

Repositório de Dados de Pesquisa para Grupo de Pesquisa: Um Projeto Piloto

Silvana Aparecida Borsetti Gregorio Vidotti

Universidade Estadual Paulista Júlio Mesquita Filho– UNESP, E-mail: silvana.vidotti@unesp.br

Caio Saraiva Coneglian

Universidade Estadual Paulista Júlio Mesquita Filho– UNESP, E-mail:

caio.coneglian@gmail.com

Sandra Milena Roa Martinez

Universidad del Cauca – UNICAUCA e Universidade Estadual Paulista Júlio Mesquita Filho – UNESP, E-mail: smroa@gmail.com

Felipe Augusto Arakaki

Universidade Estadual Paulista Júlio Mesquita Filho – UNESP, E-mail: fe.arakaki@gmail.com

Mariana Baptista Brandt

Universidade Estadual Paulista Júlio Mesquita Filho– UNESP, E-mail:

marianabrandt@gmail.com

Ana Maria Jensen Ferreira da Costa

Universidade Estadual Paulista Júlio Mesquita Filho – UNESP, E-mail:

anajfcoferreira@gmail.com

RESUMO

Considerando a importância dos dados na realização de análises de pesquisas e a possibilidade de uso e reuso de conjunto de dados, torna-se relevante utilizar ferramentas que auxiliem a gerenciar os dados de pesquisa coletados no desenvolvimento de projetos de pesquisa, bem como reunir, em um mesmo local, os artigos científicos, trabalhos apresentados em eventos, dissertações e teses, dentre outras publicações. Neste contexto, os repositórios de dados de pesquisa científica podem organizar, armazenar e permitir o acesso e o uso de resultados apresentados em diferentes formatos. Assim, este trabalho teve como objetivo compreender e investigar as características inerentes aos repositórios de dados de pesquisa atrelados aos elementos dos Sistemas de Informação de pesquisas correntes (CRIS - Current Research Information System). Tem-se como objetivo específico, analisar a ferramenta DSpace-CRIS para a implantação de um repositório de dados, garantindo a disponibilização de informações e dados relacionadas aos projetos de pesquisa, aos pesquisadores, às instituições envolvidas e às publicações, dentre outras. Para tanto, buscou-se embasamento teórico sobre Dados de Pesquisa, Metadados e Padrões de Metadados para Dados de Pesquisa, Perfil de aplicação, Curadoria Digital, Plano de Gestão de Dados, Repositórios de Dados de Pesquisa e informações técnicas sobre o DSpace-CRIS. Desse modo, foi possível a implantação de um repositório de dados de pesquisa utilizando o sistema de informação de pesquisas correntes - CRIS, para o Grupo de Pesquisa Novas Tecnologias em Informação. Como resultado, foi possível perceber que as funcionalidades dessa ferramenta atendem as necessidades e as características das publicações de dados de pesquisa como o enriquecimento de uma produção intelectual, permitindo gerenciar dados e informações. Nesse sentido verifica-se que o DSpace-CRIS é uma opção favorável para ser utilizada pelos grupos de pesquisa, que neste repositório permitiu a integração de projetos tanto com as produções intelectuais quanto com os dados coletados durante as pesquisas, possibilitando o (re) uso dos dados de pesquisa, a

interoperabilidade entre diferentes ambientes e a disseminação dos recursos gerados pela comunidade do GPNTI.

Palavras-chave: Repositórios de Dados de Pesquisa. Sistema de Informação de Pesquisas Correntes – CRIS. Dados de Pesquisa. Perfil de Aplicação.

1 INTRODUÇÃO

A evolução e a disseminação do uso das Tecnologias de Informação e Comunicação (TIC) alterou a forma como se dão as interações dos indivíduos com as informações científicas, o que, além de promover o desenvolvimento de ferramentas tecnológicas com uma alta capacidade de processamento e de armazenamento, também facilitou o acesso e uso das publicações científicas. Esse cenário teve consequências imediatas e significativas para as pessoas e para as instituições. Porém, a quantidade de dados de pesquisa dispersos em diferentes contextos e domínios, tornou-se um desafio para o gerenciamento das pesquisas e o reuso desses dados em outras pesquisas.

Diante desses fatos, as instituições de ensino e de pesquisa, começaram a se mobilizar para compreender e implementar ferramentas com a finalidade de gerenciar grandes quantidades de informações e de dados. Iniciativas vinculadas a Ciência da Informação e a Ciência da Computação como *Big Data*, *e-Science* e *Data Science*, vem contribuindo com esta nova configuração da sociedade.

Em comum, identifica-se a importância dos dados para a realização de análises de pesquisas e a possibilidade de utilizar um conjunto total de dados nos processos analíticos, rompendo com a ideia do uso de amostras para compreender um contexto geral.

Desta forma, no âmbito acadêmico, disponibilizar os dados coletados e utilizados em uma pesquisa passa ser tão importante quanto a própria publicação da produção científica, em, por exemplo, artigos, livros, teses, dissertações e trabalhos em eventos científicos. Neste contexto, os repositórios de dados de pesquisa passam a ser de extrema relevância, pois, na maioria dos casos, eles permitirão o armazenamento, a preservação e a difusão dos dados produzidos.

Uma série de iniciativas em todo o mundo, como a PLOS ONE¹ e a Nature² por exemplo, estão incentivando os cientistas a disponibilizarem os dados produzidos em suas pesquisas em

¹ Disponível em: <<http://journals.plos.org/plosone/s/data-availability>>. Acesso em: 06 set. 2017.

² Disponível em: <<https://www.nature.com/sdata/policies/repositories>>. Acesso em: 06 set. 2017.

ambientes de acesso aberto, em especial, daquelas financiadas por órgãos públicos ou agência de fomento. Nesse contexto, destaca-se o Programa Horizonte 2020, promovido pela Comissão Europeia, em que se busca financiar a ciência no continente europeu, exigindo que os dados produzidos estejam disponíveis abertamente. (EUROPEAN COMMISSION, 2012).

No Brasil, a disponibilização dos dados relativos às pesquisas financiadas pelas agências de fomento ainda não é tratada como uma obrigatoriedade. Percebem-se iniciativas como a da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) de solicitar um Plano de Gerenciamento de Dados junto à documentação nas “Chamadas de Proposta de Pesquisa”³, tratando da necessidade de compartilhar os dados e prever seu gerenciamento. Essas iniciativas estão no contexto do Open Data (dados abertos) que tem como propósito - de serem usados, reutilizados e redistribuídos por qualquer pessoa. (OPEN KNOWLEDGE INTERNATIONAL, 2017).

Neste âmbito, considerando a necessidade dos grupos de pesquisa de gerenciar, armazenar, preservar e disseminar os projetos, os dados e as publicações de pesquisa, visando o posterior uso e reuso, questiona-se: Como estruturar e quais ferramentas adotar para um repositório de dados de pesquisa para grupos de pesquisa?

