

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-5 – Política e Economia da Informação

DIREITOS AUTORAIS DE DADOS CIENTÍFICOS NO CONTEXTO DA CIÊNCIA ABERTA: ESTUDO DO REPOSITÓRIO DE DADOS DO CONSÓRCIO MADROÑO

SCIENTIFIC DATA COPYRIGHT IN THE CONTEXT OF OPEN SCIENCE: STUDY OF THE CONSÓRCIO MADROÑO'S DATA REPOSITORY

Elizabete Cristina de Souza de Aguiar Monteiro – Universidade Estadual Paulista (Marília)

Ricardo Cesar Gonçalves Sant'Ana – Universidade Estadual Paulista (Marília)

Antonio Hernández-Pérez – Universidad Carlos III de Madrid

Modalidade: Trabalho Completo

Resumo: No contexto da Ciência Aberta estão inseridos novos fatores em sua agenda como a gestão de dados científicos e os princípios FAIR (*Findable, Accessible, Interoperable, Reusable*) onde são estabelecidos subprincípios em relação aos dados e aos metadados. A elaboração de Planos de gestão de dados pelos pesquisadores documentando todo o processo de gestão dos dados contribuem para aplicação dos princípios FAIR e fornecem *insights* para a disseminação e reutilização dos dados. O objetivo desse artigo é identificar as convergências e divergências entre a operacionalização dos repositórios de dados que compõem o Consórcio Madroño (*e-cienciaDatos*) no aspecto dos direitos autorais dos dados neles depositados com a aplicação dos princípios da Ciência Aberta e FAIR. Foi utilizada a pesquisa documental e exploratória e uso do método estudo de caso e análise de conteúdo. O universo da pesquisa foi composto pelos bibliotecários que gerenciam os dados e ministram treinamentos e pelas políticas do repositório. A coleta ocorreu por meio da triangulação com a aplicação de questionário, entrevista estruturada e análise documental. Os resultados demonstram que há convergências entre a Ciência Aberta e as boas práticas dos princípios FAIR com a operacionalização do repositório. Conclui-se que os conjuntos de dados são protegidos pela propriedade intelectual e o *e-cienciaDatos* subscreve seu compromisso com a Ciência Aberta demonstrando a relevância dos direitos de propriedade intelectual para os atores envolvidos destacando a integridade e a ética na disponibilização e reuso dos dados.

Palavras-chave: Repositórios de dados; Direitos autorais; Ciência aberta; Princípios FAIR; Plano de Gestão de Dados.

Abstract: The aim of this paper is to identify the convergences and divergences between the data repositories' operationalization which compose the Consórcio Madroño (*e-cienciaDatos*) in the data's copyright aspect deposited in them with the application of Open Science and FAIR principles. Documentary and exploratory research and the use of the case study method were used. The research's universe consisted of librarians who manage the data and provide training and by repository's policies. The collection took place through triangulation with the application of a questionnaire, structured interview and document analysis. The results show that there is convergence between Open Science and the good practices of FAIR principles with the operationalization of the repository. It is concluded that *e-cienciaDatos* subscribes to its commitment to Open Science by demonstrating the relevance of intellectual property rights to the actors involved, highlighting the integrity and ethics in the availability and reuse of data.

Keywords: Data repository; Copyright; Open Science; FAIR principles; Data management plan.

1 INTRODUÇÃO

A Ciência Aberta é um movimento internacional de prática científica com a disponibilização das informações científicas e, atualmente, de dados científicos, das metodologias e dos processos de maneira aberta, proporcionando ampla divulgação à sociedade e potencialização da reprodutibilidade da ciência (MORA, 2009; BEZJAK, 2018).

Os fatores que estão adjacentes à Ciência Aberta como publicação aberta, metadados abertos, dados abertos, código aberto, recursos educativos abertos, revisão por pares aberta, impacto e métricas abertas, acesso aberto à comunicação científica e aos dados científicos, os repositórios abertos, incluindo repositórios de dados, e os princípios FAIR estão em discussão em nível nacional e internacional incorporando novos elementos a sua agenda (CONSÓRCIO MADROÑO, 2016, 2017; MONTEIRO; SANT’ANA, 2019, no prelo).

Financiadores de pesquisas como a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) no Brasil e internacionais como, *National Science Foundation* (NSF), *National Institutes of Health* (NIH), NASA dos Estados Unidos, *Wellcome Trust*, Fundação Bill e Melinda Gates e *Horizon 2020* na União Européia estão, gradativamente, instituindo em suas políticas de financiamento a orientação para publicação dos dados científicos em periódicos ou repositórios de dados, a elaboração de Plano de gestão de dados (PGD), o estabelecimento de políticas de acesso aberto e a articulação para que a gestão de dados científicos possam atender aos princípios conhecidos como FAIR (*Findable, Accessible, Interoperable, Reusable*) (BORGMAN et al., 2015; COUTO CORRÊA, 2016; FAPESP, 2017; EUROPEAN COMMISSION, 2018).

Dados científicos são heterogêneos e diferentes nas diversas disciplinas. São os materiais produzidos ou coletados durante a condução de uma pesquisa, de origens computacionais, experimentais ou observacionais, geralmente aceitos na comunidade científica como necessários para validar os resultados da pesquisa (NATIONAL SCIENCE BOARD, 2005; ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, 2007).

No contexto apresentado os repositórios de dados e os bibliotecários que, por competências inerentes a sua atuação, adquirem papel estratégico na gestão dos dados, na aplicação dos princípios FAIR e dos princípios da Ciência Aberta e no apoio técnico aos pesquisadores. A biblioteca tem papel de destaque no apoio à gestão de dados e, na Espanha, tem notável consolidação (ABADAL, et al., 2018).

Nessa dinâmica os aspectos legais e éticos relacionados à propriedade intelectual dos conjuntos de dados científicos devem ser considerados. A propriedade intelectual é a área do direito que protege, por meio de leis, os inventores ou responsáveis por qualquer produção intelectual, seja no domínio industrial, científico, literário ou artístico, facultando aos seus titulares direitos econômicos os quais ditam a forma de comercialização, circulação, utilização e produção (UNIVERSIDAD CARLOS III DE MADRI (UC3), [201-a]).

