

BOAS PRÁTICAS CIENTÍFICAS NA ELABORAÇÃO DE PLANOS DE GESTÃO DE DADOS

GOOD SCIENCE PRACTICES IN THE DRAWING UP OF DATA MANAGEMENT PLANS

¹Clara Alcina Guandalini

²Ariadne Chloë Mary Furnival

³Ana Carolina Simionato Arakaki

Universidade Federal de São Carlos^{1,2,3}

Correspondência

Clara Alcina Guandalini
Universidade Federal de São Carlos
São Carlos, SP - Brasil

E-mail: clara.alcina@gmail.com

 ORCID: <https://orcid.org/0000-0002-4382-1779>



JITA: FJ. Knowledge management

e-Location: 019034

RESUMO

As instituições e as agências de pesquisa vêm exigindo uma conduta no contexto da agenda em prol da Ciência Aberta. Sendo que, a elaboração de um Plano de Gestão de Dados pode ser caracterizada como um procedimento inicial para o desenvolvimento da pesquisa científica, documento que propõe ao pesquisador gerenciar os dados brutos de sua pesquisa, valorizando o compartilhamento em conjunto com a abertura dos dados e seu reuso em benefício aos pesquisadores e à sociedade. Diante disso, o objetivo desse trabalho foi verificar a elaboração de um Plano de Gestão de Dados como parte das novas condutas para as boas práticas na pesquisa científica. A pesquisa caracterizou-se como exploratória, empregou-se métodos documentais e bibliográficos para a identificação, coleta e sistematização das informações. Os resultados visaram a disponibilidade desse documento em repositórios de dados abertos e a relação das instituições de ensino superior, agências financiadoras e pesquisadores com o levantamento de ações que são realizadas na criação e desenvolvimento de Plano de Gestão de Dados, principalmente para o desenvolvimento de políticas institucionais para o gerenciamento dos dados de pesquisa, porém, respeitando também as áreas do conhecimento, visto que nem todos os dados são gerados do mesmo modo. Por fim, considera-se a presença de um novo paradigma na publicação e divulgação científica, o que influencia também sua comunicação, seja com o meio acadêmico ou com a sociedade.

PALAVRAS-CHAVE

Plano de gestão de dados. Dados de pesquisa. Gerenciamento de dados. Ciência aberta.

ABSTRACT

Institutions and research agencies have demanding conduct in the context of the Open Science agenda. The elaboration of a Data Management Plan can be characterized as an initial procedure for the development of scientific research, a document that proposes the researcher to manage the raw data of his research, valuing sharing together with the opening of data and its reuse for the benefit of researchers and society. Therefore, the objective of this work was to verify the elaboration of a Data Management Plan as part of the new conducts for the good practices in scientific research. The research was characterized as exploratory, documentary and bibliographic methods were used for the identification, collection and systematization of information. The results aimed at the availability of this document in open data repositories and the relationship of higher education institutions, funding agencies and researchers with the survey of actions that are carried out in the creation and development of Data Management Plan, mainly for the development of institutional policies for the management of research data, but also respecting areas of knowledge, since not all data is generated in the same way. In addition, it is evident that there is a new paradigm in scientific publication and dissemination, which also influences its communication, be it with the academic environment or with society.

KEYWORDS

Data management plan. Search data. Data management. Open Science.

1 Introdução

A produção de conhecimento científico envolve princípios de integridade, respeitando-se uma conduta ética por parte dos pesquisadores. Pensando nisso, as boas práticas que são apresentadas no decorrer do presente texto refletem a essa integridade científica. O termo “boas práticas” pode ser entendido como condutas adotadas para uma maior divulgação, disseminação e desenvolvimento da ciência, visando a sua abertura à sociedade.

O crescimento constante da produção científica faz com que as instituições de ensino e agências financiadoras se atentem cada vez mais com essa conduta no contexto da agenda em prol da Ciência Aberta. Dessa forma, as boas práticas na pesquisa científica podem incluir a elaboração de um Plano de Gestão de Dados (doravante, PGD ou ainda, *Data Management Plan* - DMP), que propõe ao pesquisador gerenciar os dados brutos de sua pesquisa. Pode-se dizer que o PGD constitui uma nova forma de realizar a ciência, valorizando o compartilhamento em conjunto com a abertura dos dados e seu reuso em benefício aos pesquisadores e à sociedade: a necessidade de desenvolver um PGD compõe as novas demandas e recomendações ligadas à abertura dos dados científicos.

O movimento *Open Data* (“dados abertos” em português) é um termo que propõe a transparência na divulgação dos dados, tendo como propósito estabelecer sua possível reutilização. Esse movimento surge no contexto de outro mais amplo, o da Ciência Aberta (*Open Science*) que inclui a promoção dos dados abertos na ciência e consequentemente retrata uma forma mais democrática e acessível em se realizar ciência, além de enfatizar a necessidade e importância de replicar e reproduzir as pesquisas, especialmente as experimentais. Esses movimentos na ciência conduzem a novas perspectivas no desenvolvimento das pesquisas científicas e ocasionam o surgimento de questões nas publicações científicas relacionados com: arquivamento dos dados, suporte na abertura dos dados, vantagens e desvantagens dos dados abertos, como realizar a abertura de dados, e como descrever os dados abertos em um PGD.

Este estudo tomou como ponto de partida que a elaboração de um PGD já é requerimento no meio científico, dado que instituições e principalmente, as agências de fomento à pesquisa – nacionais e internacionais – estipulam tal requerimento. Diante disso, o objetivo desse trabalho foi verificar a elaboração de um Plano de Gestão de Dados como parte das novas condutas para as boas práticas na pesquisa científica.

A pesquisa caracteriza-se como exploratória, pois possibilitou a análise e discussão do tema por meio de métodos documentais e bibliográficos para a identificação, coleta e sistematização das informações, sendo que o levantamento documental foi realizado principalmente a partir de páginas da internet: de instituições, projetos, repositórios e bases de dados. Levou-se em consideração primeiramente a busca de artigos oriundos de profissionais da área, instituições que possuem iniciativas *Open* e instituições governamentais e acadêmicas, a partir do pressuposto que estariam atentas a essas ações na comunicação

científica. A busca de publicações também se deu com auxílio do Portal de Periódicos CAPES, a ferramenta de busca do Google Acadêmico e outras bases, como *Scielo* e BRAPCI.

Antes de abordar os PGDs especificamente, precisamos contextualizá-los com uma breve delimitação do movimento de Ciência Aberta, dentro do qual surge uma preocupação com os dados científicos abertos tal como as formas de gerenciá-los, o que faremos nas seções a seguir.