Assim, esta pesquisa tem como objetivo compreender e investigar as características inerentes aos repositórios de dados de pesquisa atrelados aos elementos dos Sistemas de Informação de pesquisas correntes (CRIS - *Current Research Information System*). Tem-se como objetivo específico, analisar a ferramenta DSpace-CRIS para a implantação de um repositório de dados, garantindo a disponibilização de informações e dados relacionadas aos projetos de pesquisa, aos pesquisadores, às instituições envolvidas e às publicações, dentre outras.

Como resultado, apresenta-se o projeto piloto do repositório de dados de pesquisa para o Grupo de Pesquisa Novas Tecnologias em Informação (GPNTI), que congrega pesquisadores da área de Ciência da Informação da linha de Informação e Tecnologia da Universidade Estadual Paulista (Unesp).

³ Programa FAPESP de Pesquisa sobre Mudanças Climáticas Globais - Chamada de Propostas de Pesquisa 2017. Disponível em: <<http://www.fapesp.br/11068>>. Acesso em: 26 jul. 2017.

2 DADOS DE PESQUISA

O conceito de dados de pesquisa, de natureza diversificada e complexa, traz uma ampla gama de significados, podendo variar conforme o tipo de objeto de pesquisa, os métodos de coleta, os pesquisadores e as disciplinas. Assim, para Borgman (2010), é importante definir e reconhecer a diversidade de formas, de modo a promover processos de compartilhamento, publicação e curadoria de dados de pesquisa.

Nesse cenário, o programa Horizonte 2020 (EUROPEAN COMMISSION, 2012), tem como desafio definir os termos “dados de pesquisa” e “repositório de dados de pesquisa”, objetivando validá-los e aplicá-los em diferentes disciplinas acadêmicas.

Os dados de pesquisa são “[...] fatos, números, letras e símbolos que descrevem um objeto, ideia, condição, situação, ou outros fatores.” (COMMITTEE FOR A STUDY ON PROMOTING ACCESS TO SCIENTIFIC AND TECHNICAL DATA FOR THE PUBLIC INTEREST et al., 1999, p. 15, tradução nossa). A *National Science Board* e a *National Science Foundation* (2005) complementam que os dados podem ser qualquer informação armazenável em formato digital, como por exemplo, os elementos supracitados, além de: imagens, vídeos, áudios, *softwares*, algoritmos, equações, animações, modelos e simulações, gerados por diferentes formas, tais como de observação, cálculo ou experimentos.

Para Pampel et al. (2013, p. 1, tradução nossa), o termo “dados de pesquisa” refere-se a:

[...] dados digitais sendo uma parte (descritiva) ou o resultado de um processo de pesquisa. Este processo abrange todas as etapas da pesquisa, que vão desde a geração de dados de pesquisa, que podem ser em um experimento nas ciências, um estudo empírico nas ciências sociais ou observações de fenômenos culturais, até a publicação dos resultados da pesquisa. Os dados digitais de pesquisa são produzidos em diferentes tipos de dados, níveis de agregação e formatos de dados, informados pelas disciplinas de pesquisa e seus métodos.

Além disso, instituições como o *National Institutes of Health* (2003, não paginado) e *Organisation for Economic Co-Operation and Development* (2007, não paginado), reafirmam que os dados de pesquisa são: “[...] todo aquele material que foi gravado durante a pesquisa, reconhecido pela comunidade científica e que serve para certificar os resultados da pesquisa que se realizam [...] que deve vir de uma única fonte.”

Os dados de pesquisa podem estar em variados formatos reconhecidos: numéricos, descritivos, visuais, auditivos ou tácteis, também em formatos de disciplinas ou ferramentas

específicas. Ainda, entre as classificações dos dados de pesquisa existentes, destacam-se as baseadas na origem dos dados, nos métodos de coleta ou no objetivo da coleta, e que foram identificadas por Floridi (2005, 2017) como: 1) dados primários, 2) dados secundários, 3) metadados, 4) dados operacionais e 5) dados derivados. Outra classificação com base na origem dos dados refere-se a dados primários e a dados secundários.

Além disso, conforme o relatório da *National Science Board* e *National Science Foundation* (2005), os dados podem ser categorizados segundo os métodos de coleta em: observacionais, computacionais, experimentais ou registros. A partir dessas categorias, os dados podem ter um valor imediato e duradouro, atingir o valor ao longo do tempo, ter um valor transitório ou ser de fácil geração e que não precisa ser preservado.

Além disso, com respeito às estratégias de publicação dos dados de pesquisa abertamente acessíveis existem (PAMPEL; DALLMEIER-TIESSSEN, 2013):

- a) A publicação dos dados da pesquisa como uma documentação textual conhecido como um artigo de dados;
- b) A publicação de dados de pesquisa como o enriquecimento de um artigo, chamado publicação enriquecida;
- c) A publicação dos dados da pesquisa, como um objeto de informação independente em um repositório de dados de pesquisa.

Independente da estratégia de publicação dos dados gerados durante um determinado projeto de pesquisa, estes dados são, segundo Sayão e Sales (2012, p. 189): “[...] os objetos digitais requerem metodologias de gestão que são muito diferentes das que são utilizadas no universo da impressão tradicional.”. Nesse sentido, a curadoria digital pode contribuir para a gestão desses objetos subsidiando a análise criteriosa, tanto para identificar quanto para indicar elementos a fim de promover o acesso, o uso e a preservação do conteúdo.

3 CURADORIA E PLANO DE GESTÃO DE DADOS

A curadoria digital vai além da preocupação com a preservação dos dados de pesquisa, é uma forma de gestão que deve intervir em todo o ciclo de vida dos dados, iniciando-se com a seleção do conteúdo destinado ao depósito, e acompanhada com a aplicação de metodologias, como por exemplo, métodos de higienização de dados, indicação do melhor formato para armazenamento, formas de permissão de acesso, ações para promover fidedignidade dos dados

por meio de tecnologia, boas práticas de arquivamento e preservação digital dentre outras técnicas.

O ponto crucial da curadoria de dados de pesquisa é que ela precisa estar inserida organicamente nos fluxos de geração de conhecimento científico para ter validade [...]. Isto demanda uma reformulação nos perfis profissionais dos bibliotecários e arquivistas que agora se tornam profissionais de dados.(SAYÃO; SALES, 2016b, p. 70).

A Curadoria digital tem uma abordagem multidisciplinar e segundo o *Digital Curation Centre* (DCC) “Além de reduzir a duplicação de esforços na criação de dados de pesquisa, [...] melhora o valor de longo prazo dos dados existentes, tornando-o disponível para novas pesquisas de alta qualidade.”(DIGITAL CURATION CENTRE, 2017, não paginado, tradução nossa).