A propriedade intelectual possui duas vertentes: Propriedade industrial e Direitos autorais, sendo esse último foco deste artigo. Direito autoral é um dos ramos da ciência jurídica e permeia um conjunto de direitos morais (reconhecimento da autoria explicitado nas citações e o direito à integridade) e os patrimoniais (econômicos e de exploração) sobre as criações do espírito, expressas por quaisquer meios ou fixadas em quaisquer suportes, tangíveis ou intangíveis, que se concede aos criadores de obras intelectuais (UNIVERSIDAD CARLOS III DE MADRI (UC3), [201-a]).

A prática de citação dos dados promove o reconhecimento dos pesquisadores assim como a citação de suas publicações tradicionais. Os princípios e práticas de citação de dados estão em discussão juntamente com a necessidade de orientação na reutilização e replicação de dados.

Relevante destacar que nem todos os dados abertos, especialmente dados de pesquisas, são ou devem ser completamente "abertos" ou "livres", mas devem ser, pelo menos, FAIR conforme a redação do artigo 10.1 do *FAIR4Health Consortium* (2019) "*as open as possible, as closed as necessary*". Existem dois motivos principais para manter os dados da pesquisa "fechados": 1) Razões legais e éticas; 2) Razões econômicas (HERNÁNDEZ-PÉREZ; MÉNDEZ RODRÍGUEZ, 2019).

Dessa forma, o tema deste estudo é Direitos autorais dos conjuntos de dados científicos nos Repositórios de dados que compõem o repositório *e-cienciaDatos* do Consórcio Madroño. O objetivo desse artigo é identificar as convergências e divergências entre a operacionalização do *e-cienciaDatos* no fator direitos autorais dos conjuntos dos dados com a aplicação dos princípios FAIR e princípios da Ciência Aberta.

2 DIREITOS AUTORAIS NO CONTEXTO DA CIÊNCIA ABERTA

A propriedade intelectual é distinguida entre Direitos Morais e Direitos Patrimoniais. O direito moral é o reconhecimento pela autoria com sua explicitação nas citações. A *Joint*

Declaration of Data Citation Principles – Final declaração do Grupo de trabalho do *The Future of Research Communications and e-Scholarship* (FORCE11) destaca que a citação de dados se configura em boas práticas de pesquisa (MARTONE, 2014). Dentre seus princípios destaca-se, “As citações de dados devem facilitar a concessão de crédito acadêmico e a atribuição normativa e legal a todos os autores dos dados.” (MARTONE, 2014, tradução nossa).

Os direitos patrimoniais regulamentam as relações de uso com aplicação das licenças de uso, as quais são utilizadas com o intuito de publicar, reutilizar e replicar os dados. Para a reutilização e replicação dos dados está sendo indicado o uso de licenças como a *Creative Commons* e *Open Data Commons* para os *datasets* e licenças para *software*. Uma licença deixa claro todas as liberações e restrições no uso dos dados armazenados nos repositórios de dados e, um de seus requisitos documentado no PGD, é a importância do reconhecimento do autor ou detentor dos dados. Licenças como a *Open Data Commons* e a *Creative Commons* (CC) estão sendo utilizadas pelos repositórios de dados.

Na Espanha, a norma principal que orienta as questões de direitos autorais é a *Ley de Propiedad Intelectual* (LPI)¹ que sofreu várias modificações, dentre elas, pela Lei 23/2006 de 7 de julho, a fim de adaptar os regulamentos espanhóis às novas circunstâncias criadas pela sociedade da informação (UC3M, [201-b]). Em ambientes digitais e na Internet, além da LPI, é preciso considerar outras normas como as que regulam a prestação de serviços na sociedade da informação, sendo adaptada para ambientes digitais (UC3M, [201-b]). Em 2019, foi publicada a *Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE*.

O Grupo de Trabalho de *Depósito y Gestión de datos en Acceso Abierto* do projeto RECOLECTA elaborou o *Marco legal de los datos de investigación* sobre a produção, gestão e reutilização de dados científicos na Espanha. Pela previsão da legislação, das Diretivas e o Marco legal:

As coleções de dados e os bancos de dados são protegidos por propriedade intelectual, de acordo com o art. 12 da TRLPI² mencionados por meio do chamado direito *sui generis*, na medida em que constituem criações intelectuais. “A proteção se refere apenas à sua estrutura como uma forma

¹ Texto Consolidado aprovado pelo Real Decreto Legislativo 1/1996 de 12 de abril

² Decreto Legislativo Real 1/1996, de 12 de abril, que aprova o texto consolidado da Lei de Propriedade Intelectual, regularizando, esclarecendo e harmonizando as disposições legais vigentes sobre o assunto.

de expressão da seleção ou organização de conteúdo ", não os dados em si (UC3M, [201-c]).

Os dados científicos são os elementos chave no contexto atual da Ciência Aberta (ABADAL et al., 2018). São heterogêneos e diferentes nas diversas disciplinas. Dados podem estar em estado bruto ou analisado, ser numéricos, descritivos ou visuais, ser experimentais ou observacionais, ser fatos, observações ou experiências nos quais o argumento, teoria ou prova é baseado (UNIVERSITY OF MELBOURNE, 2009).

Os dados incluem, mas não se limitam a

cadernos de laboratório, cadernos de campo, dados primários de pesquisa (incluindo dados em papel ou em suporte de computador), questionários, fitas de áudio, vídeos, desenvolvimento de modelos, fotografias, filmes e verificações e respostas do teste. [...] As informações sobre a proveniência dos dados também podem incluir: como, quando, onde foram coletadas e com que (por exemplo, instrumentos). O código de software usado para gerar, comentar ou analisar os dados também pode ser considerado como dados [...] Os dados podem ser efêmeros e multifacetados e devem, portanto, ser considerados como o conjunto de possibilidades emergentes, materialidades e idéias em constante mudança e constantemente reavaliados que encontram incorporação na obra de arte final. (UNIVERSITY OF MELBOURNE, 2009, tradução nossa).

A Ciência Aberta é um conceito complexo que exige adoção de novas práticas e mudança no comportamento dos pesquisadores em relação à realização de pesquisas e compartilhamento de seus resultados, incluindo os dados (PONTIKA et al., 2015).

Para a Comissão Europeia, "*Open Science*" é a transformação, a abertura e a democratização da ciência, da pesquisa e da inovação, com o objetivo de tornar a ciência mais eficiente, transparente e interdisciplinar, de mudar a interação entre ciência e sociedade e de possibilitar a ampliação do impacto social e da inovação (RAMJOUÉ, 2015).