2 Ciência Aberta

O acesso aberto na comunicação científica surge, a partir de 2001, com o intuito de disseminar, aberta e gratuitamente, os frutos da pesquisa científica, tanto na forma das publicações tradicionais, como artigos, relatórios, teses e dissertações, quanto posteriormente, dados científicos. A *Budapest Open Access Initiative* (BOAI), uma reunião organizada em 2001 pelo *Open Society Institute* (OSI), visou discutir o acesso aberto na literatura científica e quais iniciativas poderiam impulsionar esse movimento. Esse evento tornou-se mundialmente conhecido, sendo que a declaração da BOAI explicita que o acesso aberto da literatura científica significa:

[...] disponibilidade livre na Internet, permitindo a qualquer usuário ler, fazer *download*, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral desses artigos, recolhe-los para indexação, introduzi-los como dados em software, ou usá-los para outro qualquer fim legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis ao próprio acesso a uma conexão à Internet. As únicas restrições de reprodução ou distribuição e o único papel para o *direito autoral* neste domínio é dar aos autores o controle sobre a integridade do seu trabalho e o direito de ser devidamente reconhecido e citado. (BOAI, 2001?)

Dez anos depois, em 2012, os organizadores do BOAI elaboraram uma publicação salientando que o acesso aberto às publicações científicas ainda estava longe de ser alcançado, e apresentaram mais recomendações para os próximos dez anos de como a iniciativa ainda poderia crescer e se consolidar. Como observam Furnival e Silva-Jerez (2017, p.153), o acesso aberto às publicações científicas de fato:

[...] faz parte de um cenário mais amplo em prol da abertura ao conhecimento em geral (acesso aberto, dados abertos, recursos educacionais abertos, software livre, licenças abertas) e constitui essencialmente um movimento na direção da informação e conhecimento como um bem público.

Entendemos que a referida abertura ao conhecimento como a Ciência Aberta, sendo que a expansão da *internet* e resultantes das Tecnologias da Informação e Comunicação (TIC) tem impulsionado o contexto da Ciência Aberta para uma nova proporção, possibilitando outras maneiras de compartilhamento e colaboração primeiramente entre pesquisadores e a sociedade. A Ciência Aberta pode ser considerada como movimento social,

visto que demonstra novos conceitos nas práticas científicas, assim como afirma Albagli (2015, p.14):

O movimento pela ciência aberta, em seu formato atual, reflete, na verdade, novos modos de pensar e de exercer a cientificidade, com repercussões diretas sobre os compromissos, normas e arcabouços institucionais que interferem diretamente na prática científica e nas suas relações com a sociedade.

O conceito de Ciência Aberta na atualidade vai além de disponibilizar um conteúdo em acesso aberto, a prática da ciência torna-se acessível oferecendo não só o acesso ao produto final da pesquisa e sim a todo seu desenvolvimento e dessa forma, se permite que os dados possam ser utilizados, reusados e distribuídos. Essa mudança na realização de pesquisas implica numa maior colaboração, facilitando a comunicação científica e por consequência maior avanço, como apontam Chan, Okune e Sambuli:

[...] a ciência aberta e colaborativa também promete aumentar a visibilidade e o impacto da pesquisa em nível local, facilitar a participação de pesquisadores em colaborações locais e internacionais, estimular o engajamento do público com a ciência por meio de atividades tais como ciência cidadã, e promover a cultura do compartilhamento do conhecimento, bem como uma nova reflexão sobre inovação social. (CHAN; OKUNE; SAMBULI, 2015, p. 103).

Para que esse novo formato da ciência ocorra de forma viável, e sendo que a ciência é fundamentada em dados, têm surgido iniciativas para promover e viabilizar o compartilhamento aberto dos dados científicos gerados em pesquisas e que na maioria das vezes, subsidiam a escrita de artigos e relatórios científicos. Os dados científicos abertos implicam em novos desafios para ciência, tanto tecnológicos como políticos, assim como Albagli aponta:

[...] no desenvolvimento da ciência aberta, para além dos aspectos técnicos e tecnológicos (como o desenvolvimento de ferramentas livres, disponibilidade de plataformas computacionais abertas, e infraestrutura tecnológica para compartilhamento de dados), são as questões de ordem cultural, política e institucional (formais e informais) que mais interferem no caráter aberto ou proprietário dessas práticas. (ALBAGLI, 2015, p.17)

Em 2011, o *Research Information Network* (RIN) e o *National Endowment for Science, Technology and the Arts* (NESTA), ambos do Reino Unido, publicaram o relatório “*Open to All? Case studies of openness in research*”. Este relatório tem como objetivo oferecer uma melhor percepção sobre os métodos da Ciência Aberta aos pesquisadores, instituições e financiadores. Destarte o relatório analisou diferentes casos de estudos e quais foram as consequências de acordo com os pesquisadores participantes, acompanhando os benefícios e barreiras na abertura de dados nas pesquisas. Os benefícios se destacaram em: eficiência na pesquisa; qualidade de pesquisa e rigor acadêmico; visibilidade e escopo de engajamento; novas questões de pesquisa; colaboração e construção de comunidades; e impacto social e econômico. Já as barreiras e restrições foram identificadas como: falta de evidência de benefícios; falta de incentivos, recompensas e apoio; falta de tempo, habilidades e outros recursos; culturas de independência e competição; preocupações sobre qualidade e usabilidade; e restrições éticas, legais e outras sobre a abertura.

Já o relatório “*Science as an open enterprise: open data for open science*” (2012) da *Royal Society*, visou estabelecer primeiramente algumas recomendações de princípios fundamentais, à medida que a comunicação científica vêm vivenciando o impacto de novas tecnologias:

Este relatório analisa o impacto de tecnologias novas e emergentes que estão transformando a conduta e a comunicação da pesquisa. As recomendações foram elaboradas para melhorar a conduta da ciência, responder às expectativas do público em mudança e à cultura política e permitir que os pesquisadores maximizem o impacto de suas pesquisas. (ROYAL SOCIETY, 2012, tradução nossa).

