repositórios de universidades internacionais, com o objetivo de verificar a presença ou não, de diretrizes em suas políticas de gestão, sobre a atribuição de licenças de uso de dados de pesquisa.

A seção a seguir apresenta os procedimentos metodológicos usados na pesquisa, descrevendo o método de escolha para identificação dos repositórios pesquisados. Posteriormente, apresentam-se algumas considerações amparadas pelas análises realizadas nas políticas dos repositórios.

METODOLOGIA

A investigação constitui-se num estudo descritivo, onde realizou-se uma pesquisa documental nas políticas de repositórios de dados de pesquisa, de

universidades internacionais. O *corpus* analisado é composto pelas políticas de gestão de 18 repositórios, escolhidos por meio da identificação das 18 primeiras melhores universidades mundiais elencadas no *Ranking Webometrics.info* de 2020, a saber as universidades e seus repositórios: **Harvard University** (Harvard Dataverse); **Stanford University** (Stanford Digital Repository (SDR)); **Massachusetts Institute of Technology** (DSpace@MIT); **University of California Berkeley** (Data Repository [Berkeley]); **University of Washington** (Research Works Archive); **University of Michigan** (Deep Blue Data); **University of Oxford** (ORA-Data); **Johns Hopkins University** (JHU Data Archive); **Columbia University New York** (Columbia Academic Commons); **Cornell University** (eCommons); **University of Pennsylvania** (Scholarly Commons); **University of Cambridge** (Apollo); **University of California Los Angeles UCLA** (Merritt); **Yale University** (ISPS Data Archive); **University of Wisconsin Madison** (MINDS@UW); **University of California San Diego** (UC San Diego Library Digital Collections); **University of Minnesota System** (DRUM); **University College London** (UCL Research Data Repository (RDR)).

A partir da escolha das universidades no *Ranking Webometrics.info*, procurou-se conhecer seus respectivos repositórios de dados de pesquisa, por busca em seus sites. Quando a localização dos repositórios não foi possível pelos sites das universidades, então efetuou-se pesquisa direto no Google, utilizando-se os termos de busca: “repository + of + nome da universidade”, ou, “nome da universidade + repositório”. A coleta de dados sobre as políticas de gestão dos repositórios foi realizada nos sites, de forma a conhecer as diretrizes para atribuição de licenças de uso de dados de pesquisa.

A identificação dos repositórios e a coleta de dados referente as suas políticas de gestão foram realizadas de 16 de maio a 23 de setembro de 2020.

Os dados, relativos as licenças de uso, foram primeiramente compilados em um documento único, incorporando todas as informações presentes nas políticas dos repositórios. Para a análise das informações, foram elaborados quadros no Microsoft Word, estruturados de forma a identificar o nome dos repositórios e suas respectivas diretrizes para atribuição das licenças de uso.

RESULTADOS

Conforme Monteiro (2017), a atribuição de uma licença consiste na transferência do direito de exploração sobre uma obra, podendo este ser exclusivo ou não exclusivo. Quanto aos dados de pesquisa, a atribuição de licenças, representa o estabelecimento de uma espécie de contrato, entre o autor desses dados e seus usuários, que explicita suas permissões de uso.

Sendo assim, a análise das políticas revelou a presença desse 'contrato de uso de dados de pesquisa', em todos os 18 repositórios estudados. Observou-se a indicação de atribuição das licenças Creative Commons, Open Data Commons, AusGoal, ou licenças específicas do próprio repositório.

A atribuição das licenças de uso é uma prerrogativa dos depositantes em conformidade com as diretrizes indicadas nas políticas do repositório. As análises indicam a presença de especificidades quanto à atribuição das licenças e aos contratos estabelecidos entre depositante e repositório/universidade, depositante e usuário/*downloader* para o uso dos dados.

Embora possa haver licenciamentos sobre outros termos, por padrão, repositórios como o **Harvard Dataverse**, o **Data Repository [Berkeley]**, o **JHU Data Archive**, o **Merritt** e o **UCL Research Data Repository** indicam a atribuição CC0 para seus datasets, enquanto o **Deep Blue Data** a indica somente para seus metadados. Já o **Oxford University Research Archive – ORA**, para dados que não sejam de propriedade da Universidade, exige que o depositante conceda à Oxford University uma licença mundial não exclusiva para disponibilização dos dados no Repositório, não especificando se será uma atribuição Creative Commons ou Open Data Commons. Não se localizaram referências quanto ao licenciamento para uso de dados, nas políticas do **UC San Diego**, contudo, em contato com a Instituição, obteve-se a informação de que todos os registros de dados vêm acompanhados de uma atribuição Creative Commons. Já o **DRUM** indica que, o licenciamento se dará no momento do depósito de dados pelos autores, que poderão optar pela aplicação de uma licença Creative Commons. Indica ainda que, mesmo não sendo atribuída uma CC, os usuários poderão baixar e usar os dados de pesquisa. Os repositórios **Stanford Digital Repository – SDR**, **MINDS@UW** e **UCL Research Data Repository** indicam a atribuição de licenças tanto Creative Commons como Open

Data Commons, sendo que o UCL Research Data Repository também sinaliza a utilização de licenças AusGOAL.

Sendo assim, no universo das políticas analisadas, 77% dos repositórios indicam o uso das licenças Creative Commons, 14% das Open Data Commons, 4% sinalizam o uso das licenças AusGOAL e, 5% não indicam a atribuição de uma específica licença de uso. Embora, a atribuição CC0 para os *datasets* esteja presente na maioria das políticas, identificou-se sua indicação somente para os metadados nas políticas de um dos repositórios analisados.

