

Number 50 (june 2023)

Concepção de uma Ferramenta Brasileira de Elaboração de Planos de Gestão de Dados de Pesquisa: desafios rumo ao modelo de planos acionáveis por máquina – MaDMP

 [Versió catalana](#)

 [PDF](#)

 [Metadata](#)

Laura Vilela Rodrigues Rezende

Maestra de la Universidade Federal de Goiás (UFG) y investigadora del Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

lauravil.rr@gmail.com

ORCID: <https://orcid.org/0000-0002-8891-3263>

EXIT:

<https://www.directorioexit.info/ficha6756>

Elizabete Cristina de Souza de Aguiar Monteiro

Bibliotecária de la Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/Marília) y investigadora del Instituto



Laura Vilela Rodrigues Rezende



Elizabete Cristina de Souza De Aguiar Monteiro

Brasileiro de Informação em Ciência e
Tecnologia (IBICT)

ecsamonteiro@gmail.com

ORCID: <https://orcid.org/0000-0002-3797-8139>

EXIT:

<https://www.directorioexit.info/ficha6757>

Ketlen Stueber ⓘ

Universidade Federal do Rio Grande do Sul (UFRGS) y pesquisadora del Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

ketistueber@hotmail.com

ORCID: <https://orcid.org/0000-0002-2171-0365>

EXIT:

<https://www.directorioexit.info/ficha6754>

Fabiano Couto Corrêa da Silva ⓘ

Maestro de la Universidade Federal do Rio Grande do Sul (UFRGS) y pesquisador del Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

fabianocc@gmail.com

ORCID: <https://orcid.org/0000-0001-5014-8853>

EXIT:

<https://www.directorioexit.info/ficha3640>



Ketlen Stueber



Fabiano Couto Corrêa da Silva



Alexandre De Oliveira

José Carlos Abbud Grácio

Maestro de la Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/Marília) y pesquisador del Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

jose.gracio@unesp@unesp.br

ORCID: <https://orcid.org/0000-0001-7620-1309>

EXIT:

<https://www.directorioexit.info/ficha6613>

Alexandre Faria de Oliveira

Pesquisador de la Universidade de Brasília (UnB)
Gestor de proyectos y pesquisador del Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

alexandreoliveira@ibict.br

ORCID: <https://orcid.org/0000-0003-0470-4972>

EXIT:

<https://www.directorioexit.info/ficha6755>

DOI: <https://doi.org/10.1344/BiD2023.50.07>

Recommended citation

Rezende, Laura Vilela Rodrigues; Monteiro, Elizabete Cristina de Souza de Aguiar; Stueber, Ketlen; Silva, Fabiano Couto Corrêa da; Grácio, Alexandre Faria de Oliveira; Olivera, Alexandre Faria de (2023). "Concepció d'una eina brasilera per a l'elaboració de plans de gestió de dades de recerca : reptes per al model de plans automàtics (maDMP)". *BiD: textos universitaris de biblioteconomia i documentació*, núm. 50 (june). <<https://bid.ub.edu/en/50/vilela.htm>>. DOI: 10.1344/BiD2023.50.07

Resumo

Objectivos: Este trabalho apresenta um estudo de modelo conceitual de uma ferramenta de elaboração e gestão de Planos de Gestão de Dados acionável por máquina (maDMP – Machine Actionable Data Management Plan) para o cenário brasileiro conduzido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Como objetivos, buscou-se analisar as ferramentas existentes e decidir pela viabilidade de se desenvolver uma solução desde o início ou adaptar e remodelar uma já existente além de desenhar o modelo conceitual considerando os agentes envolvidos no ecossistema de pesquisa brasileiro.

Metodologia: Trata-se de um relato de experiência de um estudo exploratório sobre o desenvolvimento de um modelo conceitual de uma ferramenta para elaboração de Plano de Gestão de Dados no cenário brasileiro. Foi utilizado o método Design Science Research, que estabelece um processo sistemático que permite ao investigador pesquisar, descrever ou explicar um fenômeno e, também, projetar ou prescrever soluções para um problema específico (Dresch *et al.*, 2014).