Outro termo que surge nesse cenário é o de “Ciência Cidadã”, como uma ciência mais participativa, sendo que os fatores que levam a essa maior colaboração podem ser considerados provenientes da expansão das tecnologias digitais e de compartilhamento. A participação mais ativa dos cidadãos por meio da *internet* e os aplicativos nos dispositivos móveis, apontam para uma maior transparência dos dados e informações conforme interesses públicos em determinados assuntos da sociedade, incluindo o meio acadêmico e científico. Segundo Parra (2015, p. 126): “Ciência cidadã refere-se ao engajamento do público em geral nas atividades de pesquisa científica quando os cidadãos contribuem ativamente para a ciência, seja com seu esforço intelectual, com seu conhecimento local ou com suas ferramentas e recursos.” Podemos citar como exemplos de projetos de amplo alcance de ciência cidadã são *Zooniverse*, *Asteroid Zoo* e *Cochrane Crowd*, entre muitos outros pelo mundo¹.

Mediante a essa maior participação e colaboração dos cidadãos, nota-se que atualmente a produção de dados é exponencial, sendo que como cidadãos, dispomos nossos dados todos os dias para diversas empresas por meio de redes sociais, aplicativos e *sites*, usualmente de forma automática, sem preocupação sobre o destino e arquivamento de longo prazo desses dados. Aqui, pode-se introduzir o conceito de *Big Data*, que compreende:

[...] a geração, tratamento e análise de grandes volumes de dados que excedem as capacidades convencionais de processamento, e vem sendo explorado também por empresas, governos e outros segmentos interessados em extrair informações a partir de grandes quantidades de dados não estruturados. Já e-Science incorpora, além do uso intensivo de dados, a pesquisa científica colaborativa e o uso de recursos compartilhados para sua exploração. (ALBAGLI; APPEL; MACIEL, 2014).

É importante ressaltar que o conceito de *e-Science* diferencia-se de Ciência Aberta, pois é um termo que se refere primeiramente à colaboração que pode ser feita a distância entre pesquisadores e grupos de pesquisa por meio das TIC. Dessa forma, é possível analisar que o termo *e-Science* tenha contribuído para o princípio da Ciência Aberta e também que uma pesquisa realizada em colaboração mediada por infraestrutura de informática computacional pode se identificar como *e-Science*, mas nem sempre como *Open Science* (ALBAGLI; APPEL; MACIEL, 2014).

¹Veja a lista de mais que 200 projetos configurados como ciência cidadão em: <https://en.wikipedia.org/wiki/List_of_citizen_science_projects>.

Os conceitos de acesso aberto se expandiram, impulsionando um movimento de mudança no compartilhamento de dados, tanto governamentais como científicos. A partir da Ciência Aberta, se têm a consciência que os dados científicos também precisam ser mais divulgados em colaboração para a comunidade científica, de acordo com novas tecnologias que conjuntamente potencializam uma *Web* mais participativa. É visível que o movimento da Ciência Aberta engloba os dados abertos e se faz necessário a reflexão de como esses dados podem ser divulgados de forma ética e como é possível declará-los em um Plano de Gestão.

3 Dados Abertos Científicos

Os dados abertos, definidos pela *Royal Society* (2012, p.12) como: “[...] dados que atendem aos critérios de abertura inteligente, sendo que devem ser acessíveis, utilizáveis, avaliáveis e inteligíveis”, representam um novo formato de produzir o acesso aberto, sendo que essa abertura é de suma importância para que a Ciência Aberta possa se desenvolver, pois conforme aponta Molloy:

Quanto mais dados forem disponibilizados de forma útil, maior será o nível de transparência e reprodutibilidade e, portanto, mais eficiente o processo científico se tornará, em benefício da sociedade. Este ponto de vista está se tornando *mainstream* entre muitos financiadores, editores, cientistas e outras partes interessadas na pesquisa, mas as barreiras para a divulgação generalizada de dados abertos permanecem. (MOLLOY, 2011, p.1, tradução nossa).

No guia *Open Data Handbook*, da *Open Knowledge International*: “[...] dados abertos são dados que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa – sujeitos, no máximo, à exigência de atribuição da fonte e compartilhamento pelas mesmas regras”. (OPEN KNOWLEDGE INTERNATIONAL, s.d.). A *Open Knowledge Foundation* é uma importante organização que vêm corroborando com essas iniciativas desde sua fundação em 2004 e segundo seu *site*², tem como missão:

[...] ver sociedades esclarecidas em todo o mundo, onde todos têm acesso a informações importantes e a capacidade de usá-las para entender e moldar suas vidas; onde instituições poderosas são compreensíveis e responsáveis; e onde informações de pesquisa vitais que podem nos ajudar a enfrentar desafios como a pobreza e as mudanças climáticas estão disponíveis para todos. (OKF, s.d., tradução nossa).

Devido à evolução continuada das TIC, as etapas da produção dos dados, tal como aquelas referente a seu tratamento, armazenamento, compartilhamento, disseminação e uso ético, precisaram ser detalhadamente descritos e analisados (SAYÃO; SALES, 2014). Geralmente os pesquisadores investem tempo considerável coletando e analisando vastas quantidades dados, então a sua gestão se torna essencial para garantir a eficiência, viabilidade, produtividade da pesquisa. Nesse caso pode-se citar o Sistema de Gerenciamento

²OPEN KNOWLEDGE INTERNATIONAL. **About**. Disponível em: <https://okfn.org/about/>. Acesso em: 01 mar. 2018

de Informação Laboratorial (LIMS), um *software* que funciona para melhorar o desempenho de um laboratório, podendo gerir documentos, prover a entrada de dados manual ou automaticamente, auxiliar na elaboração de relatórios, etc., sendo um tipo de sistema conhecido principalmente nas ciências exatas e suas funcionalidades podem variar. Segundo Tolle, Tansley e Hey (2011, p. 21) é um sistema que reorganiza a informação no sentido de organizar os dados para poder colocá-los numa base de dados onde seria possível sua publicação para as pessoas em geral.

Além de *softwares* e ferramentas, tem havido uma proliferação de princípios e diretrizes relativos ao ciclo de gestão de dados abertos. Por exemplo, os Princípios de Panton (*Panton Principles*), elaborado em 2010 por um grupo de pesquisadores da Universidade de Cambridge na Inglaterra, recomendam algumas diretrizes para que os dados científicos realmente sejam abertos, entre elas:

- Ao publicar seus dados, pesquisadores deveriam deixar explícitos seus desejos em relação ao reuso dos dados numa declaração clara;
- Utilizar uma licença, ou uma renúncia de direitos autorais, apropriadas para dados;
- Para reutilizar e reusar os dados como base para a geração de novos dados, devem seguir a definição de aberto pela *Open Knowledge*.
- Os dados devem ser colocados em domínio público por meio de licenças como a *Public Domain Dedication Licence* ou *Creative Commons*. (MURRAY-RUST et al, 2010).