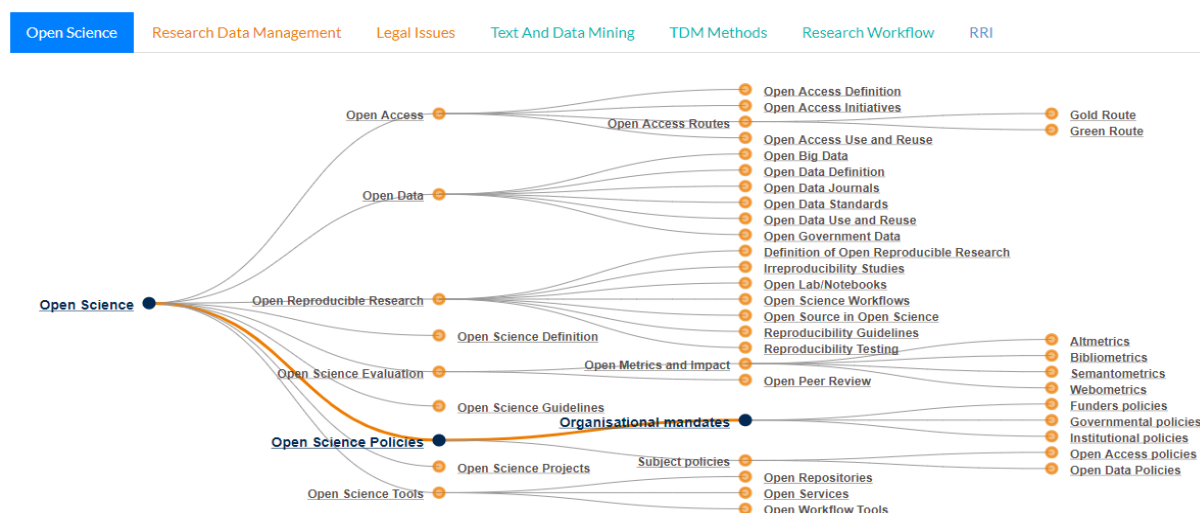
Os principais elementos da Ciência Aberta são apresentados no trabalho de Pontika et al. (2015, grifo nosso):

- **Open access:** resultados das pesquisas acadêmicas avaliados pelos pares, acessíveis, sem restrições para a leitura, reutilizáveis, **com direitos autorais e licenças** limitados ou nulos (SUBER, 2012);
- **Open data:** publicação *on-line* disponível para acesso e reutilização dos dados da pesquisa coletados ou gerados ao longo do projeto (MURRAY-RUST, 2008);
- **Open source:** *software* com acesso gratuito, **com uma licença de código fonte** que permite seu uso, criação de derivados e distribuição (ALTENHÖREN, 2005);

- **Open reproducible research:** a prática da Ciência Aberta que permite a reprodutibilidade dos resultados da pesquisa (STODDEN, 2009).

A Ciência Aberta engloba uma série de práticas e estratégias abertas que envolvem conceitos abertos como acesso aberto às publicações, dados de investigação abertos, *software* e ferramentas de código aberto, fluxos de trabalho abertos, ciência cidadã, recursos educacionais, métodos abertos e alternativos para apoiar a pesquisa como a revisão por pares de forma aberta (PONTIKA et al., 2015) conforme ilustrado na figura 1 seguindo a *Open Science Taxonomy*. Evidencia-se na figura 1 que há indicadas políticas da Ciência Aberta que englobam as políticas dos órgãos financiadores, políticas governamentais e políticas institucionais que, nesse contexto, tem forte contribuição.

Figura 1: Políticas na Open Science



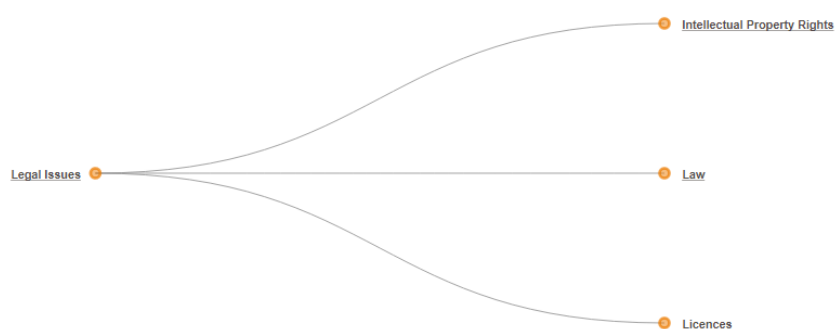
Fonte: Extraído de FOSTER, [201-].

Destaca-se que em setembro de 2018, a comunidade científica europeia presenciou uma mudança potencialmente transformadora na política de pesquisa. O *Plano S*, lançado pela *Science Europe*³, destina-se a aumentar o acesso aberto (OA) à publicação acadêmica e aos dados científicos resultantes de pesquisas produzidas com financiamento público até 01 de janeiro de 2020 (QUADERI, 2018). Após consulta à comunidade de pesquisa e de revisão de suas diretrizes, o prazo foi estendido até janeiro de 2021 (PELLS, 2019). O *Plano S* requer que os resultados de pesquisas financiadas por organizações signatárias sejam publicados em repositórios abertos ou em periódicos de acesso aberto para sua gestão e disponibilização (QUADERI, 2018).

³ <https://www.scienceeurope.org/>

Oportuno enfatizar que na *Open Science Taxonomy*, na *prática Open Data*, está elencado o *Open Data use and Reuse* (FIGURA 1). A prática “openness” não implica que os dados publicados abertamente não devam ser citados e ter seus autores reconhecidos ou liberados sem indicações de licenças de uso. Questões legais também são apresentadas na *Open Science Taxonomy* enfatizando a propriedade intelectual dos dados (FIGURA 2). Um grande e imprescindível componente do sistema operacional está relacionado com as questões legais por trás do compartilhamento, reutilização e reprodução de texto, dados, metodologias de pesquisa, práticas e experimentos (PONTIKA et al., 2015).

Figura 2: Aspectos legais direcionados à dados



Fonte: Extraído de FOSTER, [201-].

A gestão de dados científicos, no contexto de pesquisas financiadas com recursos públicos, está sendo articulada para atender aos princípios conhecidos como FAIR (*Findable, Accessible, Interoperable, Reusable*) divulgados nas comunidades científicas em todo o mundo (QUADRO 1). O FAIR exige grandes mudanças em termos de cultura e prática de pesquisa e a implementação de um ecossistema com serviços e componentes de dados como Políticas, PGDs, identificadores, padrões e repositórios sendo, esse último, essencial no ecossistema de dados FAIR pois são necessários para executar a função de oferecer dados e metadados acessíveis e reutilizáveis para usuários e por máquinas (EUROPEAN COMMISSION, 2018).

