
Elaboração da gestão de dados em um laboratório de pesquisa do campo educacional

El proceso de gestión de datos en un laboratorio de investigación educativa

Data management procedures in a educational research laboratory

Marcia Regina SILVA, Carla BARRETO DÓRIA PINTO, Ieda PELOGIA MARTINS DAMIAN, Ildeberto Aparecido RODELLO, Ana Clara RAMOS SIMÕES, Milena VILLELA MACHADO DOS REIS

Universidade de São Paulo, Av Bandeirantes, 3900 Ribeirão Preto S.P. CEP 14040-901, marciaregina@usp.br, carlabarreto@usp.br, iedapm@usp.br, rodello@usp.br, anaramosimoses@gmail.com, vilelamilena@gmail.com

Resumen

La explosión de información y el movimiento de acceso abierto han planteado desafíos importantes en la gestión de datos para laboratorios y grupos de investigación. El Laboratorio de Estudos e Pesquisas em Economia Social (LEPES) de la Universidad de São Paulo se enfoca en desarrollar investigaciones y evaluaciones de impacto de las políticas sociales en Brasil, generando una cantidad sustancial de datos de investigación. Sin embargo, el almacenamiento de estos datos en un servidor con acceso restringido ha provocado varios problemas, incluidas dificultades para acceder a los datos, falta de estandarización, problemas de seguridad y problemas de gestión de versiones. Para abordar estos desafíos, este trabajo tiene como objetivo proponer una solución integral de gestión de datos para LEPES. La propuesta abarca múltiples acciones, como educación, experiencia, curación, indexación y privacidad de datos, todo integrado en un repositorio abierto para datos de investigación. El objetivo principal es facilitar el almacenamiento de datos y permitir el acceso público, promoviendo el intercambio de datos. El enfoque adoptado en este estudio es un diseño de investigación aplicada que utiliza el estudio de caso como método. Además, el estudio se basa en la literatura bibliográfica y documental relevante. La innovación de la propuesta radica en su meticulosa alineación de los principios de gestión de datos dentro de un laboratorio que tiene un impacto directo en las políticas públicas educativas brasileñas. En cuanto a los resultados, el grupo de gestión describe los caminos que siguieron para garantizar el intercambio seguro de datos de investigación, al tiempo que se adhirió al marco legal existente. Además, la solución propuesta no solo beneficia a LEPES, sino que también ofrece información valiosa para otros grupos que enfrentan desafíos similares de gestión de datos. En general, esta investigación se esfuerza por crear un marco de gestión de datos sólido y práctico que facilite el intercambio seguro de datos, fomente la colaboración y contribuya significativamente al avance de la investigación sobre políticas sociales en Brasil.

Palabras clave: Gestión de datos. Datos de la investigación. Laboratorio de investigación. Economía social. Brasil.

Abstract

The explosion of information and the open access movement have posed significant challenges in data management for laboratories and research groups. The Laboratory of Studies and Research in Social Economy (Laboratório de Estudos e Pesquisas em Economia Social - LEPES) at the University of São Paulo focuses on developing research and impact assessments of social policies in Brazil, generating a substantial amount of research data. However, storing this data on a server with restricted access has led to various issues, including difficulties in accessing data, lack of standardization, security concerns, and version management problems. To address these challenges, this work aims to propose a comprehensive data management solution for LEPES. The proposal encompasses multiple actions, such as education, expertise, curation, indexing, and data privacy, all integrated into an open repository for research data. The primary objective is to facilitate data storage and enable public access, promoting data sharing. The approach taken in this study is an applied research design that utilizes the case study as a method. Additionally, the study draws support from relevant bibliographical and documentary literature. The proposal's innovation lies in its meticulous alignment of data management principles within a laboratory that has a direct impact on Brazilian educational public policies. As for the results, the management group outlines the pathways they followed to ensure secure sharing of research data, while adhering to the existing legal framework. Furthermore, the proposed solution not only benefits LEPES but also offers valuable insights for other groups facing similar data management challenges. Overall, this research endeavours to create a robust and practical data management framework that facilitates safe data sharing, fosters collaboration, and contributes significantly to the advancement of social policy research in Brazil.