Resultados: Apresenta-se o detalhamento do estudo comparativo das ferramentas de elaboração de PGD existentes além do desenho do modelo conceitual da solução brasileira. Entende-se que o cenário ideal para este caso é a customização da ferramenta DMPTool já existente, otimizando recursos e tempo de desenvolvimento. Trata-se de uma ferramenta robusta, com comunidade de desenvolvimento ativa e com atualizações frequentes, fazendo dela uma solução para elaboração de PGDs acionáveis por máquinas. Destaca-se o relevante papel deste estudo no entendimento das conexões e trocas de informações necessárias para o ecossistema de Ciência brasileiro, tendo a ferramenta PGD do IBICT papel centralizador e agregador destas informações.

Resum

Objectius: aquest treball presenta un estudi del model conceptual d'una eina per elaborar i gestionar plans per a la gestió de dades automàtics (maDMP, Machine Actionable Data Management Plan) per a l'escenari brasiler dirigit per l'Institut Brasileiro de Informação em Ciência e Tecnologia (IBICT). Com a objectius, s'analitzen les eines existents per decidir sobre la viabilitat de desenvolupar una solució des del principi o bé adaptar-ne i remodelar-ne una d'existent, a més de dissenyar el model conceptual considerant els agents implicats en l'ecosistema de recerca brasiler.

Metodologia: es tracta d'un informe de l'experiència amb un estudi exploratori per desenvolupar el model conceptual d'una eina per elaborar un pla de gestió de dades en l'escenari brasiler. S'ha utilitzat el mètode Design Science Research, que estableix un procés sistemàtic que permet a l'investigador recercar, descriure o explicar un fenomen i també dissenyar o prescriure solucions per a un problema concret (Dresch *et al.*, 2014).

Resultats: es presenten els detalls de l'estudi comparatiu de les eines existents per a l'elaboració de plans de gestió de dades (PGD), a més del disseny del model conceptual de la solució brasilera. S'interpreta que l'escenari ideal per a aquest cas consisteix a personalitzar l'eina DMPTool que ja existia, per tal d'optimitzar els recursos i el temps de desenvolupament. És una eina robusta, amb una comunitat de desenvolupament activa i actualitzacions freqüents, la qual cosa la converteix en una bona solució per crear PGD automàtics. Cal destacar el paper important d'aquest estudi en la comprensió de les connexions i l'intercanvi d'informació necessaris per a l'ecosistema de la ciència brasilera, en el qual l'eina PGD de l'IBICT té un paper centralitzador i aglutinador d'aquesta informació.

Resumen

Objetivos: este trabajo presenta un estudio del modelo conceptual de una herramienta para elaborar y gestionar planes para la gestión de datos automáticos (maDMP, Machine Actionable Data Management Plan) para el escenario brasileño dirigido por el Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Como objetivos, se analizan las herramientas existentes para decidir sobre la viabilidad de desarrollar una solución desde el principio o bien adaptar y remodelar una existente, además de diseñar el modelo conceptual considerando a los agentes implicados en el ecosistema de investigación brasileño.

Metodología: se trata de un informe de la experiencia con un estudio exploratorio para desarrollar el modelo conceptual de una herramienta para elaborar un plan de gestión de datos en el escenario brasileño. Se ha utilizado el método Design Science Research,

que establece un proceso sistemático que permite al investigador buscar, describir o explicar un fenómeno y también diseñar o prescribir soluciones para un problema concreto (Dresch *et al.*, 2014).

Resultados: se presentan los detalles del estudio comparativo de herramientas existentes para la elaboración de planes de gestión de datos (PGD), además del diseño del modelo conceptual de la solución brasileña. Se interpreta que el escenario ideal para este caso consiste en personalizar la herramienta DMPTool que ya existía, a fin de optimizar los recursos y el tiempo de desarrollo. Es una herramienta robusta, con una comunidad de desarrollo activa y frecuentes actualizaciones, lo que la convierte en una buena solución para crear PGD automáticos. Cabe destacar el importante papel de este estudio en la comprensión de las conexiones y el intercambio de información necesarios para el ecosistema de la ciencia brasileña, en el que la herramienta PGD del IBICT tiene un papel centralizador y aglutinador de esta información.