Quadro 1: Princípios FAIR

Princípio	Descrição
Findable	F1- (meta) dados são atribuídos a um identificador globalmente exclusivo e persistente
	F2 - os dados são descritos com metadados ricos (definidos por R1 no princípio Reusable)
	F3 - metadados de forma clara e explícita incluem o identificador dos dados que descreve
	F4 - (meta) dados são registrados ou indexados em um recurso pesquisável
Accessible	A1 - (meta) dados são recuperáveis pelo seu identificador usando um protocolo de comunicação padronizado
	A1.1 - o protocolo é aberto, gratuito e universalmente implementável

	A1.2 - o protocolo permite um procedimento de autenticação e autorização, quando necessário
	A2 - os metadados estão acessíveis, mesmo quando os dados não estão mais disponíveis
Interoperable	I1 - Os (meta) dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.
	I2 - (meta) usam vocabulários que seguem os princípios do FAIR
	I3 - Os (meta) dados incluem referências qualificadas a outros (meta) dados
Reusable	R1 meta (data) são ricamente descritos com uma pluralidade de atributos precisos e relevantes
	R1.1 - (meta) dados são liberados com uma licença de uso de dados clara e acessível
	R1.2 - (meta) dados estão associados à proveniência detalhada
	R1.3 - (meta) dados atendem aos padrões da comunidade relevantes ao domínio

Fonte: Adaptado de EUROPEAN COMMISSION (2018).

O Grupo de Trabalho da Ciência Aberta da Fiocruz elaborou um estudo no qual foi sistematizado um mapeamento e análise do processo de implantação da Ciência Aberta em oito países e na União Europeia com ênfase nas políticas e na infraestrutura de dados abertos. Neste estudo são apresentados os desafios no processo de abertura e compartilhamento dos dados tendo como destaque **as disputas por prioridade da descoberta**; a preservação de dados para futuras pesquisas; **a proteção da propriedade intelectual**; a proteção de dados sensíveis ou confidenciais, as cláusulas contratuais e regulações específicas dos países (SANTOS, 2017, grifo nosso). Mesmo com a abertura e compartilhamento dos dados o fator relacionado aos direitos autorais perpassa as discussões.

Uma pesquisa sobre Serviços de Dados de Pesquisa (RDS), realizada por um grupo de pesquisadores em colaboração com o *LIBER's Scholarly Communication & Research Infrastructures Committee and DataONE* apresenta que quase todas as bibliotecas acadêmicas da Europa estão trabalhando cooperativamente e oferecendo serviços de dados de pesquisa como, por exemplo, encontrar informações sobre PGD, padrões de metadados e práticas de citação dos dados (LIBEREUROPE, 2016).

Esse é um cenário construtivo o qual torna explícito que as Comissões, os Comitês, os Órgãos de fomento, os grupos de trabalho e as políticas institucionais contém diretrizes e orientações para disciplinar a proteção dos direitos relacionados à propriedade intelectual dos conjuntos de dados respeitando os direitos individuais e coletivos fortalecendo a Ciência Aberta.

3 REPOSITÓRIO DE DADOS E-CIENCIA DATOS

O *e-cienciaDatos* é um repositório de dados resultante do acordo de cooperação entre bibliotecas da Comunidade de Madrid composto por seis universidades: *Universidad de Alcalá*, *Universidad Autónoma de Madrid*, *Universidad Carlos III de Madrid*, *Universidad Politécnica de Madrid*, *Universidad Rey Juan Carlos* e *UNED*, membros do *Consórcio Madroño*.

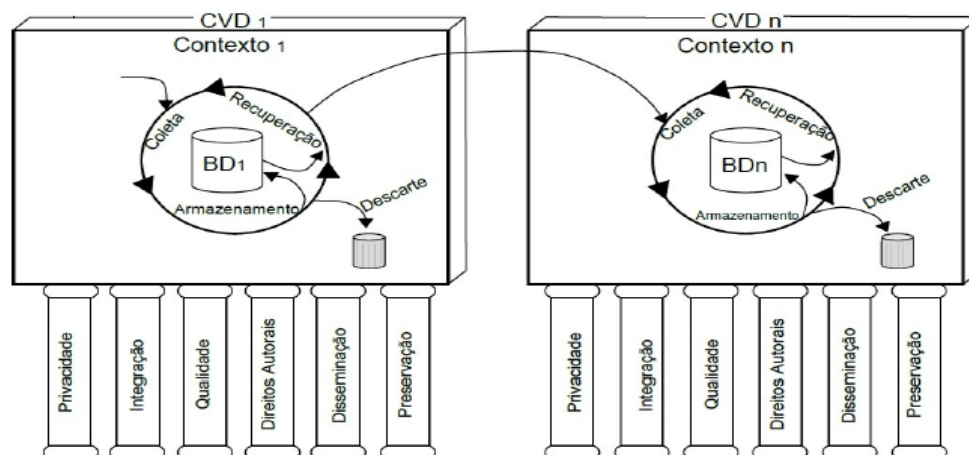
Os repositórios de dados científicos “[...] servem, entre outros propósitos, para validar os resultados da pesquisa e, portanto, devem estar vinculados de alguma forma às publicações científicas onde é mostrado para que esses dados foram usados.” (HERNÁNDEZ-PÉREZ; GARCÍA-MORENO, 2013, tradução nossa).

O repositório *e-cienciaDatos* visa organizar os conjuntos de dados e os PGDs dos pesquisadores vinculados às universidades cooperantes e oferece a gestão, a organização, o armazenamento, a preservação e a publicação dos conjuntos de dados de cada uma das universidades atribuindo um identificador de objeto digital (DOI) e referência normalizada à cada um deles com a finalidade de facilitar a verificação dos dados, sua disseminação, reutilização, impacto e acesso a longo prazo.

O PGD é um documento, mandatário ou recomendado pelos financiadores, elaborado no início do projeto de pesquisa no qual são descritas as diretrizes para o ciclo de vida dos dados (CVD) (MONTEIRO, 2017, MONTEIRO; SANT’ANA, 2018). As diretrizes orientam todos os envolvidos de alguma forma com a gestão dos dados. Nesse processo de descrição dos repositórios, deve-se observar suas políticas, os objetivos e as características da comunidade atendida. Os PGDs fornecem um índice dinâmico que articula as informações relativas a um projeto e as ligações com os vários componentes FAIR e, os repositórios, oferecem bancos de dados e serviços de dados devendo ser certificados para garantir a sua confiabilidade (EUROPEAN COMMISSION, 2018).