Keywords: Data management. Research data. Research lab. Social economy. Brazil.

1. Introdução

A Gestão de Dados (GD) se tornou um desafio para pesquisadores, laboratórios e grupos de pesquisas, pois requer engajamento e mudança de atitudes em relação as estratégias de administração, organização, armazenamento e proteção. Tais mudanças são decorrentes do Movimento da Ciência Aberta (*Open Science*) de tornar a pesquisa e os dados científicos acessíveis e da entrada em vigor no Brasil da Lei Geral de Proteção de Dados Pessoais (LGPD) - Lei nº 13.709/2018, que trouxe preocupações em relação às particularidades da metodologia científica para a condução da pesquisa.

Em relação à esta gestão, as agências de fomento têm seguido direcionamentos de escopo mundial para armazenamento e compartilhamento de dados de pesquisa, orientando pesquisadores e instituições quanto a necessidade de elaboração de projetos de GD de suas pesquisas.

Porém, ainda existem muitas razões para que a opacidade se instale nesse segmento da ciência e crie obstáculos para que os dados sejam compartilhados e reusados, como a falta de infraestruturas tecnológicas e gerenciais e de políticas institucionais que assegurem estabilidade, persistência e interoperabilidade dos dados; o controle de qualidade e a padronização, uma vez que a natureza heterogênea e fragmentada dessas coleções exige estratégias diversificadas para sua gestão; políticas voltadas para publicação de dados, ausência de esquemas de reconhecimento da autoria e políticas de recompensa pela organização e pela disseminação dos dados e falta de interesse dos pesquisadores em divulgar dados além dos limites profissionais mais próximos, dados sobre hipóteses não confirmadas, resultados negativos e dados considerados auxiliares de estudos publicados.

Em relação ao segmento da ciência no Brasil, se apresenta, como objeto de pesquisa deste estudo, o Laboratório de Estudos e Pesquisas em Economia Social (LEPES), fundado em 2011, na Universidade de São Paulo, que desenvolve pesquisas e realiza avaliações de impacto de políticas sociais em todo o território nacional. Desde então, o Laboratório tem produzido uma quantidade expressiva de dados de pesquisas que está armazenada em um servidor de acesso restrito. Esse cenário apresenta problemas como: dificuldades de acesso, diferenças de padronização e estruturação, falta de segurança, dificuldades de localização, conflitos de gerenciamento de versões e duplicidade de dados, dentre outros. O expressivo crescimento do LEPES levou ao surgimento de várias frentes de trabalho simultâneas,

o que, do ponto de vista dos dados de pesquisas, potencializou os problemas mencionados anteriormente.

O LEPES se destaca pela excelência e amplitude de projetos desenvolvidos, impactando em diversos seguimentos sociais, como, por exemplo, a agenda “Primeira Infância”, cujo objetivo é traduzir os conceitos de qualidade presentes na Base Nacional Comum Curricular Brasileira “em instrumentos avaliativos para fazer diagnósticos de redes educacionais, bem como a desenhar programas integrados de avaliação, gestão e formação continuada (LEPES, 2023) e a agenda “Violência e Criminalidade”, que

[...] se dedica a descrever como o crime se organiza, enfatizando a participação do jovem; a estudar, através de estudos longitudinais, quais são os fatores de risco para o desvio de trajetórias dos jovens em direção a contextos mais violentos; e a estudar políticas preventivas ou que não envolvam encarceramento (tais como as de justiça restaurativa) para casos leves de transgressão e criminalidade.

A decisão de elaboração de um projeto de GD partiu dos gestores do Laboratório e da equipe de dados, estabelecida em 2020. A equipe é formada por pesquisadores, estagiários e docentes das áreas de Gestão e da Ciência da Informação que tiveram como objetivo elaborar diretrizes para o atendimento da LGPD e estabelecimento de um Plano de GD. Se faz oportuno ressaltar a importância desta decisão: a falta de uma adequada gestão de dados estava impossibilitando o crescimento potencial do LEPES, ou seja, sem a devida organização e adequação dos dados de pesquisa, o LEPES não teria condições de assumir novas frentes de trabalho e de condução de pesquisas.