Abstract

Aims: This article presents a study of a conceptual model for a machine-actionable Data Management Plan (maDMP - Machine Actionable Data Management Plan) for the Brazilian setting conducted by the Brazilian Institute of Information and Technology (IBICT). The objectives were to analyse the existing tools; to consider the feasibility of developing a new solution from the very beginning, or adapting and remodeling an existing one; and to design the conceptual model considering the agents involved in the Brazilian research ecosystem.

Methods: This article reports an exploratory study on the development of a conceptual model of a Data Management Plan for use in the Brazilian scenario. The Design Science Research method was used, a systematic process that allows researchers to study and describe a phenomenon and also design or prescribe solutions for a specific problem (Dresch *et al.*, 2014).

Results: A detailed comparative study of the existing development tools for DMPs is presented, in addition to a description of the design of the conceptual model of the Brazilian solution. The ideal scenario for this case is the improvement of the existing DMPTool tool, optimizing resources and development time. This robust instrument has accompanied the development of resources that will establish it a tool for creating DMPs that can be activated by machines. The study identifies the connections and exchanges of information necessary for the Brazilian Science ecosystem, in which the IBICT's DMP tool can play a centralizing and aggregating role.

Keywords: Scientific research , Document management systems, Information management, Brazil, Free access

Received: 13/10/2022. Accepted: 07/02/2023.

1 Introdução

Os tempos atuais exigem maior aproximação entre a comunidade científica e a sociedade em geral. Torna-se premissa básica a importância da Ciência, Tecnologia e Inovação (CTI) frente aos desafios apresentados pelas sociedades em todas as áreas do conhecimento. O movimento de abertura do fazer científico começa a se concretizar pelas nações e algumas orientações ao nível internacional começam a tomar lugar de destaque, como é o caso do documento elaborado pela UNESCO versando sobre recomendações para a Ciência Aberta.

A ciência aberta é definida como um construto inclusivo que combina vários movimentos e práticas que têm o objetivo de disponibilizar abertamente conhecimento científico multilíngue, torná-lo acessível e reutilizável para todos, aumentar as colaborações científicas e o compartilhamento de informações para o benefício da ciência e da sociedade, e abrir os processos de criação, avaliação e comunicação do conhecimento científico a atores da sociedade, além da comunidade científica tradicional. Abrange todas as disciplinas científicas e todos os aspectos das práticas acadêmicas, incluindo ciências básicas e aplicadas, ciências naturais, sociais e humanas, e se baseia nos seguintes pilares-chave: conhecimento científico aberto, infraestrutura científica aberta, comunicação científica, envolvimento aberto dos atores sociais e diálogo aberto com outros sistemas de conhecimento (UNESCO, 2021, p. 7).

O processo de abertura do fazer científico abrange várias vertentes, que vão desde o acesso aberto às publicações, disponibilização de dados, avaliações por pares abertas, tornar públicos os cadernos de laboratórios, recursos educacionais, códigos computacionais e também a popularização da Ciência. Observa-se no contexto brasileiro que as ações relacionadas com a Ciência Aberta prioritariamente ainda se concentram no acesso aberto à produção científica. Um fato que evidencia este cenário é que o diretório internacional DOAJ (Directory of Open Access Journals), que indexa periódicos de acesso aberto, apresenta um número expressivo de 1.632 revistas abertas originadas no Brasil (DOAJ, 2022).