Destaca-se que no CVD (FIGURA 3) Sant’Ana, (2016) estabelece o fator “Direitos autorais” que consiste em garantir ao pesquisador os créditos de autoria dos conjuntos de dados e a forma como os dados são liberados para o uso, registrados no PGD incluindo a negociação de condições de processamento, divulgação e reutilização (DATA..., 2016).

Figura 3: Ciclo de vida dos dados na Ciência da Informação (CVD-CI)



Fonte: Sant'Ana (2016).

Os dados do repositório *e-cienciaDatos* estão em acesso aberto, a menos que haja razões legais, éticas ou de transferência de tecnologia que o impeçam. Destaca-se que *Consórcio Madroño* e suas universidades membros têm compromisso com a Ciência aberta e o Acesso aberto e, em suas políticas, os aspectos relacionados à propriedade intelectual estão explicitados (CONSÓRCIO MADRONO, 2017).

4 MÉTODO DA PESQUISA

A metodologia utilizada teve como base a pesquisa documental e exploratória de natureza qualitativa e quantitativa com o levantamento bibliográfico e revisão de literatura para a estruturação da fundamentação teórica e metodológica. Na pesquisa documental foram analisadas as políticas e documentos dos repositórios sobre os temas deste artigo. Foi aplicado o método estudo de caso.

Os dados foram coletados a partir da análise documental dos documentos: *Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica; Política para la Gestión de datos de Investigación; Normativa Vigente y Principales Recomendaciones*; e políticas do repositório, das entrevistas e dos questionários com os bibliotecários que atuam no *e-cienciaDatos* responsáveis pela gestão dos dados e pelos treinamentos dos usuários. Os instrumentos de coleta de dados com os bibliotecários foram: 'Roteiro de entrevista semiestruturado' e 'Questionário' com perguntas abertas e fechadas. O roteiro assegurou que o foco da entrevista fosse mantido, porém, não impediu que os entrevistados se manifestem livremente, o que possibilitou a inclusão de novas perguntas.

Para a análise dos dados foi utilizado o método ‘Análise de conteúdo’ aplicando-se a técnica ‘Análise Categorial’ (BARDIN, 2010). A Análise Categorial é a técnica mais antiga e a mais utilizada entre as técnicas da Análise de Conteúdo funcionando por desmembramento do texto em categorias e reagrupamento das unidades baseado em analogias, a partir de critérios definidos (BARDIN, 2010; FRANCO, 2008).

A definição de categorias pode ser conduzida por dois caminhos: *a priori* - pré-determinado, definidas antes da coleta de dados e *a posteriori* - emergem do conteúdo analisado e implicam constante ida e volta do material de análise à teoria (BARDIN, 2010; FRANCO, 2008). Para esse artigo foram adotadas as categorias *a priori* e *a posteriori* seguindo os critérios apresentados em Bardin (2010) e em Franco (2008) no que diz respeito a definição de categorias de qualidade, sendo adotada a categoria ‘*objetividade e a fidedignidade*’ em que “[...] As diferentes partes de um mesmo material, ao qual se aplica uma determinada matriz de categorias, devem ser codificadas da mesma maneira, mesmo quando submetidas a várias análises.” (FRANCO, 2008, p. 67). As categorias selecionadas foram: Propriedade intelectual, direito autoral, licenças de uso, licenças de depósito e citação de dados. O tratamento dos resultados da análise de conteúdo permitiu gerar inferências para apresentação dos resultados, pois a análise de conteúdo é utilizada para “[...] produzir inferências acerca de dados, verbais e/ou simbólicos, mas, obtidos a partir de perguntas e observações de interesse de um determinado pesquisador.” (FRANCO, 2008, p. 17).

Na apresentação dos resultados foram considerados: a) os eixos estabelecidos na taxonomia da *Open Science* da *Fostering the Practical Implementation of Open Science in Horizon 2020 and Beyond* (FOSTER) (FIGURA 1) partindo de *Open Science* - *Open Science Policies* - *Subject Policies* - *Open Access Policies* incluindo a taxonomia *Legal Issues* que subdivide-se em *Intellectual Property Rights, Law* e *Licenses* (FOSTER, 2018)(FIGURA 2); b) os princípios FAIR em seus princípios - *Findable* - F1, F2, *Reusable* apresentados no quadro 1.

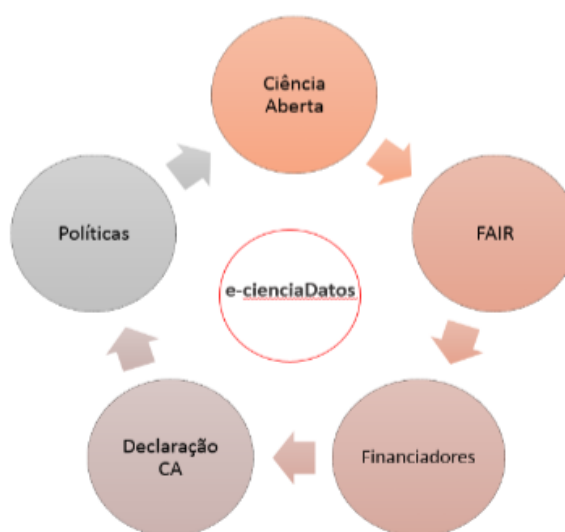
5 RESULTADOS E DISCUSSÃO

O *Consórcio Madroño* está sistematizando suas atividades para atender os fatores envolvidos com a Ciência Aberta e assessorar seus pesquisadores na publicação aberta de seus *datasets* e na geração de planos de gestão de dados. Mediante análise dos documentos, entrevistas e questionários realizou-se a análise categorial proposto no método Análise de

conteúdo para verificar os princípios relacionados ao direito autoral norteadores do repositório.

O repositório *e-cienciaDatos* promove a adoção de práticas de pesquisa abertas e publicação de dados abertos corroborando com suas Políticas, Declaração Ciência Aberta, Financiadores, diretrizes da Ciência aberta e princípios FAIR (FIGURA 4).

Figura 4: Princípios norteadores do repositório *e-cienciaDatos*



Fonte: Elaborada pelos autores.