Também há os Princípios *FAIR - Findable, Accessible, Interoperable e Reusable* – elaborados em 2014 pela *Oficina Lorenz* e publicados pela *Future of Research Communications and e-Scholarship – FORCE11*. Em português as quatro propriedades podem ser traduzidas por: Encontrável, Acessível, Interoperável e Reutilizável, que representam facetas prescritivas para os dados abertos, e determinam que:

1. Para ser encontrado, qualquer objeto de dados deve ser único e persistentemente identificável;
2. Os dados devem ser acessíveis na medida em que podem ser sempre obtidos por máquinas e seres humanos;
3. Os objetos de dados podem ser interoperáveis somente se os dados estiverem acionáveis por máquina e os formatos de dados utilizarem vocabulários compartilhados e/ou ontologias, sendo sintaticamente reproduzíveis e semanticamente acessíveis à máquina;
4. Para que os objetos de dados sejam reutilizáveis, os critérios adicionais são: estar em conformidade com os princípios 1 e 3, estar suficientemente bem descritos e ricos para que possam ser automaticamente (ou com o mínimo esforço humano) ligados ou integrados. (FORCE 11, s.d.).

Destaca-se aqui o trabalho da *Open Science Framework* (OSF), uma proposta efetiva de boas práticas científicas, oferecendo uma solução de como é possível realizá-las. Esta sistemática apresenta três emblemas, ou selos, que são: *Open Data*, *Open Materials* e *Preregistered*, como vistos na Figura 1 que segue:

Figura 1. Emblemas desenvolvidos pela OSF para boas práticas na produção científica.



Fonte: OSF (2016).

O selo *Open Data* é obtido quando se disponibiliza publicamente os dados digitalmente compartilháveis e sendo necessários para reproduzir os resultados reportados, devendo ser disponibilizados em um repositório considerado de acesso aberto, com a descrição dos dados e a aplicação de licenças abertas. Já o selo *Open Materials* é obtido disponibilizando publicamente os componentes da metodologia de pesquisa necessária para reproduzir o procedimento e a análise relatados, e os componentes devem ser descritos para que se entenda a relação dos materiais com a metodologia relatada. Por último, o *Preregistred* seria um planejamento de como a pesquisa será conduzida, o *design* do método dela, hipóteses, tamanho da amostra, descrição detalhada das variáveis, entre outras informações (OSF, 2019), todo pré-registrado num sistema institucional, de acesso aberto (por exemplo o *ClinicalTrials.gov*). Este último selo assemelha o registro de pesquisas na Plataforma Brasil, do Ministério da Saúde, embora esta plataforma ainda não seja de acesso aberto, por às vezes conter informações sigilosas sobre participantes humanos.

Tal como acontece com o acesso aberto às publicações, a elaboração e implementação de uma política de dados abertos requer uma mudança que influencia no modo de se disseminar os resultados das pesquisas científicas, e é um desafio também aos pesquisadores. Scheliga e Friesike (2014) levantaram alguns obstáculos da Ciência Aberta baseados em entrevistas que os autores realizaram com pesquisadores, dividindo os obstáculos em individuais e sistêmicos. Nos obstáculos individuais, eles apontam que “[...] o medo da livre circulação é um tópico recorrente nas entrevistas. Os cientistas temem que se liberarem seus materiais de pesquisa no início do processo de pesquisa, eles se expõem ao abuso de propriedade intelectual”. (SCHELIGA; FRIESIKE, 2014, tradução nossa). Já nos obstáculos sistêmicos: “Muitos pesquisadores afirmaram que os esforços científicos abertos não são recompensados pelo sistema acadêmico atual”. (SCHELIGA; FRIESIKE, 2014, tradução nossa), o que aponta para a avaliação da produtividade científica ainda ser muito pautada na publicação de artigos, apenas, e principalmente em revistas com altos fatores de impacto que frequentemente, não são acesso aberto no primeiro momento de publicação. Mas também houve respostas positivas no referido estudo: “Muitos pesquisadores entrevistados expressaram seu endosso à ideia de ciência aberta. Eles descrevem que o compartilhamento de conhecimento através da Internet lhes proporciona oportunidades ampliadas de intercâmbio, *feedback* e colaborar com cientistas internacionalmente.” (SCHELIGA; FRIESIKE, 2014, tradução nossa).

Após essa delimitação do contexto sobre Ciência Aberta e dados abertos, demonstra-se de fundamental importância a elaboração de um Plano de Gestão de Dados (PGD), esclarecendo como é possível realizá-lo a partir de uma compreensão sobre sua relevância no meio acadêmico, resultando em boas práticas científicas.

4 Plano de Gestão de Dados

O Plano de Gestão de Dados (PGD), em inglês Data Management Plan, é um documento que visa descrever o tratamento dos dados durante um projeto de pesquisa e o que ocorrerá com esses dados após a finalização da pesquisa; ou seja, esses planos devem abordar o ciclo de vida dos dados, desde sua descoberta, coleta e organização até como serão preservados. (MICHENER, 2015, tradução nossa). O *template* do PGD geralmente será desenvolvido e sugerido pelas instituições de ensino, órgãos públicos e agências de fomento à pesquisa, que por sua vez, exigem que seus pesquisadores incluam o PGD no seu projeto, que detalha o ciclo de vida dos dados da sua pesquisa, visando, dessa forma, desenvolver um padrão que possa ser disponibilizado com detalhes de como a gestão de dados será realizada, constituindo assim, uma prática na comunicação científica.

Michener (2015, tradução nossa) apresenta dez diretrizes de como criar um bom PGD:

1. Determinar os requisitos do patrocinador de pesquisa;
2. Identificar os dados a serem coletados;
3. Definir como os dados serão organizados;
4. Explicar como os dados serão documentados;
5. Descrever como a qualidade dos dados será assegurada;
6. Apresentar uma estratégia sólida de armazenamento e preservação de dados;
7. Definir as políticas de dados do projeto;
8. Descrever como os dados serão divulgados;
9. Atribuir funções e responsabilidades;
10. Preparar um orçamento realista.

As regras citadas acima realmente precisam ser consideradas ao elaborar um PGD de acordo com as políticas estabelecidas pela instituição na qual o pesquisador está vinculado, pois servem como um excelente guia, evitando assim equívocos e oferecendo o que é essencial para compor o PGD de forma compreensível. Também é importante observar que os dados a serem disponibilizados devem estar certificados com uma licença aberta para garantir que possam ser reutilizados de modo legal.