A incidência maior de atribuição das Creative Commons, em 77% dos repositórios analisados, pode ser justificada, considerando que as CC desde sua fundação no ano de 2001, permanecem em constante utilização, além de fornecerem diferentes tipos de licenças de uso. Conforme Lemos (2005), a Creative Commons cria instrumentos jurídicos que permitem ao detentor dos direitos legais de uma obra, explicitar de forma clara e precisa, que sua obra é livre para ser distribuída, copiada e utilizada. Ainda, segundo Lemos (2005), as CC são fundamentadas na prerrogativa que cada indivíduo tem, como autor, o direito de permitir o acesso as suas obras e a seus trabalhos, autorizando que outros os utilize e façam criações sobre estes.

Pode-se afirmar que as CC têm uma apresentação visual de fácil compreensão e memorização para seu uso, constituindo-se na melhor solução prática de licenciamento (SWAN, 2016). Além de ser um instrumento de fácil compreensão e fornecerem um pacote de licenças abrangentes para as necessidades dos autores, as CC podem ainda, ser interpretadas por máquinas.

Seguindo a Creative Commons, as licenças Open Data Commons são a segunda mais indicada nas políticas analisadas. Embora disponibilize apenas três tipos de licenças, a Open Data Commons apresenta similaridades com a CC e, segundo Silva (2019), juntamente com esta, são as licenças mais utilizadas para dados abertos.

Somente um dos repositórios indica em suas políticas a atribuição da licença AusGoal, o UCL Research Data Repository. E esta indicação se dá por meio das diretrizes do Figshare para o Repositório, guiando a escolha de licenças de uso.

CONSIDERAÇÕES FINAIS

O estudo abrangeu práticas de acesso aberto em repositórios universitários internacionais, com foco na análise de suas políticas de gestão, para atribuição de licenças de uso de dados de pesquisa. Pode-se identificar nos repositórios analisados a atribuição de licenças para uso dos dados de pesquisa, sejam elas apresentadas em suas políticas ou na recuperação dos dados na plataforma dos repositórios.

A análise mostrou que mesmo que possa haver, nas políticas dos repositórios, diretrizes para licenciamentos sobre outros termos, por padrão, o maior percentual de atribuição de licenças recai para o uso das Creative Commons, seguido em segundo lugar pelas licenças Open Data Commons. Em menor número, representando somente a utilização por um dos 18 repositórios analisados, está a atribuição das licenças AusGOAL. Por fim, alguns não indicam em suas políticas o uso de uma licença específica.

REFERÊNCIAS

LEMOS, Ronaldo. **Direito, tecnologia e cultura**. [Rio de Janeiro: FGV, 2005].

Disponível em:

<http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/2190/Ronaldo%20Lemos-Direito-Tecnologia-e-Cultura.pdf?sequence=1> Acesso em: 21 jan. 2021.

MONTEIRO, Elizabete Cristina de Souza de Aguiar. **Direitos autorais nos repositórios de dados científicos**: análise sobre os Planos de Gerenciamento dos Dados. 2017. Dissertação (Mestrado em Informação e Tecnologia) – Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2017. Disponível em:

<https://repositorio.unesp.br/handle/11449/149748> . Acesso em: 23 maio 2020.

MONTEIRO, Gabriela; LUCAS, Elaine Rosângela de Oliveira. Dados científicos abertos: identificando o papel das políticas de gestão e das agências de fomento.

AtoZ: novas práticas em informação e conhecimento, Curitiba, v. 8, n. 1, p. 13-20, jan./jun. 2019. Disponível em: <https://revistas.ufpr.br/atoz/article/view/67253/38725> . Acesso em: 14 mar. 2020.

OPEN KNOWLEDGE PORTUGAL. Dados abertos: [licenças]. [20--]. Disponível em: <http://dadosabertos.pt/licencas.php> . Acesso em: 19 jun. 2020.

SANTA ANNA, Jorge; DIAS, Célia da Consolação; MACULAN, Benildes Coura. A gestão dos dados de pesquisa nas universidades e o papel dos serviços informacionais oferecidos nas bibliotecas: uma revisão narrativa. **Múltiplos olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, p. 1-16, 2019. Disponível em:

<https://periodicos.ufmg.br/index.php/moci/article/view/19169> . Acesso em: 21 maio 2020.

SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 1, n. 21, p. 90-115, maio/ago., 2016. Disponível em:

<http://www.uel.br/revistas/uel/index.php/informacao/article/download/27939/20122>

Acesso em: 09 nov. 2019.

SILVA, Fabiano Couto Corrêa da. **Gestão de dados científicos**. Rio de Janeiro: Interciência, 2019. 128 p.

SWAN, Alma. **Diretrizes para as Políticas de desenvolvimento e promoção do Acesso Aberto**. Brasília: Unesco, 2016. Disponível em:

https://unesdoc.unesco.org/ark:/48223/pf0000246018_por Acesso em: 21 jan. 2021.

VANZ, Samile Andrea de Souza et al. Acesso aberto a dados de pesquisa no Brasil: prática e percepções dos pesquisadores: relatório 2018. Porto Alegre, RS: UFRGS, 2018. Disponível em: <http://hdl.handle.net/10183/185195> .Acesso em: 05 mar. 2020.

WIKI AUSTRÁLIA AUSGOAL. **Austrália**: AUSGoal - Governos Australianos abrem o quadro de acesso e licenciamento. 2015. Disponível em:

[https://wiki.creativecommons.org/wiki/Australia:_AUSGoal_-](https://wiki.creativecommons.org/wiki/Australia:_AUSGoal_-_Australian_Governments_Open_Access_and_Licensing_Framework)

[_Australian_Governments_Open_Access_and_Licensing_Framework](https://wiki.creativecommons.org/wiki/Australia:_AUSGoal_-_Australian_Governments_Open_Access_and_Licensing_Framework) Acesso em: 21 jan. 2021.