A análise dos documentos (FIGURA 4) propiciou pontuar os principais aspectos da Ciência Aberta (CA) apontados no trabalho de Pontika et al. (2015):

- **Open data:** *datasets* estão disponíveis *on-line* para acesso e reutilização. Cada conjunto de dados tem informações sobre a licença aplicada. Destaca-se que a política do repositório deixa claro que são utilizadas somente licenças livres e, caso se tenha a necessidade de outra licença ou uma combinação de licenças, deve-se entrar em contato com os administradores do *Consórcio Madroño*. A opção de embargo também será aplicada caso necessário e com a indicação do prazo;
- **Open source:** é recomendado pelo repositório o uso de formatos abertos, *software* livre e aplicação de licenças livres para *software* aprovadas por *Open Source Initiative* (OSI), *GNU General Public License*, licença MIT entre outras além do uso de recursos educativos abertos. Esses fatores refletem na visualização ou reuso dos dados;
- **Open reproducible research:** nos documentos *Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica*, *Política para la Gestión de datos de Investigación*, *Normativa Vigente y Principales Recomendaciones* e Políticas

do *Consórcio Madroño* estão discriminadas as práticas para a reprodutibilidade aberta da pesquisa.

Esses aspectos deixam claro as convergências com os princípios da Ciência Aberta.

Seguindo a análise, é possível detalhar os aspectos indicados na *Open Science Taxonomy - Open Science Policies - Organisational Mandates* - subdividido em: *Funders Policies, Governmental Policies; Institutional Policies* conforme apresentados na figura 1. Os aspectos indicados na *Open Science Taxonomy - Open data* subdividido em: *Intellectual Property Rights, Law e Licenses* (FIGURA 2) também são regulamentados pelas políticas do repositório.

O repositório dispõe de vários documentos que regulamentam o CVD que indica as políticas governamentais e das agências de fomento (FIGURA 5). Os documentos representados na figura 5 são: Políticas - instruções gerais sobre os dados depositados; *Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica; Política para la Gestión de datos de Investigación; Normativa Vigente y Principales Recomendaciones; Licencia de uso; Licencia de depósito.*

Figura 5: Documentos norteadores do repositório *e-cienciaDados*



Fonte: Elaborada pelos autores.

Na política “Licença de depósito” está explícito que os dados estão em acesso aberto, a menos que haja razões legais, éticas ou de transferência de tecnologia que o impeça. A política “Licença de uso” esclarece que ela é um instrumento legal através do qual o proprietário dos direitos de propriedade intelectual outorga que terceiros utilizem os

conjuntos de dados sem infringir estes direitos. As Políticas indicam as licenças para cessão e reutilização dos conjuntos de dados.

No documento *Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica* são enfatizadas as recomendações para a adoção dos princípios da Ciência Aberta, a aplicação das boas práticas seguindo os princípios FAIR e, apresenta, as indicações sobre a cessão de direitos autorais às revistas e relevância da citação dos autores.

Em relação aos direitos morais o repositório indica o *link* das informações sobre padrões de citação dos conjuntos dos dados e disponibiliza a referência de cada *datasets* deixando claro a importância do reconhecimento de sua autoria.

A menção dos principais programas de financiamento de pesquisas identificados na análise dos documentos foram o Programa de investigação e inovação da União Europeia (UE) *Horizon 2020*⁴ e *Horizon Europe*⁵. Ressalta-se que na política do **Horizon 2020** os direitos de acesso e a proteção da propriedade intelectual são apresentados.

Mediante a análise sobre os princípios FAIR foi possível identificar os princípios *Findable* e *Reusable* conforme apresentado no quadro 2. Evidencia-se que nos princípios FAIR em seu princípio *Reusable* - R1.1 - (meta)dados são liberados com uma licença de uso de dados clara e acessível.

⁴ <https://ec.europa.eu/programmes/horizon2020/en>

⁵ https://ec.europa.eu/info/designing-next-research-and-innovation-framework-programme/what-shapes-next-framework-programme_en

Quadro 2: Princípios FAIR no e-cienciaDados

Princípios FAIR	Descrição dos Princípios	Metadados	Repositório e-cienciaDados
Findable	F1- (meta) dados são atribuídos a um identificador globalmente exclusivo e persistente	Title Author	DOI ORCID
	F2 - os dados são descritos com metadados ricos (definidos por R1 no princípio Reusable)	Title Author	nomeados autores dos datasets ORCID indicação das instituições aos quais estão vinculados
Reusable	R1.1 - (meta) dados são liberados com uma licença de uso de dados clara e acessível	Rights Terms of use	Creative commons CC0, CC-BY, CC-BY-SA Open Data Commons (ODC) Public Domain Dedication and Licence (PDDL) Open Source Initiative (OSI) GNU General Public License Versions MIT license

Fonte: Elaborada pelos autores.

Nas práticas da Ciência aberta estão evidentes que os dados devem ser unicamente identificáveis e listados na seção de referência de um documento. Tanto o *DataCite*⁶ como o ORCID⁷ suportam isso. O *DataCite* tornou possível o uso generalizado de (DOI), está ligado a uma descrição de metadados do objeto e a um local digital, como um URL, onde todos os detalhes sobre o objeto estão acessíveis, além de ligado aos artigos publicados (TAYLOR et al., 2017). O ORCID identifica exclusivamente os pesquisadores e os liga às suas publicações, garantindo que seus trabalhos sejam identificados corretamente (TAYLOR et al., 2017). Esses identificadores podem estar ligados aos artigos científicos, aos resultados de pesquisa, incluindo equipamentos, experimentos, patentes e conjuntos de dados (TAYLOR et al., 2017).

Tanto o *DataCite* como o ORCID viabilizam o desenvolvimento de métricas de citação. Este acompanhamento claro do trabalho (DOIs) com pesquisadores (ORCIDs) significam que os índices de citação são mantidos possibilitando a cienciometria (por exemplo, h-index) e altmetria (baseado em mídia) e os autores recebem crédito pelo seu trabalho por meio de citações ou “menções” (TAYLOR et al., 2017).

Salienta-se que para o pesquisador que almeja ter melhor desempenho em seu *status* profissional e ter reconhecimento perante seus pares (agradecimentos e citações), o compartilhamento de seus dados propicia oportunidades para ser reconhecido como o autor,

⁶ www.datacite.org

⁷ www.orcid.org

ser citado por membros mais proeminentes de seu campo de atuação e receber o reconhecimento formal (KIM; STANTON, 2013, 2016).

A gestão e publicação dos conjuntos de dados dos repositórios seguem as normas, diretivas e Lei referentes à propriedade intelectual. Nos documentos analisados constatou-se que os conjuntos de dados dispõem da proteção autoral prevista na lei. Nos treinamentos dos bibliotecários aos usuários a questão dos direitos autorais é apresentada expondo as possibilidades da cessão, da disponibilização e da reutilização dos dados. Interessante destacar que a cessão dos direitos autorais não é exclusiva ao repositório.