O *Institute for Social Research* (ICPSR) da Universidade de Michigan nos EUA, visando que muitas agências de financiamento começaram a exigir PGDs, fornece orientações para auxiliar os pesquisadores em como criá-los. No documento “*Guidelines for Effective Data Management Plans*” (2012), apresentam-se algumas recomendações do que deve ser considerado ao desenvolver um PGD, além de apresentar exemplos, quais recursos podem ser utilizados e algumas outras informações. O que deve ser considerado, segundo

essas recomendações, são: descrição dos dados, acesso e compartilhamento, repositório de domínio como o ICPSR, os metadados, direito de propriedade intelectual, ética e privacidade; formato, arquivamento e preservação, armazenamento e *backup*, segurança, responsabilidade, dados existentes, períodos de seleção e retenção; audiência; organização de dados, garantia da qualidade, despesas e requisitos legais.

A elaboração de um PGD ocorre por diversos motivos, e um dos mais importantes atualmente é incluir boas práticas no meio científico. Quando nos referimos às “boas práticas”, se pressupõe que nessa gestão haja uma atuação responsável garantindo assim confiabilidade e integridade.

Analisando as diretrizes e recomendações delineadas anteriormente, são apresentados a seguir alguns dos elementos/componentes que podem ser considerados como fundamentais ao elaborar um PGD, visto que aparecem de forma recorrente em diversos exemplos pesquisados aqui. São eles:

Identificar quais tipos de dados presentes na pesquisa e sob quais licenças devem ser “protegidos”: Esse tópico pode variar muito dependendo da área de conhecimento, pois há diferentes tipos de dados, como: observacionais, computacionais, experimentais, etc. Dessa forma é necessário observar quais dados serão trabalhados para poder escolher como será sua preservação, sendo importante nesse quesito escolher uma licença correspondente para obtenção de *copyright* e que futuramente os dados possam ser reutilizados com a devida responsabilidade.

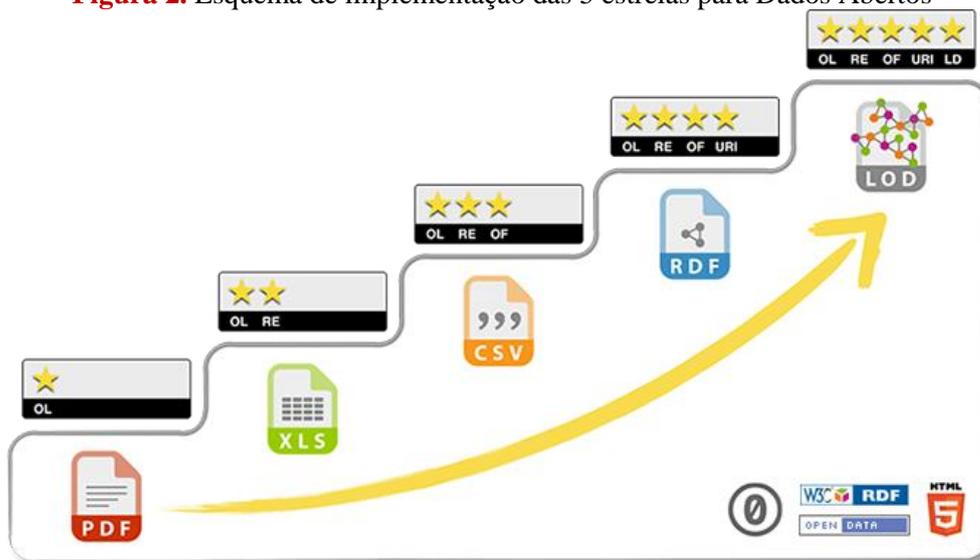
Formato dos dados: Nesse item, tem-se como objetivo que esses dados possam estar abertos. Como citado anteriormente, um dos propósitos dos dados abertos é que possam ser reutilizados. Os formatos de dados (textual, numérico, multimídia, linguagem de software, etc.) podem variar e com isso, sua preservação também se modifica, além disso, na era da “informação digital” é necessário que possam ser interpretados por máquinas, para garantir a sua preservação.

Mesmo considerando a compatibilidade retrospectiva de muitos pacotes de *software* - que permite que dados criados em versões anteriores sejam lidos em versões atuais do software - e a interoperabilidade entre softwares concorrentes, a opção mais segura para garantir o acesso de longo prazo é converter os dados para formatos padronizados. (SAYÃO; SALES, 2015, p. 59).

Este item também nos lembra das “5 estrelas dos dados abertos”³, esquema proposto por Tim Berners-Lee, e que tem como objetivo apontar o quanto os dados estão disponíveis e conectados na *web*, conforme demonstrado na figura 4.

³ Disponível em: <<https://5stardata.info/pt-BR/>>. Acesso em: 26 nov. 2018.

Figura 2. Esquema de implementação das 5 estrelas para Dados Abertos



Fonte: HAUSENBLAS (2012).

Nesse esquema tem-se que a partir de uma estrela se pressupõe que os dados estejam disponíveis na *Web* não importando seu formato, mas com uma licença aberta. Com duas estrelas, os dados estão mais estruturados para serem legíveis por máquina, e com três estrelas, já há a possibilidade de um formato não-proprietário. Quatro estrelas indicam que foram utilizados URIs para identificação de recursos, e finalmente, cinco estrelas sinalizam que além dos dados estarem legíveis por máquinas, também se tem a possibilidade de se conectar trazendo interoperabilidade.

Metadados: Os metadados são importantes, pois têm a função de descrever os dados: como foram utilizados na pesquisa, por quais caminhos passaram, se tenham sido processados e como.

Perceber os benefícios dos dados abertos requer uma comunicação eficaz através de uma abertura mais inteligente: os dados devem estar acessíveis e prontamente localizados; eles devem ser inteligíveis para aqueles que desejam examiná-los; os dados devem ser avaliados para que possam ser feitos julgamentos sobre sua confiabilidade e a competência daqueles que os criaram; e eles devem ser utilizáveis pelos outros. Para que os dados atendam a esses requisitos, eles devem ser suportados por metadados explicativos (dados sobre dados). (ROYAL SOCIETY, 2012, p.7, tradução nossa).

Os metadados também são referidos como a documentação dos dados, a elaboração da qual é considerada como “uma das melhores práticas na criação, organização e gestão de dados, além de ser uma estratégia importante para a preservação digital dos dados. (SAYÃO; SALES, 2015, p. 27).