Diante de tal cenário, esse artigo tem como objetivo relatar as principais ações realizadas para a implantação de um projeto de GD de pesquisa no LEPES. A maioria das pesquisas que envolve a gestão estratégica está direcionada ao contexto organizacional, de modo que, existe uma lacuna de desenvolvimento de pesquisas de GD voltadas a pesquisa. Quanto aos seus objetivos, esta pesquisa configura-se como exploratória e descritiva, tendo como estratégia o estudo de caso.

Além desta introdução, que buscou contextualizar o tema e apresentar a problemática e os objetivos, apresenta-se no referencial teórico conceitos e definições relacionadas aos dados de pesquisa, em seguida descreve-se o protocolo para a implantação da gestão de dados de pesquisa no LEPES e, finalmente, as considerações finais.

2. Dados de pesquisa: contexto

Os dados de pesquisa podem ser conceituados como “Informação registrada necessária para apoiar ou validar as observações, descobertas ou resultados de um projeto de pesquisa” (Rice e Southall, 2016, p. 20, tradução nossa) e podem contribuir para fomentar investigações científicas e para o ensino da ciência. Para Semeler e Pinto (2019, p. 127), tais dados podem ser vistos como “um produto para pesquisa, bem como um insumo para a pesquisa. Esse tipo de dados assume papel relevante no que tange aos novos processos de comunicação científica”.

A GD, por sua vez, compreende, conforme, Plotkin (2020), as atividades envolvidas com o processo decisório e com a autoridade com que tudo que se relaciona aos dados. Ela envolve procedimentos estratégicos que englobam tanto a definição de políticas de governança como o estabelecimento de rotinas relacionadas ao uso de ferramentas e tecnologias, como: seleção e implementação de banco de dados, integração de sistemas, limpeza e normalização de dados, backups, adoção de medidas de segurança que garantam a privacidade e integridade dos dados. Hoje, existe uma diversidade de repositórios de dados de pesquisa que contribuí para o armazenamento e acesso de dados que subsidiaram muitas pesquisas.

Conforme Sales e Sayão (2023, p. 2-3), a gestão de dados

[...] se concretiza aos olhos do pesquisador na forma de um amplo espectro de serviços e ferramentas que apoiam todo o ciclo de vida dos dados, no âmbito de um projeto de pesquisa, cujos benefícios são diretamente perceptíveis por ele, como por exemplo: maior visibilidade para a pesquisa; mais citações e prestígio; reconhecimento da autoria dos dados; maior nível de colaboração em escala global; organização dos dados para o próprio uso do pesquisador e de seus colegas próximos; reconhecimento em termos de promoção e financiamento; proteção lógica e física dos dados; e preservação de longo prazo.

Os dados de pesquisa figuram-se como elementos de destaque no cenário científico, uma vez que carregam significados importantes para o desenvolvimento de novas pesquisas, além de dar credibilidade aos trabalhos publicados. Os pressupostos da Ciência Aberta ressignificaram a Ciência, contribuindo para o desenvolvimento científico com o incentivo para o compartilhamento de materiais de pesquisa, códigos, instrumentos, metodologias etc. A ideia de reuso passou a ser o ideal da ciência, como forma de aproveitamento de recursos e tempo dispensados por pesquisadores. A possibilidade de acesso a dados

de pesquisa vem tornando o processo de pesquisa mais transparente.

Além disso, os pesquisadores enfrentam dificuldades no âmbito de laboratórios e grupos de pesquisas quanto a manipulação e recuperação dos dados. Por isso, há perspectivas de que a GD seja uma prática estabelecida na ciência. As tecnologias e ferramentas vigentes contribuem para que a captura, curadoria e análise de dados de pesquisa sejam um processo desenvolvido para além das grandes instituições científicas. Isso contribui para que grupos de pesquisa possam elaborar planos de GD que abarquem desde a coleta até a disponibilização dos dados em repositórios de dados de acesso público.