Para o DCC, a gestão do ciclo de vida da curadoria digital aborda três tipos de fases ou ações. A primeira está relacionada com ações que permeiam todo o ciclo de vida como a descrição e representação da informação; o planejamento da preservação; a participação e observação da comunidade; a preservação e curadoria. A segunda fase ocorre de forma sequencial como as ações de conceitualização; criação ou recepção; valorização e seleção; ingestão; ação de preservação; armazenamento; acesso, uso e reutilização; e transformação. Por fim, na terceira fase estão ações ocasionais como a migração e a reavaliação.

No que tange aos Planos de Gestão de Dados (PGD), este instrumento se apresenta como um elemento útil e necessário na curadoria dos dados descrevendo aspectos relativos ao ciclo de vida dos dados. Para autores como Sayão e Sales (2015), “[..] ele se constitui em um documento formal que estabelece um compromisso de como esses dados serão tratados durante todo o desenvolvimento do projeto, e também após a sua conclusão.”

Os PGD permitem que os dados possam ser encontráveis, acessíveis, interoperáveis e reutilizáveis (*Findable, Accessible, Interoperable and Re-usable - FAIR*). Para isso, devem ser incluídas informações como:

[...] o tratamento de dados de pesquisa durante e após o final do projeto; -quais dados serão coletados, processados e / ou gerados; -qual metodologia e padrões serão aplicados; -se os dados serão compartilhados / criados de acesso aberto e; - como os dados serão curados e preservados (inclusive após o final do projeto). (EUROPEAN COMMISSION, 2012, p. 4, tradução nossa)

Com respeito às informações que deveriam conter os PGD, há por exemplo, organizações como a *Digital Curation Centre* (2014), que tem disponibilizado um *checklist*; diretrizes para planos efetivos de gerenciamento de dados da *Inter-university Consortium for Political and*

Social Research (2017); e o guia de gestão de dados de pesquisa para bibliotecários e pesquisadores proposto por Sayão e Sales (2015).

4 REPOSITÓRIO DE DADOS DE PESQUISA

Uma estratégia para que os dados de pesquisa possam estar disponíveis e serem recuperados para reuso é o seu armazenamento em um repositório. O termo “Repositórios de Dados de Pesquisa” também é encontrado na literatura como: Repositórios de Dados, Repositórios de Dados Científicos e Repositórios Digitais de Dados de Pesquisa. Em inglês os termos encontrados foram: *Research Data Repositories*, *Scientific Data Repository* e *Data Repository*. Neste trabalho, optou-se pela terminologia “Repositório de Dados de Pesquisa”.

Assim como os Repositórios Institucionais armazenam documentos completos de produções intelectuais, o Repositório de Dados de Pesquisa é a ferramenta necessária para armazenar e gerenciar os dados utilizados e produzidos durante uma pesquisa. Segundo Rodrigues e Saraiva (2010, p. 23): "Nos repositórios de dados pode ir-se [...] além desta visão de repositório de objectos, uma vez que cada conjunto de dados tem características próprias e por isso pode requerer um tratamento diferenciado." Costa e Braga (2016) apresentam um panorama dos repositórios de dados de pesquisa a partir diretório *Registry of Research Data Repositories* (re3data.org) e destacam que é uma tendência na construção de repositórios de dados de pesquisa temático. Exemplos de repositórios de dados de pesquisa são *Dryad Digital Repository*⁴, a iniciativa do Repositório Científico de Acesso Aberto de Portugal (RCAAP)⁵, entre outros.

Complementarmente, Sayão e Sales (2016a, p. 95) afirmam que:

Além de oferecer uma base tecnológica para a execução dos processos de contextualização dos dados, os repositórios têm um papel importante nas interações que envolvem a validação do trabalho de pesquisa e na própria dinâmica social da comunicação científica.

Os repositórios são também ferramentas que registram o andamento da pesquisa por meio dos dados coletados e que podem contribuir para a metodologia científica, por exemplo, dados referentes a experimentos sem sucesso também podem ser registrados, evitando que outros pesquisadores façam a mesma coleta desnecessariamente (SAYÃO; SALES, 2016a). Nos repositórios institucionais, em que somente os produtos finais das pesquisas são registrados, essas

⁴ Disponível em: <<http://datadryad.org/>>. Acesso em: 06 set. 2017.

⁵ Disponível em: <<https://repositorium.sdum.uminho.pt/handle/1822/46094>>. Acesso em: 06 set. 2017.

informações das etapas da pesquisa e da coleta de dados não são gerenciadas e podem ser perdidas.

Sayão e Sales (2016a) destacam como benefícios dos Repositório de Dados de Pesquisa: visibilidade e compartilhamento dos dados, crédito ao autor dos dados, preservação digital, memória científica e transparência, segurança dos dados, disponibilidade, curadoria digital, serviços inovadores, reuso dos dados, redes de repositórios e indicadores de qualidade e produtividade da instituição.

Com relação aos metadados e documentação dos dados de pesquisa, Kindling e Schirmbacher (2013) apontam que são essenciais para permitir sua posterior utilização, pois descrevem o contexto de tais dados e as ferramentas com as quais eles foram gerados, armazenados, processados e analisados.

Segundo Qin, Ball e Greenberg (2012) os metadados são fundamentais para descoberta, controle, qualidade e uso de dados. Entretanto, para garantir que a descrição do recurso informacional seja recuperada e interoperável por diferentes sistemas de informação é necessário estabelecer uma padronização dos metadados. Nesse sentido, o padrão de metadados Dublin Core é uma solução viável para compor uma descrição mínima, uma vez que propõe metadados básicos para localização de recursos no ambiente Web.

5 CRIS - SISTEMA DE INFORMAÇÃO DE PESQUISAS CORRENTES

O gerenciamento dos dados como recursos pode contribuir com o desempenho das instituições de ensino e de pesquisa, das agências de fomento, das empresas e outras organizações interessadas no desenvolvimento da ciência. No contexto da ciência aberta e gestão de dados de pesquisa, além dos repositórios de dados de pesquisa, surge outra tecnologia que vem implementar o gerenciamento de dados, abarcando o conceito denominado Sistema de Informação de Pesquisas Correntes (*Current Research Information System - CRIS*).

Os CRIS são ferramentas para representar o fluxo de informações dos processos de pesquisas:

[...] que possam gerenciar toda a informação relevante da pesquisa, começando com as oportunidades de financiamento, passando pelo estágio de escrita e submissão de propostas, seguindo com as propostas bem sucedidas que então se tornam projetos ativos que serão gerenciados até a conclusão - em que estágio geram resultados, muitos dos quais publicações ou algum outro produto da atividade de pesquisa. (JOINT, 2008).

O DSpace-CRIS insere características de CRIS aos repositórios digitais, sendo uma ferramenta em código aberto (*open-source*) de gestão de pesquisas do repositório DSpace. Existem outras aplicações como o Converi, ferramenta da Thompson Reuters e o Pure, da Elsevier, que apresentam como benefícios o suporte ao gerenciamento do ciclo de vida completo das pesquisas, gerando relatórios analíticos com informações gerenciais para gestores das universidades e outras instituições de pesquisa.