Escolha de um repositório ou banco de dados para arquivar os dados: Existem atualmente diversos repositórios para armazenar dados. Alguns recomendados na literatura consultada foram: <https://figshare.com/> e <https://datahub.io>. Para pesquisar mais repositórios,

indica-se o *Registry of Research Data Repositories* (www.re3data.org), que é um cadastro global de repositórios de dados de pesquisadores, órgãos financiadores, editores e instituições acadêmicas. Foi fundado em 2012 e é financiado pela Fundação Alemã de Pesquisa (DFG). É importante citá-lo, pois é uma ferramenta notável para identificar qual repositório selecionar para os dados de pesquisa, sendo que seu conteúdo é licenciado sob uma Licença Internacional *Creative Commons*, e é de suma relevância para o conceito de dados abertos. O *Data Portals* (<http://dataportals.org>) também é uma opção similar.

Analisando todo o contexto apresentado, há uma compreensão de que os PGDs têm uma função importante na comunicação científica e influencia o relacionamento entre autores e organizações. Pode-se considerar que o PGD é uma atividade que auxilia as boas práticas científicas, pois propõe que os dados sejam utilizados e disponibilizados de forma coerente com as novas circunstâncias, advindas do meio digital, de produção científica que estamos vivenciando atualmente. O *DMPTool* é um serviço *online* que disponibiliza informações para auxiliar os pesquisadores a preparar um Plano de Gestão de Dados, sendo gratuito e de código aberto. Além disso, o *site* também oferece um suporte que visa que o pesquisador cumpra os requisitos das agências financiadoras. O desenvolvimento do *DMPTool* foi idealizado devido à percepção de demanda por PGDs exigidos pelas agências financiadoras de pesquisa, e então surgiu em 2011, com a ajuda de instituições contribuintes. Dessa forma, pode-se encontrar no site vários modelos de PGDs fornecidos pelas instituições. Outra ferramenta que também oferece modelos de planos de gerenciamento de dados, é o *DMPonline*, criado em parceria com o Centro de Curadoria Digital (DCC) e pelo Centro de Curadoria da Universidade da Califórnia (UC3) a fim de propor uma colaboração entre os financiadores e universidades, visando fornecer os PGDs atualizados e que atendam aos requisitos necessários em todo o ciclo de vida de projetos científicos.

5 A Ciência da Informação nas Políticas de Dados Abertos Científicos

Os dados de pesquisa no contexto digital, que se têm atualmente, geram um fluxo que cresce exponencialmente, sendo difícil seu armazenamento de maneira correta e reutilizável para novos projetos, a gestão de dados mostra-se então importante no âmbito científico. Visto isso, há novas perspectivas na produção científica que exige uma gestão de dados eficiente. Sayão e Sales (2012, p. 182) citam um quarto paradigma:

No quarto paradigma temos a ciência unificando experimentos, teorias e simulações, através do uso intensivo de dados capturados por instrumentos cada vez mais sofisticados ou gerados por simulação, processados por software e armazenados em computadores na forma de bases de dados.

O termo “Curadoria Digital” surge nesse cenário de gestão de recursos digitais como uma prática que têm como finalidade a preservação e o acesso, envolvendo um conjunto de técnicas. (SAYÃO; SALES, 2012). Os dados de pesquisas exigem cada vez mais a elaboração de um PGD, como abordado anteriormente. Dessa forma, pode-se compreender

que a gestão de dados científicos e a *e-Science* também estão relacionados com a Curadoria Digital.

De uma forma geral, a curadoria de dados científicos adiciona velocidade ao ciclo da comunicação científica na medida em que oferece aos pesquisadores dados prontos para o reuso, ou seja, dados tratados, acompanhados por metadados semânticos e estruturais – que asseguram a fidedignidade de seu significado e a reconstrução correta de sua apresentação, somados a metadados que asseguram a integridade, precisão e autenticidade. (SAYÃO; SALES, 2012, p. 188).

Recomenda-se que os dados científicos sejam armazenados em repositórios de dados específicos, caso a instituição ofereça a opção de um repositório digital ou pode-se realizar uma busca de qual pode ser mais apropriado dependendo da pesquisa. Realizando uma busca pelo Brasil no *re3data.org*, foi possível encontrar 8 repositórios de dados recomendados: Banco de Dados de Exploração e Produção; *WorldClim - Global Climate Data*; *GLOBE (Global Collaboration Engine)*; *IODP (International Ocean Discovery Program)*; Repositório de Dados de Levantamentos Biológicos (*PPBio Data Repository*); Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) *Dataverse Network*; Base de Dados Científicos da Universidade Federal do Paraná e Centro de Documentação e Acervo Digital da Pesquisa da Universidade Federal do Rio Grande do Sul (*CEDAP Research Data Repository*).

As publicações científicas conduzem a um tema também muito importante para as boas práticas na comunicação científica que envolvem os profissionais da informação: a atenção ao direito autoral.

Dentre as diversas práticas acadêmicas e normas éticas e jurídicas que estruturam a prática da publicação científica, o direito autoral é hoje tema fundamental para a compreensão dos desafios da comunicação científica e, sobretudo, do papel que as bibliotecas têm nessa missão. (SIQUEIRA, 2015, p.30).

O direito autoral influencia as estratégias de inovação, com o movimento de dados abertos e desenvolvimento de gestão desses dados no meio científico torna-se um assunto essencial nos dias atuais, principalmente pelo fato de a produção científica estar inserida no meio digital. Segundo Siqueira (2015, p. 31):

O direito autoral se funda na ideia de que ao autor de uma obra deve ser concedido o privilégio da exclusividade da exploração de seus proveitos econômicos, bem como o direito ao reconhecimento moral de sua autoria, por um período limitado de tempo, de modo que sirva de incentivo à criação e de estímulo à circulação de obras intelectuais.

Com o desenvolvimento de gestão de dados científicos abertos, a comunicação científica é beneficiada trazendo a oportunidade de melhorias no conhecimento gerado pelas pesquisas. Dessa forma, o direito autoral auxilia que os pesquisadores compartilhem seus dados de forma mais segura, podendo mostrar-se também como uma barreira considerando algumas obrigações e a falta de uma lei mais específica nesse sentido. “Há em parte da

comunidade acadêmica e científica o sentimento de que se perdeu o equilíbrio entre o acesso ao conhecimento e a proteção ao autor”. (SIQUEIRA, 2015, p. 32).