Para este trabalho, optou-se por trazer um arcabouço prático e de implementação da gestão dos dados de pesquisa, priorizando a etapa inicial de planejamento. A vertente da Ciência

Aberta relacionada à gestão e abertura de dados começa a se consolidar no cenário brasileiro, mesmo que de maneira incipiente. Percebe-se que ainda não existe integração de informações provenientes dos agentes envolvidos em todo o ecossistema de Ciência, Tecnologia e Inovação (CTI). Além dos pesquisadores e instituições de pesquisa, as agências de fomento são atores fundamentais rumo à abertura dos dados científicos uma vez que podem solicitar que as pesquisas financiadas tornem seus dados gerados ou produzidos disponíveis para a comunidade científica e sociedade em geral.

Dito isto, a etapa de planejamento da pesquisa deve considerar quais ações serão desenvolvidas para que os dados produzidos ou gerados durante o processo investigativo sejam abertos de maneira adequada, além de também prever onde esses dados serão armazenados ao final da pesquisa. Um documento utilizado para descrever estas informações é o Plano de Gestão de Dados (PGD). Atualmente, existem ferramentas de criação destes planos, porém percebe-se um caráter demasiadamente estático e textual, sem informações que poderiam ser buscadas automaticamente de sistemas utilizados pelos agentes envolvidos em todo o ecossistema científico, ao nível nacional e internacional.

Outro ponto relevante é que, por ser um documento estático, em geral os PGDs são considerados como documento adicional a ser preenchido pelos pesquisadores somente quando lhes é solicitado, por parte das agências de fomento. Entende-se que, ao invés de ser considerado pelos pesquisadores como mais uma demanda burocrática, os PGDs deveriam ser parte integrante de todas as etapas da pesquisa, estabelecendo conexões entre os agentes envolvidos no processo investigativo, servindo como uma espécie de integrador dinâmico, passível de atualizações automáticas durante todo o gerenciamento de dados de pesquisa.

Existem algumas iniciativas de definição e desenvolvimento de PGD acionável por máquina, ou seja, estruturado de maneira a possibilitar que os agentes envolvidos no processo investigativo e de gestão dos dados atuem de forma integrada e colaborativa tendo várias informações sendo preenchidas de maneira automatizada. Este trabalho apresenta um estudo de definição do modelo conceitual de ferramenta de elaboração e gestão de PGDs acionável por máquina (maDMP - Machine Actionable Data Management Plan) para o cenário brasileiro conduzido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

2 Planejando a gestão de dados de pesquisa

Visando o compartilhamento das pesquisas produzidas bem como seus processos, metodologias, códigos e dados, é preciso que os pesquisadores realizem o planejamento de toda a investigação, inclusive o que será feito com seus dados e insumos gerados. Visando ilustrar de maneira mais didática e como uma tentativa de atualizar e adaptar o modelo de

comunicação científica proposto por Pinto e Costa (2018, p.155), a figura 1 a seguir apresenta uma sugestão de como poderia ser o ciclo de vida da pesquisa e opções de disseminação dos seus resultados e dados científicos gerados. Na fase inicial, tem-se a etapa do planejamento da pesquisa, quando é possível que o pesquisador utilize o Plano de Gestão de Dados (PGD) no intuito de detalhar os procedimentos da pesquisa, técnicas de coleta e quais os dados serão gerados, além de fornecer informações referentes ao local de armazenamento e publicação dos resultados ao final das pesquisas. Ao término do processo investigativo, é possível compartilhar os resultados e dados gerados ou não. Exemplificando como caso negativo, tem-se as questões que envolvem propriedade intelectual e industrial, havendo a necessidade de manter sob sigilo ou embargo produtos gerados como resultados de pesquisas científicas. Partindo para a parte central e direita da figura 1, na etapa de finalização da pesquisa, ao decidir por disseminar ou compartilhar os resultados e dados gerados, o pesquisador tem como opção a comunicação formal, que envolve os periódicos científicos e todo seu processo de avaliação por pares como principal meio disponível. Para além da publicação formal e no escopo do presente estudo, um dos modos de abertura do fazer científico é o depósito dos dados de pesquisa gerados durante o processo investigativo. Os repositórios de dados científicos são considerados os melhores locais de armazenamento uma vez que são concebidos visando conferir informações descritivas detalhadas sobre os conjuntos de dados visando favorecer a localização, inteligibilidade e compartilhamento.