Há uma parcela considerável do trabalho científico que não está visível nem para a sociedade, nem para os pares no contexto da comunidade científica. A descontinuidade provocada pela invisibilidade de um segmento da pesquisa científica fragiliza a ciência, como empreendimento social e humanístico, e tem desdobramentos como: deixar registros incompletos e tendenciosos dos processos de geração de conhecimento, criando uma lacuna nas memórias acadêmicas das instituições; duplicar esforços que alonguem desnecessariamente o ciclo de comunicação científica e, sobretudo, impossibilitar que se efetivem os princípios básicos da reprodutibilidade dos experimentos científicos e da autocorreção da ciência (Ferguson et al., 2014; The Royal Society, 2012).

Conforme ressaltam Ferguson et al. (2014, p. 1443),

Estudos recentes indicam que mais de 50% das descobertas científicas não aparecem na literatura publicada, ao invés disso, residem nas gavetas e nos computadores pessoais dos pesquisadores.

Esse fenômeno é conhecido como “cauda longa da ciência”, em que um grande número de pequenas equipes de pesquisadores e laboratórios independentes gera uma ampla variedade de coleções de dados. Apesar do dimensionamento individual, essas pequenas coleções de dados estão sendo reconhecidas como ativos informacionais de alto valor, que, coletivamente, têm o potencial de ser mais relevantes do que a soma de suas partes (Wyborn e Lehnert, 2016). Os dados da cauda longa representam a maior parcela de dados produzida pela ciência e constituem um território de constante criatividade e inovação que precisam ser revelados, integrados e compartilhados.

Os dados dão garantias e evidências da veracidade dos resultados das pesquisas e dos artigos científicos publicados e servem de alicerce para

o progresso da ciência (Molloy, 2011; Borgman, 2015).

Os dados de pesquisas podem ser descritos como “registros factuais usados como fonte primária para a pesquisa científica e que são comumente aceitos pelos pesquisadores como necessários para validar os resultados do trabalho científico” (OCDE, 2007, p.13). São muitos os tipos de dados de pesquisa, como números, imagens, textos, vídeos, áudio, software, algoritmos, equações, animações, modelos, simulações. A abrangência dos dados de pesquisa deve ser refletida no armazenamento e curadoria das coleções desses dados (Sayão e Sales, 2014).

Portanto, o plano de GD envolve conhecer a natureza e processos de criação de dados de pesquisas para pensar em estratégias de gestão de armazenamento e curadoria que garantam que o sistema seja robusto e tenha credibilidade.

Porém, se observa uma diferença acentuada nos processos de GD. Enquanto um grande cuidado é devotado à formação de coleções, à preservação e ao reuso de dados provenientes dos grandes projetos, comparativamente, pouca atenção é dedicada aos dados que são gerados pela maioria dos cientistas que desenvolvem projetos em menor escala. Novas estruturas sociais e desenvolvimentos técnicos podem aumentar significativamente a disponibilidade e o valor dos dados dos “pesquisadores individuais” e de seus projetos de pesquisa. O desafio que se coloca para a política científica e para as instituições de pesquisa é desenvolver infraestruturas e práticas

como repositórios digitais disciplinares, que tornem esses dados úteis para a sociedade (Heidorn, 2008).

3. Protocolo para a implantação da gestão de dados de pesquisa no LEPES

O LEPES, buscando alinhar-se às novas diretrizes da *Open Science* e assumindo o compromisso de compartilhamento dos dados que gera, buscou, por meio de sua equipe de dados, criar uma agenda que envolveu a elaboração de protocolo para a organização das pastas de um sistema de armazenamento em nuvem, de protocolos para renomeação de arquivos, para renomeação de variáveis dos *datasets*, de documentação do *script* de automatização da renomeação, de ajustes de dados e um guia de boas práticas. Todos esses produtos fazem parte do Plano de GD elaborado para o LEPES.