Sendo assim, há convergências entre a Ciência Aberta e as boas práticas dos princípios *Findable* e *Reusable* dos princípios FAIR com a operacionalização do repositório. Não foram identificadas divergências entre os mesmos.

6 CONSIDERAÇÕES FINAIS

A Ciência Aberta é um movimento com a premissa de que a documentação científica e os conjuntos de dados e metadados estejam em acesso aberto o quanto possível. O repositório *e-cienciaDatos* gerencia os conjuntos de dados de sua comunidade e, suas políticas, sua *Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica*, a *Política para la Gestión de datos de Investigación*, a *Normativa Vigente y Principales Recomendaciones* estão em consonância com as diretrizes da Ciência Aberta e princípios FAIR.

O aspecto de direitos autorais está explicitado nos documentos analisados os quais estão seguindo as normas, as diretivas e a lei de propriedade intelectual confirmando que os conjuntos de dados têm direitos morais e patrimoniais.

O *Consórcio Madroño* e seus membros subscrevem seu compromisso com o Acesso Aberto e Ciência Aberta ao coletar e disseminar publicações e dados resultantes das pesquisas de forma aberta.

REFERÊNCIAS

ABADAL, E. et al. Maredata: red temática sobre datos de investigación en abierto. In: MERLO VEGA, J. A. (ed.). **Ecosistemas del Acceso Abierto**. Salamanca: Ediciones Universidad de Salamanca, 2018. p. 105-114. Disponível em: <http://hdl.handle.net/10366/138585>. Acesso em: 10 jan. 2019.

ALTENHÖREN, R. Open Source Software: definition, licensing models and organizational consequences (introduction). *In: IFLA GENERAL CONFERENCE AND COUNCIL - LIBRARIES: VOYAGE OF DISCOVERY*, 71., 2005, Oslo, 2005. **Proceedings [...]**. The Hague: IFLA, 2005. Disponível em: <http://archive.ifla.org/IV/ifla71/papers/121e-Altenhoener.pdf>. Acesso em: 10 jan. 2019.

BARDIN, L. **Análise de conteúdo**. Rev. e atual. Lisboa: Edições 70, 2010.

BEZJAK, S. et al. **Manual de formação em Ciência Aberta**. FOSTER, 2018. Disponível em: book.fosteropenscience.eu/pt. Acesso em: 10 jan. 2019.

BORGMAN, C. L. et al. Knowledge infrastructures in science: data, diversity, and digital libraries. **International Journal on Digital Libraries**, New York, v. 16, n. 3, p. 207–227, Sep. 2015. Disponível em: <http://link.springer.com/article/10.1007%2Fs00799-015-0157-z#page-1>. Acesso em: 2 nov. 2015.

CONSÓRCIO MADROÑO. **Publicación de datos científicos**. Madrid, 2016. Disponível em: <https://www.consorcioadrono.es/investigam/curso-publicacion-de-datos-cientificos/>. Acesso em: 10 jan. 2019.

CONSÓRCIO MADROÑO. **Declaración del Consorcio Madroño en apoyo a la Ciencia Abierta a la información académica y científica**. Madri, 2017. Disponível em: https://www.consorcioadrono.es/docs/declaracion_ciencia_abierta.pdf. Acesso em: 15 dez. 2018.

COUTO CORRÊA, F. **Gestión de datos de investigación**. Barcelona: Editorial UOC, 2016. Disponível em: <http://bit.ly/2uwefAX>. Acesso em: 2 jul.2017.

DATA Management Plan 4: data access and security. Utrecht: Utrecht University, 2016. Disponível em: <http://libguides.library.uu.nl/c.php?g=202272&p=1330220&preview=1776d74a25e3418c4d63aafdbf99598f#s-lg-box-wrapper-4739017>. Acesso em: 18 ago. 2016.

EUROPEAN COMMISSION. **Turning FAIR into reality**. Luxemburgo: Final Report and Action Plan from the European Commission Expert Group on FAIR Data. Publications Office of the European Union, 2018. Disponível em: https://ec.europa.eu/info/sites/.../turning_fair_into_reality_1.pdf. Acesso em: 10 jan. 2019.

FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO (FAPESP). Programa FAPESP de Pesquisa sobre Mudanças Climáticas Globais: Chamada de Propostas de Pesquisa 2017. São Paulo, 2017a. Disponível em: <http://www.fapesp.br/11068>. Acesso em: 15 dez. 2017.

FONSECA JÚNIOR, W. C. Análise de conteúdo. *In: DUARTE, J.; BARROS, A. Métodos e técnicas de pesquisa em comunicação*. 2. ed. São Paulo: Atlas, 2014. p. 280-304.

FOSTERING THE PRACTICAL IMPLEMENTATION OF OPEN SCIENCE IN HORIZON 2020 AND BEYOND. **Open Science taxonomy**. 2018. Disponível em: https://www.fosteropenscience.eu/themes/fosterstrap/images/taxonomies/os_taxonomy.png. Acesso em: 19 mar. 2019.

FRANCO, M. L. P. B. **Análise de conteúdo**. 2. ed. Brasília, DF: Liber Livros, 2005.

GANDELMAN, H. **De Gutenberg à internet: direitos autorais das origens à era digital**. 5. ed. Rev. e atual. Rio de Janeiro: Record, 2007.

HERNÁNDEZ-PÉREZ, T.; GARCÍA-MORENO, M. A. (2013). Datos abiertos y repositorios de datos: nuevo reto para los bibliotecarios. **El profesional de la información**, Barcelona, v. 22, n. 3, p. 259-263, mayo-jun. 2013. Disponível em: <https://recyt.fecyt.es/index.php/EPI/article/view/epi.2013.may.10/17811>. Acesso em: 18 ago 2017.

HERNÁNDEZ-PÉREZ, T.; MÉNDEZ RODRÍGUEZ, E. **Improving Health Research in EU through FAIR Data**. Seville: FAIR4HEALTH, 2019. Disponível em: <https://www.fair4health.eu/storage/files/Resource/15/D23%20Guidelines%20for%20implementing%20FAIR%20Open%20Data%20policy%20in%20health%20research.pdf>. Acesso em: 20 jul. 2019.