No que diz respeito aos elementos contemplados em um CRIS, este sistema:

[...] contém dados / metadados ou informações sobre gerentes de projetos, projetos em andamento e concluídos, departamentos de pesquisa, organizações de financiamento, programas e financiamentos, pesquisadores, resultados de pesquisas (publicações, patentes, produtos), eventos, instalações, (Semântica) e fornece uma abordagem integrada para gerenciar informações de pesquisa. (EUROCRIS, 2010, não paginado, tradução nossa).

Nesse sentido, por meio dos metadados, um CRIS quando implementado possibilita rastrear outros ambientes informacionais digitais para interligar dados de pesquisa e pesquisadores. Para o gestor da pesquisa, o CRIS fornece uma visão geral das atividades de pesquisa, dentro e fora da sua instituição, possibilitando tomada de decisões estratégicas.

Assim, os repositórios de dados de pesquisa em conjunto com sistemas CRIS possibilitam maior qualidade na gestão da pesquisa científica, beneficiando os pesquisadores, administradores de pesquisa, conselhos de pesquisa, organizações tecnológicas e empreendedoras e o público em geral, pois facilita o acesso à informação científica, além de abarcar questões que tangem ao tratamento dos dados de pesquisa.

Esse acesso é favorecido pelo uso de metadados específicos para dados de pesquisa disponibilizados em repositórios, uma vez que os metadados, além de descreverem os conteúdos, têm função relativa à própria curadoria contida nos planos de gestão de dados e nos elementos de gestão da pesquisa.

Um CRIS pode ser implementado usando elementos do *Common European Research Information Format* (CERIF), criado no ano de 2006 para manter um padrão de interoperabilidade de informações científicas no contexto europeu. O CERIF abrange o nível conceitual, lógico e físico das entidades e suas relações, abrangendo quatro tipos de entidades: 1) Entidades básicas que fornecem a estrutura chave para o CERIF (Projeto, Pessoa, Organização/Unidade); 2) Entidades de resultados que estão estreitamente relacionadas às entidades base e são resultantes do processo de pesquisa (Resultado/Publicação, Resultado/Patente,

ResultadoProduto); 3) Entidades de segundo nível ou secundárias que envolvem as entidades base e resultado para permitir a representação do contexto da pesquisa (Idioma, Evento, Citação, Equipamento, País, Financiamento, entre outros); e 4) Entidades de ligação que são as relações ou links entre duas entidades do CERIF (Pessoa 'é autor de' ResultadoPublicação).

Nesse contexto, o modelo CERIF apresenta-se como alternativa no processo de interoperabilidade entre outros sistemas, pois sua construção baseou-se no Dublin Core, garantindo a troca de informações com repositórios digitais que utilizam o *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH), um mecanismo para a coleta de registros de metadados em repositórios.

Destaca-se que os dados de pesquisa podem estar ligados a diversas fontes como repositórios institucionais e temáticos, que tenham publicações como teses, dissertações, artigos e trabalhos em eventos. Os benefícios da integração com um CRIS, facilitam a localização e a relação de um dado de pesquisa a um projeto de pesquisa e a seus pesquisadores. O uso especificamente do DSpace-CRIS possibilita ainda, relacionar o pesquisador com identificadores persistentes como *Open Researcher and Contributor ID* (ORCID), *ResearcherID* e *ScopusID* entre outros identificadores que forem necessários. Os projetos de pesquisa podem estar relacionados também com as agências de fomento, subsidiando-as para analisar o retorno dos projetos financiados. Outra característica do DSpace-CRIS são as redes de colaboração e estatísticas de citação da *Web of Science*, *Scopus*, *PubMed*, *Google Scholar*, além de estatísticas internas.

6 PROCEDIMENTOS METODOLÓGICOS

Para o desenvolvimento do projeto piloto do Repositório de Dados de Pesquisa do Grupo de Pesquisa Novas Tecnologias em Informação (Repositório GPNTI-CRIS), utilizou-se metodologia qualitativa, com caráter exploratório, devido a busca bibliográfica e documental que contemplam as recomendações teóricas e aplicadas de definições de repositórios de dados de pesquisa. Caracteriza-se também como uma pesquisa aplicada, uma vez que desenvolveu um projeto piloto de repositório, iniciando com a implantação e o povoamento de recursos.

Para implementação do sistema, foi instalado o *software* DSpace-CRIS. Posteriormente, foram definidas e implementadas as características e funcionalidades do sistema e a inserção de um projeto de pesquisa. Vale destacar que o GPNTI possui uma série de produções intelectuais,

em que para comprovar as funcionalidades do *software*, foi escolhido um destes projetos com várias publicações e com resultados de pesquisa em diferentes formatos. Nesse sentido, realizou-se a inserção destes dados no Repositório, visando demonstrar a viabilidade da proposta.

O *software* DSpace-CRIS foi escolhido pois amplia as possibilidades de um *software* de repositório digital com as características de um sistema CRIS. Tal escolha foi motivada pela necessidade de se ter uma plataforma que fosse capaz de gerenciar e de relacionar os dados utilizados e produzidos em uma pesquisa de um projeto específico, e assim abarcasse as produções científicas resultantes deste processo. Como modelos de aplicações foram utilizados o da Universidade de *Hong Kong*⁶, a proposta do PT_CRIS⁷, o Eurocris⁸ entre outras.

Partindo desse pressuposto, buscou-se validar a utilização da plataforma DSpace-CRIS enquanto um sistema para repositórios de dados, que atende as necessidades e as principais características para a disponibilização de dados no âmbito de um grupo de pesquisa. Vale destacar que para a implantação foram utilizados os dados de pesquisadores do grupo. A partir dos procedimentos discutidos, foram realizados o desenvolvimento e a implantação do projeto piloto Repositório GPNTI-CRIS, que será apresentado na próxima seção.

7 RESULTADOS E DISCUSSÕES: Desenvolvimento e Implantação do Repositório GPNTI-CRIS

O Repositório de Dados de Pesquisa do Grupo de Pesquisa Novas Tecnologias em Informação, ou Repositório GPNTI-CRIS, foi criado como um projeto piloto para disponibilizar os dados de trabalhos científicos publicados pelos seus integrantes, buscando seguir a tendência global de disponibilização de dados. Para tanto, foi utilizado o *software* DSpace-CRIS que possibilita a disponibilização dos dados, permitindo a publicação enriquecida e relacionando informações entre os projetos, os pesquisadores, os trabalhos e os resultados de pesquisa.

Dessa forma, a primeira etapa foi a instalação e a disponibilização do ambiente em uma URL disponível abertamente (<http://gpnti.marilia.unesp.br:8084/jspui/>), para isso utilizou-se o manual de instalação do DSpace CRIS produzido por Silveira, Shintaku e Bollini (2016).

A partir da instalação do DSpace CRIS, iniciou-se o processo de customização, em que foram inseridas as características do Grupo de Pesquisas ao qual o repositório está vinculado.