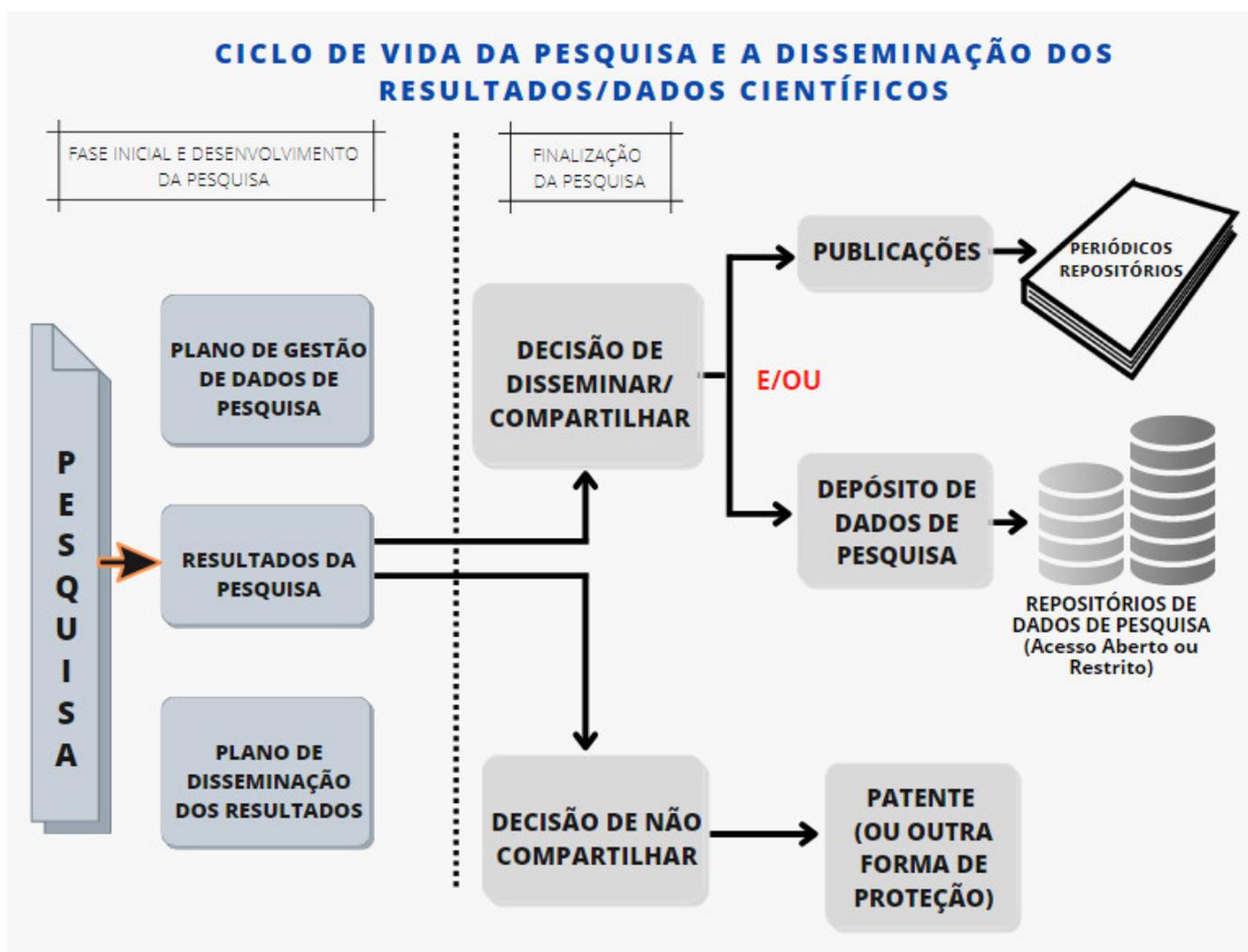


Figura 1. Ciclo de vida da pesquisa e a disseminação dos resultados/dados científicos.