Serviços de dados de pesquisa, em inglês *Research Data Services* (RDS), começaram a surgir nas bibliotecas acadêmicas devido ao aumento de exigências para gerenciar dados e seu compartilhamento por órgãos financiadores. (TENOPIR et al., 2014). Tenopir et al. (2014) abordam a necessidade de gerenciamento de dados e como as bibliotecas estão preparadas ao desenvolver e planejar o RDS, sendo que as políticas da biblioteca necessitam estar alinhadas também com a percepção dos bibliotecários, possibilitando assim um bom serviço de RDS para a comunidade acadêmica. Focando em bibliotecas de pesquisa acadêmicas norte-americanas, os autores levantam diversas questões de como esse RDS vêm sendo realizado e oferecido, sendo que a biblioteca pode oferecer serviços mais informativos e consultivos como auxiliar os pesquisadores a encontrar repositórios para depositar os dados e indicando exemplos de PGDs, ou também serviços mais práticos como atuar no repositório da instituição e ajudar na escrita de PGDs. O foco principal do estudo é saber quais bibliotecas (nos EUA e Canadá) já oferecem serviços relacionados à gestão de dados e também identificar se os bibliotecários estão inseridos nesse meio, além de comparar a política da biblioteca sobre o RDS e as percepções dos bibliotecários à medida que esse é implementado. Os autores concluem que:

É claro que algumas bibliotecas de pesquisa acadêmica estão oferecendo uma variedade de serviços de gerenciamento de dados de pesquisa e mais planos para fazê-lo nos próximos dois anos. Mais comumente, esses serviços são extensões de serviços informativos ou consultivos tradicionais, como ajudar professores e alunos a localizar conjuntos de dados ou repositórios. Um número pequeno, mas crescente, de bibliotecas está cada vez mais envolvido com dados de pesquisa, desde a ajuda com planos de gerenciamento de dados até a preparação e preservação de dados de pesquisa para depósito em repositórios de dados. (TENOPIR et al., 2014, p. 89, tradução nossa).

Research Data Librarian é um termo que têm aparecido em algumas descrições de vagas de emprego para bibliotecários. Uma vaga publicada na *London School of Economics and Political Science* (LSE) em maio de 2018, define em sua descrição que o profissional: “[...] é responsável por fornecer consultoria especializada e treinamento para pesquisadores e funcionários tanto no uso de recursos de dados de pesquisa quanto na criação e preservação de seus próprios dados de pesquisa.” (LSE, 2018). É perceptível a tendência dos profissionais da informação estarem cada vez mais envolvidos com os dados científicos abertos, pois se trata de um assunto que necessita de uma gestão e elaboração de práticas para sua recuperação e aproveitamento futuros. Assim como as bibliotecas podem ser inseridas em projetos de dados abertos, conforme a opinião de Ferrer-Sapena; Peset; Alexandre-Benavent:

Nossa opinião é que as bibliotecas não podem permanecer indiferentes ao movimento de dados abertos. Ao longo dos últimos anos, os padrões abertos, o software de código aberto, o acesso gratuito a publicações e atualmente o acesso gratuito aos dados foram tratados quase que diariamente. Um exemplo em que seu papel ativo pode ser visto é o projeto DISC-UK Datashare (*Data Information Specialist Committee*) do Reino Unido, cujo objetivo é que as bibliotecas de

pesquisas do país atuem como preservadoras dos dados que suportam a gestão e atividades de pesquisa através da criação de repositórios abertos institucionais e tecnologias web 2.0. (FERRER-SAPENA; PESET; ALEIXANDRE-BENAVENT, 2011, p. 268, tradução nossa).

No “Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores”, os autores indicam que o guia também é direcionado para os bibliotecários, pois:

Os bibliotecários estão bem posicionados para trabalhar com os dados pelos seus conhecimentos em gestão de informação, metadados, descoberta de recursos, preservação digital, além disso, eles sempre estabeleceram uma relação longa e produtiva com os pesquisadores. (SAYÃO; SALES, 2015, p.6).

Após a apresentação dessas circunstâncias fica notável que bibliotecários podem atuar no gerenciamento de dados, principalmente nas instituições de ensino superior, promovendo suas competências relacionadas à representação, recuperação e ao acesso das informações, além de se mostrarem atualizados sobre as demandas de acesso aberto no meio digital.

6 Considerações Finais

A partir do levantamento realizado sobre os conceitos de Ciência Aberta, dados científicos abertos e Planos de Gestão de Dados, foi possível observar que realmente há uma consciência por parte das instituições de ensino sobre organizar dados de pesquisa e disponibilizá-los utilizando e exigindo dos pesquisadores os PGDs, porém é necessário que haja também um auxílio dessas instituições, seja em orientações ou apresentando modelos prontos e ainda um desenvolvimento dessa cultura das iniciativas *Open* para que seja mais difundida e propague uma confiança maior aos pesquisadores.

Também é importante destacar o fato de que bibliotecários podem atuar nessa área, sendo que os PGDs envolvem decisões de como gerir os dados pesquisa, ou seja, o profissional da informação pode exercer uma função de suporte informacional ao usuário e ainda exercer suas habilidades ativamente nas instituições de ensino, podendo fazer parte de programas científicos e das políticas dos repositórios digitais das mesmas.

Conclui-se que é necessária uma política de gerenciamento de Plano de Dados científicos, com prioridade em sua abertura para reuso de forma ética. Dessa forma, seria interessante que houvesse um acordo entre instituições para uma possível padronização, porém respeitando também as áreas do conhecimento, visto que nem todos os dados são gerados do mesmo modo. Além disso, fica evidente que há um novo paradigma na publicação e divulgação científica, o que influencia também sua comunicação, seja com o meio acadêmico ou com a sociedade.

Referências

ALBAGLI, S.; APPEL, A. L.; MACIEL, M. L. L. E-science, ciência aberta e o regime de informação em ciência e tecnologia. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 7, n. 1, 2014. Disponível em: <http://ridi.ibict.br/bitstream/123456789/854/1/124-540-1-PB.pdf>. Acesso em: 19 abr. 2018.

ALBAGLI, S. Ciência aberta em questão. In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília; Rio de Janeiro: IBICT; UNIRIO, 2015. p. 9-26.

ALBAGLI, S.; CLINIO, A.; RAYCHTOCK, S. Ciência Aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, p.434-450, 5 dez. 2014. Disponível em: <http://revista.ibict.br/liinc/article/view/3593>. Acesso em: 26 nov. 2018.

AVENTURIER, P. **Plano de Gestão de Dados**: uma introdução. 2017. Disponível em: <https://publicient.hypotheses.org/1660>. Acesso em: 07 de mar. 2018.

BOAI. Dez anos da Iniciativa de Budapeste em Acesso Aberto: a abertura como caminho a seguir. 2012. Disponível em: <http://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese-brazilian-translation>. Acesso em: 19 jun. 2018.