Durante a trajetória do LEPES, a constante elaboração de projetos gera um grande volume de dados brutos, sendo todos relevantes para as pesquisas. Tendo em vista que a falta de padronização e organização pode levar a situações propensas a perda de desses dados, seja por falha no equipamento ou por falta de comunicação, foi elaborado, com base em estudos de caso do Banco Mundial e das diretrizes internacionais da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), um protocolo para orientar todas as ações nesse âmbito, procurando melhor organizar e evitar a perda de dados, além de compartilhar e otimizar o trabalho realizado.

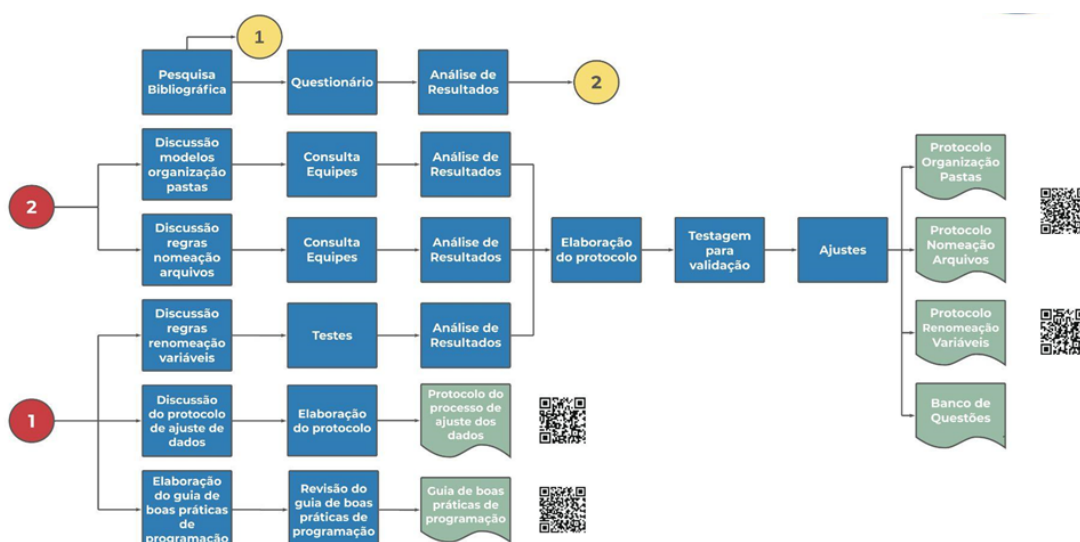


Figura 1. Processo de criação de protocolos de GD do LEPES

Descreve-se, então, os passos e orientações para a elaboração dos modelos (protocolos) e

ferramenta que compõem o projeto para a elaboração do Plano de GD do LEPES:

1. pesquisa bibliográfica em bases de dados nacionais e internacionais;
2. estudo de caso – identificação dos problemas relacionados a falta de GD no LEPES;
3. elaboração e aplicação de questionário para identificação do perfil dos usuários do LEPES;
4. análise e discussão dos resultados;
5. elaboração dos protocolos e testagem;
6. apresentação dos protocolos aos integrantes do LEPES; e
7. implementação dos modelos no LEPES

A Figura 1 (na página anterior) representa o caminho percorrido pelo projeto para a elaboração do Plano GD. A implementação dos protocolos aconteceu de forma gradativa, de acordo as seguintes etapas:

1. criação e organização das pastas no serviço de armazenamento em nuvem pela equipe de dados;
2. transferência dos arquivos da pasta anterior para a recém-criada;
3. execução do *script* para renomeação dos arquivos antigos de acordo com o padrão de nomenclatura estabelecido;
4. organização manual dos arquivos nas pastas de destino e;
5. monitoramento pós-implantação.