⁶ Disponível em: <<http://hub.hku.hk/>>. Acesso em 02 set. 2017.

⁷ Disponível em: <<https://ptcris.pt/>>. Acesso em 02 set. 2017.

⁸ Disponível em: <<http://dspacecris.eurocris.org/>>. Acesso em 02 set. 2017.

Com a instalação e customização do *software*, iniciaram-se as definições dos processos e das características que deverão estar presentes no projeto piloto do Repositórios de Dados de Pesquisa. Esta fase é necessária pois, o Repositório GPNTI-CRIS deverá contemplar as características tanto de repositórios de dados de pesquisa, quanto aspectos de um CRIS. As características essenciais do repositório estão apresentadas no Quadro 1:

Quadro 1 - Características Repositórios de Dados GPNTI.

Característica	Elemento	Descrição ou Decisão de Escolha
Plataforma	DSpace CRIS	Escolhida por ser uma ferramenta gratuita e possuir código aberto, além do DSpace ser também utilizado no Repositório Institucional da Universidade Estadual Paulista (Unesp)
Representação da informação	Padrão de Metadados/ Perfil de Aplicação	Para definição dos metadados foram utilizados como base elementos do Dublin Core (simples e qualificado), foi considerado também, os elementos propostos no DSpace-CRIS e alguns elementos específicos para a identificação e a recuperação dos dados de pesquisa no Repositório. O CERIF foi utilizado para auxiliar na descrição das entidades relacionadas ao CRIS.
		Ver perfil de aplicação no Quadro 2.
Elementos gerenciados pelo CRIS	Pesquisadores; Instituições; Projetos de pesquisa; Dados e recursos de pesquisa; Coleções; Assuntos	Foram elencados os principais elementos necessários à gestão das pesquisas do grupo. Além disso, elementos semânticos como assunto e coleções foram inseridos para permitir a busca temática e organização por categorias (coleções).
Interoperabilidade	Protocolo OAI-PMH	O protocolo OAI-PMH foi instalado no repositório para que seja possível a interoperabilidade de recursos com outros provedores.
Relação com outras bases	Repositórios Institucionais	Verificação das produções intelectuais dos pesquisadores vinculados ao grupo de pesquisa em Repositórios Institucionais, para que o Repositório GPNTI fique atualizado automaticamente.
Indexação	Google Scholar	O repositório segue os padrões do Google Scholar para que os seus recursos possam ser encontrados por esta base.
Tipologia dos dados de pesquisa	Dados primários	O repositório armazenará os dados específicos coletados e usados para o desenvolvimento das pesquisas, associados aos projetos.
	Observacionais	Os dados de pesquisa armazenados incluem medições, questionários, entre outros.
	Registros	O repositório armazena conjuntos de dados governamentais, públicos e privados.
Conteúdo dos dados de pesquisa	Tipo	Artigos, livros, capítulo de livros, <i>datasets</i> , objetos de aprendizagem, imagens, mapas, partituras, <i>blueprint</i> , slides, áudios, gravações, softwares, relatórios técnicos, teses, vídeo, <i>working paper</i> , outros.
Comunidades e Coleções	Temática	As comunidades e coleções foram definidas a partir da temática utilizada pelos membros do grupo de pesquisa, as coleções foram criadas de acordo com as especificações de cada temática.
Plano de gestão de dados	Dados administrativos; Informações relativas aos processos de coleta dos dados de pesquisa; metadados e documentação;	O plano de gestão de dados está em desenvolvimento, pois as agências de fomento estão discutindo e em vias de propor um plano, que deverá ser adotado pelo grupo de pesquisa. No entanto, os principais guias para a elaboração de planos de gestão (DIGITAL CURATION CENTRE, 2014; INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH,

	aspectos éticos, legais e de responsabilidade; armazenamento; preservação e compartilhamento	2017; SAYÃO; SALES, 2015) apresentam um núcleo comum, composto por: Dados administrativos; Informações relativas aos processos de coleta dos dados de pesquisa; metadados e documentação; aspectos éticos, legais e de responsabilidade; armazenamento; preservação e compartilhamento.
--	--	---

Fonte: Elaborado pelos autores.

Conforme Quadro 1, as principais características analisadas e implantadas no projeto piloto do Repositório GPNTI-CRIS foram: Plataforma; Padrão de Metadados; Perfil de Aplicação; Gestão da pesquisa (CRIS); Interoperabilidade; Relação com outras bases Indexação; Tipologia e Conteúdo dos dados de pesquisa; Comunidades e Coleções; e Plano de gestão de dados, as quais são a base para que um repositório atenda a disponibilização de dados no âmbito de um grupo de pesquisa, abrangendo também os elementos de um CRIS. A seguir apresenta-se o Perfil de Aplicação, no quadro 2, construído para a descrição dos recursos no Repositório GPNTI-CRIS:

Quadro 2 – Perfil de Aplicação GPNTI-CRIS

Metadado	Termo utilizado	Descrição
Método de coleta	dc.collection.method	Método de coleta dos dados.
Orientador	dc.contributor.advisor	Nome do orientador do documento.
Autor	dc.contributor.author	Autor do documento descrito.
ORCID do autor	dc.contributor.authorOrcid	ORCID do autor do documento descrito.
Instituição	dc.contributor.institution	Instituição a qual o autor do recurso está filiado no momento da publicação.
Cobertura espacial	dc.coverage.spatial	Características espaciais do conteúdo.
Cobertura temporal	dc.coverage.temporal	Características temporais do conteúdo.
Data de depósito	dc.date.accessioned	Data de inserção do item no DSpace.
Data de disponibilização	dc.date.available	Data que o item fica disponível para o público.
Data de CopyRight	dc.date.copyright	Data de copyright.
Data criação	dc.date.created	Data criação ou fabricação conteúdo.
Data de publicação	dc.date.issued	Data da publicação ou distribuição.
Data de submissão	dc.date.submitted	Recomendando para Teses e Dissertações.
Data de atualização	dc.date.updated	Data da última atualização.
Resumo	dc.description.abstract	Resumo.
Agência fomento	dc.description.sponsorship	Agência de fomento/financiamento.
Número do processo	dc.description.sponsorshipId	Número do processo.
Versão	dc.description.version	Versão da última edição / atualização.
Extensão	dc.format.extent	Extensão.
Formato	dc.format.medium	Formato do arquivo.
Tipo	dc.format.mimetype	Registro tipo identificador MIME.
Referência	dc.identifier.citation	Referência do documento descrito.
Identificador	dc.identifier	Identificador do registro.
Idioma	dc.language	Idioma.
Publicador	dc.publisher	Publicador.
Processamento	dc.processing	Processamento.
Relação conferência	dc.relation.conference	Relação com uma conferência.
Relação com dataset	dc.relation.dataset	Relação com um dataset.