Fonte: os autores, baseada em Pinto; Costa, 2018.

Desta forma, torna-se necessário planejar a gestão de dados a ser realizada durante a pesquisa visando o compartilhamento e divulgação científica adequados, tendo os planos de gestão de dados como documentação relevante para o detalhamento de todas estas etapas. Destaca-se que as opções de publicação dos resultados de pesquisa e dados devem ser entendidas como maneiras complementares de abertura do fazer científico.

2.1 Planos de gestão de dados

Um Plano de Gestão de Dados (PGD) é um documento que descreve todo o ciclo de vida dos dados de pesquisa, desde a sua coleta até a documentação completa do processo de pesquisa e registra as decisões tomadas em relação aos padrões de metadados, formatos, licenças, período de armazenamento, custos associados, dentre outros fatores que envolvem as

características e processos da pesquisa desde a sua concepção até o armazenamento, preservação e consequente disponibilização.

O PGD requer uma sequência documentada de ações destinadas a identificar, assegurar recursos, coletar, manter e utilizar os conjuntos de dados. Isto inclui a obtenção de financiamento e identificação dos recursos técnicos e de pessoal para o completo ciclo de gestão de dados. O alcance e a quantidade de detalhes em um PGD dependem do projeto e do público para o qual está sendo desenvolvido.

Os PGDs são documentos fundamentais para garantir a preservação e acessibilidade dos dados gerados em pesquisas científicas. Esses planos têm sido exigidos por agências de fomento e instituições de pesquisa para assegurar que os dados gerados sejam tratados de forma adequada, desde a concepção até a preservação.

2.2 Conferindo dinamismo aos Planos de Gestão de Dados: maDMPs

A questão cerne que se apresenta como problemática dos PGDs é que os pesquisadores entendem que se trata de mais uma demanda a ser feita por eles, muitas vezes de maneira mandatória e, da forma como ocorre hoje, estes documentos se apresentam estáticos e sem interligação com outros agentes envolvidos no fazer científico. No contexto das agências de fomento que solicitam PGDs o controle, conferência e análises desses documentos acabam sendo feitos manualmente desprovidas de inter-relação entre identificadores que poderiam interligar pesquisadores, pesquisas, financiamento, dados e publicações.

Nos últimos anos, esta problemática impulsionou estudos voltados para o aprimoramento dos PGDs na tentativa de lhes conferir dinamismo, interoperabilidade com outros sistemas e capacidade de atualização automática de informações ao longo do processo investigativo. Estes PGDs são conhecidos como maDMP, acrônimo do inglês *Machine Actionable Data Management Plans*. Os maDMPs buscam melhorar os planos tradicionais, potencializar a qualidade dos dados e metadados de pesquisa por meio da automação e assim, reduzir a sobrecarga administrativa gerada pelo preenchimento e avaliação dos textos presentes dos modelos tradicionais. Ao interligar sistemas de informação no ecossistema científico, os maDMPs trazem benefícios para financiadores, equipes de conselhos de ética, especialistas jurídicos, pesquisadores e equipes de apoio, editores, operadores de repositório, administradores institucionais e provedores de infraestrutura (Miksa, *et al.*, 2019).

No tocante às suas funcionalidades, Miksa, *et al.* (2019) apresentam dez princípios para a aplicação dos maDMPs considerando a variabilidade de escopo e especificidades implícitas em

cada princípio, além das questões técnicas, organizacionais e sociais demandadas conforme a origem e realidade dos locais de aplicação, desenvolvimento da ferramenta e dos planos. Para o funcionamento dos planos acionáveis por máquina, os provedores de infraestrutura devem fornecer sistemas de suporte correspondentes às demandas necessárias de sua execução. A seguir os dez princípios mencionados (Miksa, *et al.*, 2019):