- 1 Não utilizar acentuação. Ç(cedilha), símbolos a ou °, caracteres especiais ou de pontuação (^~*!#%\$&)
- 2 Não utilizar espaços em branco ou – (hifen). Os campos devem ser separados por underscores "_" sem espaço de digitação
- 3 Utilizar apenas letras maiúsculas
- 4 Evitar preposições, conjunções e artigos (o, a, e, de, do, da...)
- 5 Os nomes dos arquivos devem facilitar a sua busca, devendo ser orientado pelo seu conteúdo como relatório de..., apresentação X, proposta Y
- 6 A data da última edição do arquivo será utilizada como versionamento. O padrão adotado é o ISO 8601-2:2019, no formato ano-mês-dia (AAAAMMDD)
- 7 Versionamento pode ser inserido antes ou depois do nome do arquivo

Quadro 1. Regras para a nomeação de arquivos do LEPES

A nomeação de arquivos é uma etapa importante do projeto de GD. A adoção de um modelo único de nomenclatura facilita a recuperação de dados interequipes e dos dados em geral, além de evitar o desaparecimento de arquivos e solucionar problemas relacionados ao versionamento e duplicidade. Em suma, minimiza as dificuldades encontradas pelos usuários nos processos de busca pelos arquivos. Foram elaboradas regras gerais para a nomeação de arquivos (Quadro 1).

A Figura 2 representa a proposta de estruturação e organização das pastas no drive virtual compartilhado. Para isso, foram adotadas as seguintes diretrizes: 1) Limite de 10 (dez) níveis para criação de subpastas; e 2) Mesma convenção de nomeação de arquivos, não sendo obrigatório o versionamento.

A proposta de uma nova estrutura de organização impacta diretamente no modo de trabalho no LEPES. Com isso em mente, o modelo de atendimento às equipes para realizar essa transição foi planejado para ser realizado por etapas.

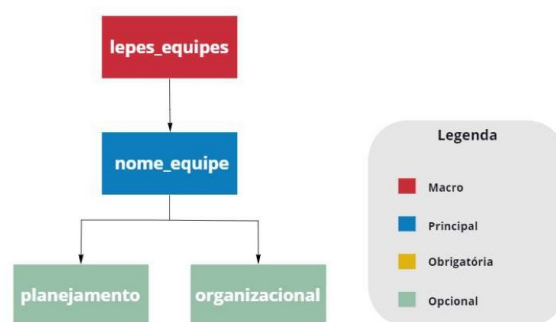


Figura 2. Estrutura de organização das pastas - Pasta Equipes

- 1 Escolha da data para a transferência dos arquivos do projeto para o Drive de destino
- 2 Líder da equipe seleciona a data desejada e insere o caminho dos arquivos que serão transferidos
- 3 Líder da equipe sinaliza se nas pastas em questão existem dados sensíveis. Em caso afirmativo, os arquivos serão transferidos para uma pasta de acesso restrito
- 4 Concluída a transferência, no mesmo dia é realizada a nomeação dos arquivos através do aplicativo Praticidados
- 5 Concluída a segunda etapa, é realizado um novo agendamento para a transferência manual dos arquivos para a estrutura de pastas

Quadro 2. Orientações de preenchimento da Planilha

O ponto de partida para a divulgação e capacitação das equipes do LEPES foi a realização de um seminário para todos os membros. Nesta ocasião, foi disponibilizado um documento com orientações de preenchimento para os líderes das equipes, contendo as principais orientações, conforme Quadro 2.

Uma ferramenta computacional, denominada PRATICIDADOS, foi desenvolvida pela equipe de GD para a executar a atualização do inventário com os caminhos novos dos arquivos, facilitando os processos de GD do LEPES. Quando terminada a renomeação, uma planilha dentro da pasta informada é automaticamente criada, contendo as seguintes informações: diretório antigo, nome do arquivo antigo, diretório atual e nome do arquivo atual (Figura 3).

Renomeação de pastas do Google Drive e criação de inventário

ATENÇÃO Essa parte não deve ser executada online

Instruções de uso:

1) Primeiramente é necessário que o usuário tenha o Google Drive instalado no seu computador e logado na sua conta, caso você não tenha, é possível fazer o download através do seguinte link: <https://www.google.com/intl/pt-BR/drive/download/>.