Faz parte de	dc.relation.haspart	Referências físicas ou logicamente contidas.
Tem versão	dc.relation.hasversion	Referências versão posterior.
É baseado em	dc.relation.isbasedon	Fonte de referências.
É um formato de	dc.relation.isformatof	Referências de forma física adicional.
Faz parte de	dc.relation.ispartof	Referências físicas ou lógica do item.
É referenciado por	dc.relation.isreferencedby	Apontado por recurso referenciado.
É substituído por	dc.relation.isreplacedby	Referências item de sucesso.
É uma versão de	dc.relation.isversionof	Referências versão anterior.
Relação publicação	dc.relation.publication	Responsável por publicar um recurso.
Substituí	dc.relation.replaces	Referências anteriores ao item.
Requer	dc.relation.requires	Recurso suporta função, entrega ou coerência do item.
Licença	dc.rights.license	Informações de licença.
Direitos	dc.rights	Informações sobre os direitos sobre o item.
Fonte	dc.source	Informações sobre a fonte.
Assunto	dc.subject	Palavra-chave.
Título	dc.title	Título do recurso.
Título alternativo	dc.title.alternative	Título alternativo.
Tipo	dc.type	Tipo dos recursos.

Fonte: Elaborado pelos autores.

No que tange a forma como os recursos são inseridos e visualizados no sistema, o repositório GPNTI-CRIS aceita documentos de diferentes formatos, os quais foram definidos em função dos tipos de dados que poderão ser armazenados. Além disso, o usuário, ao cadastrar um conteúdo, escolhe qual é o tipo de recurso que está sendo inserido, como artigo, vídeo e conjuntos de dados.

No projeto piloto, os dados de pesquisa inseridos no repositório correspondem às pesquisas desenvolvidas ou em andamento estando associadas aos projetos do grupo de pesquisa. Estes dados podem ser considerados primários, os quais foram coletados e utilizados para cumprir os objetivos e obter resultados nas pesquisas.

Como funcionalidade do DSpace-CRIS, o sistema possibilita na página do autor, vincular informações com o *Open Researcher and Contributor ID (ORCID)*, além de reunir as produções desenvolvidas por ele. O vínculo com o ORCID permite que informações cadastradas pelo indivíduo sejam inseridas no sistema, possibilitando localizar e escolher autores dos documentos inseridos pelo seu cadastro no ORCID.

A página que apresenta o “Projeto” reúne todas as informações e os recursos referentes a ele. Além disso, ao lado dos nomes dos autores do projeto, existe um ícone em que é possível acessar a página de informações de cada um deles.

A figura 1 apresenta a tela de um projeto específico que contempla informações como a sua identificação, o pesquisador coordenador, outros pesquisadores envolvidos, datas, além das produções e dos dados de pesquisa pertencentes a tal projeto.

Figura 1 - Tela de Projeto

Arquitetura da informação e experiência do usuário em ambientes informacionais digitais

Visualizar estatísticas Email de alerta RSS Feed

Informations

Primary Data

Project title Arquitetura da informação e experiência do usuário em ambientes informacionais digitais

Project Coordinator Vidotti, Silvana

Status Em andamento **Start date** 02-07-2016

Co-Investigator(s) Roa-Martinez, Sandra Milena
Ferreira, Ana Maria
Coneglian, Caio Saraiva
Rodas, Cecilio Merlotti

Description

Abstract
As ações humanas baseadas no exercício da técnica e direcionadas à estruturação, representação e organização de conteúdos informacionais estão presentes em todos os momentos da história das sociedades...

Keyword(s) Arquitetura da informação
Experiência do usuário
Ambientes informacionais digitais

Publications (Todos) Mostrar/Ocultar Filtros

Resultados 1-4 of 4 (Tempo de busca: 0.002 segundos)

Data do envio	Título	Autor(s)
1 Out-2016	ARQUITETURA DA INFORMAÇÃO E EYE TRACKING: O QUE O OLHAR E OS DADOS REVELAM	Vidotti, Silvana, Coneglian, Caio Saraiva, Roa-Martinez, Sandra Milena, Ferreira, Ana Maria, Rodas, Cecilio Merlotti, Santarem Segundo, José Eduardo

Fonte: Elaborado pelos autores, a partir da consulta em:
<<http://gpnti.marilia.unesp.br:8080/jspui/cris/project/pj00017>>. Acesso em: 07 ago. 2017.

Diante dos aspectos apontados nos Quadros 1 e 2 e na Figura 1, o projeto piloto do repositório GPNTI-CRIS foi configurado e implantado. Algumas das características já foram implementadas tais como a plataforma, o padrão de metadados, o perfil de aplicação, as comunidades, as coleções e os tipos de recursos com os seus respectivos formatos, no entanto outras características serão inseridas de acordo com o aumento nos tipos de recursos e a realização de manutenções no repositório, como o plano de gestão de dados.

Após a customização do sistema realizou-se a inserção de alguns dados e registros de projetos de pesquisa desenvolvidos no âmbito no GPNTI, que será apresentado na seção seguinte.

7.1 APLICAÇÃO DO REPOSITÓRIO

Para validar o uso do projeto piloto do repositório GPNTI-CRIS, demonstra-se a seguir uma aplicação de uso do repositório. Para o projeto piloto foram considerados projetos de pesquisa, livros, teses, artigos científicos, dados de pesquisa (vídeos, planilhas, imagens, etc.) e trabalhos completos de eventos.

A figura 2 apresenta como os dados da pesquisa são visualizados no Repositório GPNTI-CRIS, sendo possível identificar os metadados e os dados em si. Neste caso, os arquivos apresentados na figura 2 são: os procedimentos para coleta dos dados (formato .txt), a tabela com os dados (formato .txt) e o vídeo de gravação (formato .rar).

Figura 2 - Tela de visualização dos dados da pesquisa.

Título:	Experimento com Eye Tracking com usuários idosos visitando websites de agências de viagem		
Autor(es):	Rodas, Cecilio Merlotti		
Palavras-chave:	Eye Tracking Terceira idade Usabilidade Arquitetura da informação		
Data do documento:	Jan-2015		
Editora:	Grupo de Pesquisa Novas Tecnologias em Informação		
Projeto:	Arquitetura da informação e experiência do usuário em ambientes informacionais digitais		
Resumo:	Vídeo e as estatísticas obtidas por meio do software Tobii Studio.		
Descrição:	Vídeo e as estatísticas obtidas por meio do software Tobii Studio.		
URI:	http://hdl.handle.net/123456789/44		
Aparece nas Coleções:	Usabilidade		
Arquivos neste item:			
Arquivo	Descrição	Tamanho	Formato
Procedimentos Coleta.txt		258 B	Text
table.txt		79,45 kB	Text
TesteAgencia.rar		43,01 MB	Unknown

Fonte: Elaborado pelos autores, a partir da consulta em:

<<http://gpnti.marilia.unesp.br:8080/jspui/handle/123456789/44>>. Acesso em: 07 ago. 2017.