- Integrar PGDs com os fluxos de trabalho de todas as partes interessadas no ecossistema de dados de pesquisa.
- Permitir que sistemas automatizados atuem em nome das partes interessadas.
- Políticas devem considerar os sistemas e interconexões, não apenas versar sobre os atores envolvidos.
- Os componentes do ecossistema de gerenciamento de dados devem ser descritos considerando sistemas, suas interconexões e os atores envolvidos.
- Use identificadores únicos persistentes e vocabulários controlados.
- Sugere-se seguir um modelo de dados comum para maDMPs.
- PGDs devem funcionar voltados para consumo humano e de máquina.
- PGDs devem dar apoio na avaliação e monitoramento do gerenciamento de dados.
- PGDs devem ser atualizáveis, vivos e com versionamento de documentos.
- PGDs devem estar disponíveis ao público em geral.

Os princípios apresentados pelos autores possibilitam perceber claramente a evolução do modelo inicial do PGD estático em formato texto para um acionável por máquina. Outros aspectos a serem considerados na definição conceitual e implementação de um PGD acionável por máquina foram listados pelo Research Data Alliance (RDA), com base em estudos conduzidos por Miska, *et al.* (2020). Ao analisarem dez ferramentas para elaboração de planos de gestão de dados: [Haplo](#), [F1000 Research](#), [DMPTool](#), [DMPonline](#), [DMP Opidor](#), [Data Stewardship Wizard](#), [NSD DMP](#), [Argos](#), [Tu Wien](#) e [Easy Dmp](#), verificaram que as ferramentas contemplam as características básicas de cardinalidade, granularidade, versionamento, embargo, status e série. Entende-se por:

- **Cardinalidade:** Os PGDs que possuem campos preenchidos com informações pré-definidas oriundas de outros sistemas relacionados (obrigatórias ou opcionais).
- **Granularidade:** Depende do contexto da pesquisa e se refere às descrições das características dos conjuntos de dados podendo variar e ser atualizadas conforme o avanço da pesquisa.
- **Versionamento:** Geração de marcadores de data e hora de criação dos PGDs e suas respectivas versões. Os mecanismos de controle de versão atualizam cada vez que os

dados são alterados e permitem recuperar diferentes versões de um PGD com base nos registros dos marcadores.

- *Embargo*: Os dados só podem ser disponibilizados após terem recebido uma atribuição de licença de acesso e uso. Enquanto a licença não for atribuída, os dados são considerados fechados ou inacessíveis;
- *Status do PGD*: Indica se as ações do plano estão em fase de "previstas no planejamento" ou se já foram realizadas.
- *Serializações*: Serialização em Java Script Object Notation (JSON), representação ontológica de conteúdo que pode ser utilizada em qualquer outra representação: Extensible Markup Language (XML), Ontology Web Language (OWL) e JavaScript Object Notation for Linked Data (JSON-LD).

Apresentadas as características e inovações conferidas aos planos de gestão de dados acionáveis por máquina, seguem descrições sobre o percurso metodológico deste estudo que visa definir um modelo conceitual para uma ferramenta centralizada de criação de PGDs acionáveis por máquina a ser disponibilizada para a comunidade científica brasileira.

3 Percurso metodológicos

O presente estudo emprega um relato de experiência do desenvolvimento de modelo conceitual de ferramenta para elaboração de Plano de Gestão de Dados de Pesquisa. Foi utilizado o método Design Science Research, que permite ao investigador pesquisar, descrever ou explicar um fenômeno e, também, projetar ou prescrever soluções para um problema específico (Dresch, *et al.*, 2014). O método Design Science Research objetiva desenvolver soluções para dois tipos de problemas: quesitos práticos, que demandam mudanças que melhor se adequem com os objetivos dos tomadores de decisão, e problemas de conhecimento que necessitem de reinvenções no conhecimento sobre o mundo (Weiringa, 2009). Para esse estudo, identificou-se dois problemas práticos:

- As ferramentas de elaboração de PGDs que resultam em planos estáticos e textuais, sem informações que poderiam ser buscadas automaticamente de sistemas utilizados pelos agentes envolvidos em todo o ecossistema científico.
- Inexistência de uma ferramenta brasileira centralizada que atenda às demandas das diversas partes envolvidas no ecossistema de pesquisa e seja acionável por máquina.