2) Para realizar a renomeação é necessário informar o nome da pasta e o ID da mesma. Para obter o ID, basta entrar na pasta desejada e copiar a cadeia de caracteres que aparece no final da URL. Exemplo: <https://drive.google.com/drive/folders/GALR0BUK9Y0U0K9PVA>.

3) Após informar o nome e o ID correto da pasta, basta clicar em "Renomear" e o processo de renomeação começará.

4) Quando a renomeação terminar, será criada uma planilha dentro da pasta informada com as seguintes informações: diretório antigo, nome do arquivo antigo e nome do arquivo atual.

Figura 3. Aplicativo PRATICIDADOS

4. Conclusões

Além de seu rápido e considerável crescimento, os dados de pesquisa tiveram sua importância reconhecida em função dos benefícios que proporcionam a todos os envolvidos com o ambiente de investigação científica. Porém, para que estes benefícios possam ser usufruídos, esses dados precisam ser armazenados de maneira segura e estruturada para que sejam recuperados de modo fácil e eficaz, o que se faz por meio da GD.

Esta situação se torna ainda mais relevante no contexto dos laboratórios de pesquisa, como é o caso do Laboratório desse estudo. Diante de tal realidade, uma equipe foi constituída para desenvolver a GD de pesquisa do LEPES, aqui apresentada.

Se acredita que este trabalho poderá trazer grandes benefícios ao LEPES, como os relacionados à segurança, recuperação e reuso de dados de pesquisa, o que, dentre outras coisas, leva à

redução de tempo, de investimentos e ao aumento de produtividade e de trabalhos de excelência desenvolvidos. Para além das fronteiras do LEPES, se espera que esta pesquisa contribua para mais casos de sucesso de desenvolvimento e implantação da GD.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro para a realização dessa pesquisa.

Referências

- Borgman, Christine L. (2010). Research data: who will share what, with whom, when, and why? // *RatSWD Working Paper*. 61. <https://ssrn.com/abstract=1714427>.
- Ferguson, Adam; Nielson, Jessica; Cragin, Melissa; Bandrowski, Anita; Martone, Maryann (2014). Big data from small data: data-sharing in the 'long-tail' of neuroscience. // *Nature Neuroscience*. 17:11, 1442-1447. <https://doi.org/10.1038/nn.3838>.
- Heidorn, P. Bryan (2008). Shedding light on the dark data in the long tail of science. // *Library trends*. 57:2, 280-299. <https://doi.org/10.1353/lib.0.0036>.
- LEPES (2023). Laboratório de Estudos e Pesquisas em Economia Social (LEPES). <http://lepes.fearp.usp.br/> (2023-11-02)
- Molloy, Jenniffer C. (2011). The Open Knowledge Foundation: open data means better science. // *PLoS Biology*. 21:12, 1-4.
- OECD (2007). OECD Principles and Guidelines for Access to Research Data from Public Funding.
- Plotkin, David (2020). Data stewardship: an actionable guide to effective data management and data governance. London: Academic Press.
- Rice, Robin; Southall, John (2016). The data librarian's handbook. London: Facet Publishing.
- Sales, Luana Farias; Sayão, Luís Fernando (2023). Proposta de modelo de serviço de gestão de dados de pesquisa. // *AtoZ: novas práticas em informação e conhecimento*. 11, 1-13.
- Sayão, Luís Fernando; Sales, Luana Farias (2014). Dados abertos de pesquisa: ampliando os conceitos de acesso livre. // *RECIIS: Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*. 8:2, 76-92.
- Semeler, Alexandre Ribas; Pinto, Adilson Luiz (2019). Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. // *Ciência da Informação*. 48:1. <https://doi.org/10.18225/ci.inf.v48i1.4461>.
- The Royal Society (2012). Science as an open enterprise. London: The Royal Society Science Policy Centre.
- Wyborn, Lesley; Lehnert, Kerstin (2016). Exploiting the long tail of scientific data: making small data big. // *Eresearch Australasia Conference*, Melbourne, Australia.

Enviado: 2021-05-02. Segunda versão: 2021-08-13.
Aceptado: 2021-11-22.