KIM, Y.; STANTON, J. M. Institutional and individual influences on scientists' data sharing practices. **Journal of Computational Science Education**, Durham, v. 3, p. 47-56, 2013. Disponível em: https://www.researchgate.net/publication/239522383_Institutional_and_Individual_Influences_on_Scientists%27_Data_Sharing_Practices. Acesso em: 29 set. 2018.

KIM, Y.; STANTON, J. M. Institutional and individual factors affecting scientists data-sharing behaviors: a multilevel analysis. **Journal Of the Association for Information Science and Technology**, Hoboken, v. 67, n. 4, p. 776–799, 2016. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/asi.23424>. Acesso em: 29 set. 2018.

LIBEREUROPE. **Research data services in europe's academic research libraries**. The Hague, 2016. Disponível em: <https://libereurope.eu/blog/2016/10/13/research-data-services-europes-academic-research-libraries/>. Acesso em: 08 maio 2018.

MARTONE M. (ed.) **Data Citation Synthesis Group: Joint Declaration of Data Citation Principles**. San Diego CA: FORCE11, 2014. Disponível em: <https://www.force11.org/datacitationprinciples>. Acesso em: 10 jan. 2019.

MONTEIRO, E. C. S. A. **Direitos autorais nos repositórios de dados científicos: análise sobre os planos de gerenciamento dos dados**. 2017. 115 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de filosofia e Ciências, Universidade Estadual Paulista, Marília, 2017. Disponível em: <http://hdl.handle.net/11449/149748>. Acesso em: 30 abr. 2017.

MONTEIRO, E. C. S. A.; SANT'ANA, R. C. G. Plano de gerenciamento de dados em

repositórios de dados de universidades. **Encontros Bibli**, Florianópolis, v. 23, n. 53, p. 160-173, set./dez. 2018. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p160/37296>. Acesso em: 10 jan. 2019.

MONTEIRO, E. C. S. A.; SANT'ANA, R. C. G. Direitos autorais e Ciência Aberta: estudo do Repositório de dados do Consórcio Madroño. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 28., 2019, Vitória. No prelo.

MORA, F. (coord.). **Compromisos de las universidades ante la Open Science**. Madri: CRUE Universidades Españolas, 2019. Disponível em: http://www.crue.org/Documentos%20compartidos/Informes%20y%20Posicionamientos/2019.02.20-Compromisos%20CRUE_OPENSCIENCE%20VF.pdf. Acesso em: 24 jun. 2019.

MURRAY-RUST, P. Open Data in Science. **Serials Review**, New York, v. 34, n. 1, p.52-64, 2008. doi:10.1016/j.serrev.2008.01.001. Acesso em: 10 jan. 2019.

NATIONAL SCIENCE BOARD. **Long-lived digital data collections**: enabling research and education in the 21st century. Alexandria, 2005. Disponível em: <http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf>. Acesso em: 30 maio 2016.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OCDE). **OECD Principles and Guidelines for Access to Research Data from Public Funding**. Paris: OCDE, 2007. Disponível em: <http://www.oecd.org/sti/scitech/38500813.pdf>. Acesso em: 20 set. 2016.

PONTIKA, N. et al. Fostering Open Science to Research using a Taxonomy and an eLearning Portal. *In*: INTERNATIONAL CONFERENCE ON KNOWLEDGE TECHNOLOGIES AND DATA DRIVEN BUSINESS, 15., 2015, Graz. **Anais [...]**. Milton Keynes: The Open University, 2015. Disponível em: http://oro.open.ac.uk/44719/2/kmi_foster_iknow.pdf. Acesso em: 10 jan. 2019.

PELLS, R. Open access: Plan S launch delayed until 2021. **Times Higher Education**, London, May 2019. Disponível em <https://www.timeshighereducation.com/news/open-access-plan-s-launch-delayed-until-2021>. Acesso em: 31 maio 2019.

QUADERI, N. et al. **The Plan S footprint**: Implications for the scholarly publishing landscape. Jersey, 2018. Disponível em: <https://clarivate.com/g/plan-s-footprint/>. Acesso em: 22 mar. 2019.

RAMJOUÉ, C. Towards Open Science: the vision of the European Commission. **Information Services & Use**, Amsterdam, v. 35, p. 167–170, 2015. Disponível em: <https://pdfs.semanticscholar.org/c47a/f7edc0607c6561390bba2a993c89711297c8.pdf>. Acesso em: 10 jan. 2019.

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação e informação**, Londrina, v. 21, n. 2, p. 116-142, maio/ago. 2016. Disponível em:
<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27940/20124>. Acesso em: 20 out. 2016.

SANTOS, P. X. (coord.). **Livro Verde - Ciência aberta e dados abertos**: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz, 2017. Disponível em: <https://www.arca.fiocruz.br/handle/icict/24117>. Acesso em: 10 jan. 2019.

STODDEN, V. Enabling Reproducible Research: Open Licensing for Scientific Innovation. **International Journal of Communications Law and Policy**, v. 13, p.1-25, 2009.

SUBER, P. **Open Access**. Boston: MIT Press, 2012.

TAYLOR, S. J. E. **Open science**: approaches and benefits for modeling & simulation. IEEE, 2017. Disponível em: <https://ieeexplore.ieee.org/document/8247813>. Acesso em: 22 mar. 2019.

UNIVERSIDAD CARLOS III DE MADRI (UC3M). **Propiedad intelectual y derechos de autor**: cuestiones generales. Getafe, [201-a]]. Disponível em: Getafe, [201-b].
<https://www.uc3m.es/ss/Satellite/Biblioteca/es/TextoDosColumnas/1371213999139/>. Acesso em: 10 ago. 2019.

UNIVERSIDAD CARLOS III DE MADRI (UC3M). **Propiedad intelectual y nuevas tecnologías**. Getafe, [201-b]]. Disponível em:
<https://www.uc3m.es/ss/Satellite/Biblioteca/es/TextoDosColumnas/1371213998525/>. Acesso em: 10 ago. 2019.

UNIVERSIDAD CARLOS III DE MADRI (UC3M). **Marco legal de los datos de investigación**. Getafe, [201-a]]. Getafe, [201-c]. Disponível em:
<https://www.uc3m.es/biblioteca/propiedad-intelectual-industrial/propiedad-intelectual+>. Acesso em: 10 ago. 2019.

UNIVERSITY OF MELBOURNE. **Management of Research Data and Records Policy (MPF1242)**. Melbourne, 2009. Não paginado. Disponível em:
<https://policy.unimelb.edu.au/MPF1242>. Acesso em: 10 jan. 2019.