Os metadados apresentados na figura 2 são uma parte do registro completo do recurso, havendo a possibilidade do usuário acessar o registro completo. Além disso, destacamos que eventuais mudanças da política do repositório podem alterar os metadados apresentados inicialmente.

Outro ponto de destaque é a possibilidade de descrever a relação entre as produções intelectuais e os dados utilizados. Neste caso, no momento de cadastrar o registro é possível utilizar o metadado *relation* do Dublin Core, com o qualificador *dataset*, que especifica qual foi o conjunto de dados utilizados naquele trabalho. No âmbito do exemplo aqui apresentado, a relação entre um artigo e seus dados de pesquisa, que pode ser visualizado no recorte do registro completo do primeiro, em que está exposto esta relação, demonstrado na figura 3.

Figura 3 - Recorte do registro completo de um trabalho em evento.

dc.title	ARQUITETURA DA INFORMAÇÃO E EYE TRACKING: O QUE O OLHAR E OS DADOS REVELAM	pt_BR	
dc.type	Article	pt_BR	
dc.relation.conference	XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)	pt_BR	
dc.relation.dataset	http://gpnti.marilia.unesp.br:8084/jspui/handle/123456789/44	-	
item.fulltext	Com Texto completo	-	
Aparece nas Coleções:	Usabilidade		
Arquivos neste item:			
Arquivo	Descrição	Tamanho	Formato
4083-12201-1-PB.pdf		699,33 kB	Adobe PDF
			Ver/Aberto

Fonte: Elaborado pelos autores, a partir da consulta em:

<http://gpnti.marilia.unesp.br:8084/jspui/handle/123456789/45?mode=full> Acesso em: 07 ago. 2017.

Por meio da relação demonstrada pela figura 3, é possível visualizar que a ligação entre os recursos fica explicitada, permitindo que caso algum leitor do texto queira ter acesso aos dados primários, isso seja possível. Neste sentido, a ideia de publicação de dados de pesquisa como o enriquecimento de um artigo pode ser identificada neste exemplo, pois permite que o leitor seja capaz de acessar a outros recursos relacionados à produção intelectual. Destaca-se que os conjuntos de dados de pesquisa depositados por um pesquisador do grupo devem necessariamente estar vinculados a um determinado projeto de pesquisa, para que estes possam ser acessados posteriormente, pois a partir de um projeto, os dados de pesquisa podem relacionar trabalhos em eventos, monografias, dissertações, teses, entre outros tipos.

8 CONSIDERAÇÕES FINAIS

A publicação de dados de pesquisa está no centro das discussões da Ciência da Informação, em que estudos sobre *e-Science*, *Data Science*, *Open Data* e publicações ampliadas, buscam redimensionar a forma como os pesquisadores constroem e disponibilizam as suas pesquisas e como os leitores consomem essas informações.

Neste sentido, os grupos de pesquisa necessitam encontrar meios para disponibilizar os seus dados, de forma a seguir essa tendência global e a exigência de abertura de dados para acesso. Como consequência, os repositórios de dados de pesquisa surgiram como uma solução viável que é capaz de possibilitar, além da disponibilização dos dados, sua ligação com as produções intelectuais produzidas com tais dados.

A partir disso, o presente trabalho apresentou o processo de implantação de um projeto piloto do repositório GPNTI-CRIS, contendo as principais características que um repositório de dados implantado utilizando o *software* DSpace-CRIS pode ter, perpassando por questões próprias da Ciência da Informação, relativas à organização de comunidades e coleções, pelos padrões de metadados, perfil de aplicação, interoperabilidade, entre outros.

Assim, verificou-se que o DSpace-CRIS é uma ferramenta que reúne os critérios necessários para a criação de um repositório de dados, com características para interligar informações dos dados de pesquisa com outras informações (pesquisadores, projetos entre outros). Essa plataforma, ao se inserir em um sistema de repositório as características do CRIS, consegue ser um ambiente adequado para a construção de um repositório de dados para um grupo de pesquisa. Isto porque a utilização de sistema de informação de pesquisas correntes (CRIS) permite que o processo de pesquisa como um todo seja registrado e os dados e os recursos de pesquisa sejam integrados a todas as entidades do processo (pesquisadores, instituições e projetos de pesquisa). Assim, a descrição de cada um desses elementos do CRIS permite a obtenção de informações importantes para a tomada de decisões no processo de gestão da pesquisa.

Outro ponto abordado diz respeito ao uso do modelo CERIF, que foi utilizado num primeiro momento por ser o padrão para sistemas CRIS da Comunidade Europeia, visando a interoperabilidade.

Como trabalhos futuros, busca-se inserir o repositório GPNTI-CRIS na plataforma Re3Data (<http://www.re3data.org/>), visando divulgar este trabalho, como uma forma de

construção de repositórios de dados para grupos de pesquisa. Busca-se também aplicar os demais processos de curadoria digital que não foram aplicados até o momento. Com a inserção de novos dados, será possível gerar a rede de colaboração de cada pesquisador, proporcionando uma visão geral como por exemplo a relação de um pesquisador com outros.

Finalizando, o repositório GPNTI-CRIS permitiu a integração de projetos tanto com as produções intelectuais quanto com os dados coletados durante as pesquisas, possibilitando o re(uso) dos dados de pesquisa, a interoperabilidade entre diferentes ambientes e a disseminação dos recursos gerados pela comunidade do GPNTI.

AGRADECIMENTOS

Os autores agradecem aos Grupos de Pesquisa: Novas Tecnologias em Informação da Universidade Estadual Paulista Júlio de Mesquita Filho (Brasil), Inteligência Computacional do Departamento de Sistemas da Universidade do Cauca (Colômbia), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES e à Associação Universitária Ibero-americana de Pós-graduação - AUIP.

Research Data Repository for Research Group: a Pilot Project

ABSTRACT

Considering the importance of data in research analysis and the possibility of data reuse, it becomes relevant to use tools for management of collected research dataset from research projects and also manage, in the same platform, scientific papers, event papers, dissertations, thesis and others publications. In this context, scientific research data repositories can be used to organize, store, give access and use of results presented in different formats. This paper aims to understand and investigate the features of a research data repository align with elements of a Current Research Information System (CRIS). As a specific goal, the research analyses the software DSpace-CRIS to implement a research data repository, to warrant the availability of information and data related to research projects, researchers, institutions, publications, among others. The paper brings theory about Research Data, Metadata and metadata frameworks, Application Profile, Digital Curation, Data Management Plan and technical information about DSpace-CRIS. Thus, it was possible to implantation of a research data repository using a Current Research Information System – CRIS for the Research Group New Technologies in Information. As a result, it was noted that the features and functionalities of this tool meets the needs and characteristics of research data publications such as the enrichment of an intellectual production that allows management of data and information. In this sense, we believe that DSpace-CRIS is a propitious option to be used by research groups, which in this repository allowed the integration of projects with both the intellectual productions and the data collected during the research, allowing the (re) use of research data, interoperability among different environments and the dissemination of resources generated by the GPNTI community.