CHAN L.; OKUNE, A.; SAMBULI, N. O que é ciência aberta e colaborativa, e que papéis ela poderia desempenhar no desenvolvimento? In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília; Rio de Janeiro: IBICT; UNIRIO, 2015. p. 91-120.

DELFANTI, A.; PITRELLI, N. Ciência aberta: revolução ou continuidade? In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília; Rio de Janeiro: IBICT; UNIRIO, 2015. p. 59-70.

DUDZIAK, E. A. **Gestão de dados de pesquisa**: o que precisamos saber hoje! 2018. Disponível em: <http://www.sibi.usp.br/noticias/gestao-de-dados-de-pesquisa-o-que-precisamos-saber-hoje/>. Acesso em: 17 jan. 2018.

FECHER, B.; FRIESIKE, S. **Open Science: One Term, Five Schools of Thought**. 2014. p.17-47.

FERRER-SAPENA, A.; PESET, F.; ALEIXANDRE-BENAVENT, R. Acceso a los datos públicos y su reutilización: open data y open government. **El profesional de la información**. p. 260-269, 2011, v. 20, n. 3.

FURNIVAL, A. C. M.; SILVA-JEREZ, N. S. Percepções de pesquisadores brasileiros sobre o acesso aberto à literatura científica. **Inf. & Soc.:est.**, João Pessoa, v. 27, p.153-166, 2017. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/32667/pdf>. Acesso em: 18 jun. 2018.

FORCE 11. **Guiding Principles For Findable, Accessible, Interoperable And Re-Usable Data Publishing Version B1.0**. Disponível em: <https://www.force11.org/fairprinciples>. Acesso em: 30 out. 2018.

HYLA SOFT. **Sistemas de Gerenciamento de Informação Laboratorial (LIMS)**. Disponível em: https://www.hylasoft.com/pt_br/solution/laboratory-information-management-system-lims. Acesso em: 30 out. 2018.

ICPSR. **Guidelines for Effective Data Management Plans**. 2012. Disponível em: <https://www.icpsr.umich.edu/files/datamanagement/DataManagementPlans-All.pdf>. Acesso em: 04 abr. 2018.

ICPSR. **Data Management & Curation**. Disponível em: <https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/index.html>. Acesso em: 04 abr. 2018.

MACHADO, J. Dados abertos e ciência aberta. In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília; Rio de Janeiro: IBICT; UNIRIO, 2015. p. 201-228.

MICHENER, W, K. Ten Simple Rules for Creating a Good Data Management Plan. **PLOS Computational Biology**. 2015. Disponível em: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004525>. Acesso em: 07 mar. 2018.

MOLLOY, J. C. The Open Knowledge Foundation: Open Data Means Better Science. **PLoS Biology**, United Kingdom; v.9, december, 2011.

OPEN DATA COMMONS. **Legal tools for Open Data**. Disponível em: <https://opendatacommons.org/licenses/>. Acesso em 20 ago. 2018.

OPEN DATA GOVERNMENT WORKING GROUP. **8 Principles of Open Government Data**. 2007. Disponível em: https://public.resource.org/8_principles.html. Acesso em: 18 fev. 2018.

OPEN DEFINITION. **The Open Definition**. Disponível em: <http://opendefinition.org>. Acesso em: 29 mar. 2018.

OPEN KNOWLEDGE BRASIL. **Guia de Dados Abertos**. Disponível em: http://opendatahandbook.org/guide/pt_BR/. Acesso em: 01 mar. 2018.

OPEN KNOWLEDGE INTERNATIONAL. Open Definition. **Conformant Licenses**. Disponível em: <http://opendefinition.org/licenses/#Dat>. Acesso em 20 ago. 2018.

OPEN SCIENCE FRAMEWORK. Disponível em: <https://osf.io/tvyxz/wiki/home/>. Acesso em: 15 fev. 2018.

PANTON PRINCIPLES. **Principles for open data in science**. Elaborado por: Murray-Rust, Peter; Neylon, Cameron; Pollock, Rufus; Wilbanks, John; (19 Feb 2010). Disponível em: <https://pantonprinciples.org>. Acesso em: 29 de mar. 2018.

PARRA, H. Z. M. Ciência cidadã: modos de participação e ativismo informacional. In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília; Rio de Janeiro: IBICT; UNIRIO, 2015. p. 121-142.

ROYAL SOCIETY. **Science as an open enterprise**. London: The Royal Society, 2012.

RE3DATA.ORG. **About**. Disponível em: <https://www.re3data.org/about>. Acesso em: 08 maio 2018.

RE3DATA.ORG. **Browse by country**. Disponível em: <https://www.re3data.org/browse/by-country/>. Acesso em: 20 ago. 2018.

RESEARCH INFORMATION NETWORK. NATIONAL ENDOWMENT FOR SCIENCE, TECHNOLOGY AND THE ARTS. **Open to All? Case studies of openness in research**. 2010. Disponível em: <http://www.rin.ac.uk/our-work/data-management-and-curation/open-science-case-studies>. Acesso em: 08 fev. 2018.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.3, 2012, p. 179-191.

SAYÃO, L. F.; SALES, L. F. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **RECIIS – Rev. Eletron. de Comun. Inf. Inov. Saúde**. 2014. p.76-92.

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: <http://www.cnen.gov.br/component/content/article?id=160>. Acesso em: 01 jun. 2018.

SHELIGA, K; FRIESIKE, S. **Putting open science into practice: a social dilemma?** First Monday, v. 19, 2014.

SIQUEIRA, L. P. B. P. Direitos autorais e comunicação científica: desafios para bibliotecas. **Bibliotecas Universitárias: pesquisas, experiências e perspectivas**, v. 2, n. 1, 2015. Disponível em: <https://www.brapci.inf.br/v/a/21564>. Acesso em: 21 ago. 2018.

TENOPIR, C. et al. Research data management services in academic research libraries and perceptions of librarians. **Library & Information Science Research**, v. 36, n. 2, p.84-90, abr. 2014. Elsevier BV. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0740818814000255>. Acesso em: 20 ago. 2018.

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE. LSE. **Job description**. Research Data Librarian. 2018.

TOLLE, K.; TANSLEY, S.; HEY, T. (Org.). Jim Gray e a eScience: um método científico transformado. In: HEY, T.; TANSLEY, S.; TOLLE, K. (Org.). **O quarto paradigma: descobertas científicas na era da eScience**. São Paulo: Oficina de Textos, 2011. p. 17-29.