Keywords: Research data repositories. Current Research Information System – CRIS. Research Data. Application Profile.

REFERÊNCIAS

BORGMAN, C. L. Research Data: Who will share what, with whom, when, and why?. In: CHINA–NORTH AMERICA LIBRARY CONFERENCE, 5., 2010, China. **Proceedings...** . China: National Library Of China, 2010. p. 1 - 21. Disponível em: < <http://bit.ly/2vFOMSA>>. Acesso em: 04 ago. 2017.

CERIF-1.6. 2013. Disponível em: <<http://www.eurocris.org/cerif/feature-tour/cerif-16>>. Acesso em: 26 jan. 2017.

COMMITTEE FOR A STUDY ON PROMOTING ACCESS TO SCIENTIFIC AND TECHNICAL DATA FOR THE PUBLIC INTEREST et al. **A question of balance:** Private rights and the public interest in scientific and technical databases. Washington, D.C.: National Academy Press, 1999. Disponível em: <https://books.google.com.br/books?id=D2SdAgAAQBAJ>>. Acesso em: 04 ago. 2017.

COSTA, Michelli; BRAGA, Tiago. Repositórios de dados de pesquisa no mundo. **Cadernos BAD**, Portugal, n. 2, jul-dez, pp. 80-95, 2016. Disponível em: <<https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1585>>. Acesso em: 03 set. 2017.

DIGITAL CURATION CENTRE. **Checklist for a Data Management Plan.** 2014. Disponível em: <<http://www.dcc.ac.uk/resources/data-management-plans/checklist>>. Acesso em: 26 jul. 2017.

DIGITAL CURATION CENTRE. **What is digital curation?** 2017. Disponível em: < <http://www.dcc.ac.uk/digital-curation/what-digital-curation> >. Acesso em: 20 jul.2017.

EUROCRIS. **CRIS concept and CRIS benefits.** 2010. Disponível em: <<http://bit.ly/2uj45HL>>. Acesso em: 19 jul. 2017.

EUROPEAN COMMISSION. **Towards Better Access to Scientific Information:** Boosting the Benefits of Public Investments in Research. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. 2012. Disponível em: <<http://bit.ly/2vcBPFg>>. Acesso em: 04 ago. 2017.

FLORIDI, L. Semantic Conceptions of Information. **The Stanford Encyclopedia of Philosophy** (Spring 2017 Edition), Edward N. Zalta (ed.). Disponível em: <<https://plato.stanford.edu/archives/spr2017/entries/information-semantic/>>. Acesso em: 04 ago. 2017.

KINDLING, M.; SCHIRMBACHER, P. Die digitale Forschungswelt als Gegenstand der Forschung. **Information: Wissenschaft und Praxis**, v. 64, p. 127–136, 2013. Disponível em: <<http://bit.ly/2veusLV>>. Acesso em: 04 ago. 2017.

INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH. **Guidelines for Effective Data Management Plans**. 2017. Disponível em: <<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/>>. Acesso em: 26 jul. 2017.

JOINT, N. Current research information systems, open access repositories and libraries: ANTAEUS. **Library Review**, v. 57 n. 8, p. 570-575, 2008. Disponível em: <<http://www.emeraldinsight.com/doi/full/10.1108/00242530810899559>>. Acesso em: 04 ago. 2017.

NATIONAL INSTITUTES OF HEALTH. **Final NIH Statement on Sharing Research Data**. MD, USA: National Institutes of Health Bethesda, 2003. Disponível em: <<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>> . Acesso em: 04 ago. 2017.

NATIONAL SCIENCE BOARD; NATIONAL SCIENCE FOUNDATION. **Long-lived digital data collections: Enabling research and education in the 21st century**. Washington, D.C: National Science Foundation (U.S.), 2005. Disponível em: <<http://bit.ly/2vxgB6b>> . Acesso em: 04 ago. 2017.

OPEN KNOWLEDGE INTERNATIONAL. O que são Dados Abertos? Disponível em: <http://opendatahandbook.org/guide/pt_BR/what-is-open-data/>. Acesso em: 06 set. 2017.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding**. Paris: OECD Publishing, 2007. Disponível em: <<https://doi.org/10.1787/9789264034020-en-fr>>. Acesso em: 04 ago. 2017.

PAMPEL, H.; DALLMEIER-TIESSSEN, S. Open Research Data: From Vision to Practice. In: BARTLING, S.; FRIESIKE, S. (Orgs.). **Opening Science**. Springer International Publishing, 2014. p. 213–224. Disponível em: <http://link.springer.com/chapter/10.1007/978-3-319-00026-8_14>. Acesso em: 28 jun. 2017.

PAMPEL, H. et al. Making research data repositories visible: The re3data. org registry. **PloS one**, v. 8, n. 11, p. e78080, 2013. Disponível em: <<http://bit.ly/2vHTqGA>>. Acesso em: 28 jun. 2017.

QIN, J.; BALL, A.; GREENBERG, J. Functional and Architectural Requirements for Metadata: Supporting Discovery and Management of Scientific Data. In: DCMI INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 12., 2012, Kuching. The Kuching Proceedings... Kuching: DCMI, 2012. p. 62 - 71. Disponível em: <<http://bit.ly/2hCHBeV>>. Acesso em: 25 jun. 2017.

RODRIGUES, E.; SARAIVA, R. **Os Repositórios de Dados Científicos**: Estado da arte. Portugal: RCAAP, 2010. Disponível em: <<http://bit.ly/2vx3FgI>>. Acesso em: 25 maio. 2017.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade**, v. 22, n. 3, p. 179-191, 2012. Disponível em: <https://www.icict.fiocruz.br/sites/www.icict.fiocruz.br/files/Curadoria%20digital_Luis%20Fernando%20Sayao.pdf>. Acesso em: 11 ago. 2017.

SAYÃO, L. F.; SALES, L. F. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN. 2015. Disponível em: <<http://www.cnen.gov.br/component/content/article?id=160>>. Acesso em: 26 jul. 2017.

SAYÃO, L. F.; SALES, L. F. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, v. 21, n. 2, p. 90-115, 2016a. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/27939/20122>>. Acesso em: 11 ago. 2017.

SAYÃO, L. F.; SALES, L. F. Curadoria digital e dados de pesquisa. **AtoZ**: novas práticas em informação e conhecimento, v. 5, n. 2, p. 67 – 71, 2016b. Disponível em: <<http://dx.doi.org/10.5380/atoz.v5i2.49708>>. Acesso em: 04 ago. 2017.

SILVEIRA, L.; SHINTAKU, M. BOLLINI, A. **Guia de instalação DSpace-CRIS**. Brasília: IBICT, 2016. Disponível em: <<http://bit.ly/2vez0BY>>. Acesso em: 04 ago. 2017.