No quadro 1 constam as diretrizes aplicadas ao presente estudo.

ORIENTAÇÃO	DESCRIÇÃO	APLICAÇÃO NO MAPGD
1. Design como artefato	Design Science Research deve produzir um artefato viável com processos de criação, modelo, método ou instanciação.	PGD IBICT: ferramenta para gestão e elaboração de PGDs.
2. Relevância do problema	O objetivo do Design Science Research é desenvolver soluções baseadas em tecnologia para problemas relevantes em empresas e instituições.	Ferramenta PGD acionável por máquina alinhada com necessidades de gestão dos dados das partes interessadas.
3. Avaliação do design	A utilidade, qualidade e eficácia do artefato deve ser rigorosamente demonstrada através de métodos de avaliação.	Estudo comparativo entre as ferramentas PGDs.
4. Contribuições da pesquisa	O Design Science Research deve fornecer contribuições claras e verificáveis nas áreas do artefato, fundamentos e/ou metodologias de design.	Elaboração de conteúdo para divulgação do desenvolvimento da ferramenta PGD.
5. Rigor de pesquisa	Design Science Research é baseada na aplicação de métodos rigorosos na construção e avaliação do artefato.	Serão realizadas avaliações e testes da ferramenta PGD.
6. Design como processo de busca	A busca por um artefato efetivo requer a utilização de meios disponíveis para	Foram avaliadas ferramentas de elaboração de PGDs, utilizou-se conceitos de PGDs acionáveis por máquina e os princípios FAIR para

ORIENTAÇÃO	DESCRIÇÃO	APLICAÇÃO NO MAPGD
	alcançar os objetivos desejados.	pesquisa e desenvolvimento da solução.
7. Comunicação de pesquisa	Design Science Research deve ser apresentado efetivamente para públicos voltados para tecnologia e gestão	Publicação dos resultados e disponibilização para a comunidade

Tabela 1: Design Science Research para concepção e desenvolvimento da ferramenta de criação de PGD acionável por máquina.

Fonte: Elaborado pelos autores baseado em Hevner, March, Ram e Park (2004).

4 Análise e resultados

Os resultados apresentam-se em duas etapas. Primeiramente foi feita uma análise detalhada de conformidade dos principais softwares de elaboração de PGDs disponíveis quanto à melhor adequação e customização para acionamento por máquina. Em seguida apresenta-se a proposta do modelo conceitual para uma ferramenta de elaboração de PGD acionável por máquina, baseada na DMPTool, no cenário brasileiro.

4.1 Análise das ferramentas de elaboração de PGD para se tornar acionável por máquina

As ferramentas selecionadas para elaboração de PGDs foram: DMPTool, DMPOnline, Argos, DataWiz, EasyDMP. Tais ferramentas foram analisadas para verificação de suas funcionalidades, possibilidades de customização e usos. A ferramenta DMPTool foi selecionada para customização e implementação e a seguir as devidas justificativas:

- Segue as recomendações da **RDA DMP Common Standards**, para ser acionável por máquina.
- DMPRoadmap é uma base de código aberto para uma ferramenta de PGD acionável por máquina, gerenciada em conjunto pela Digital Curation Center (DCC) e a University of California Curation Center (UC3), representando esforços para convergir em uma única solução os melhores recursos de versões anteriores do DMPonline e DMPTool.
- DMPRoadmap é usada em vários serviços internacionais, incluindo DMPAssistant no Canadá, DMPTuuli na Finlândia, DMPOPIDoR na França, PGDonline na Espanha e a

versão do DEIC do DMponline na Dinamarca.

- Oferece suporte amplo para a comunidade que está implementando a ferramenta, crucial para os desenvolvedores.
- Possui parceria com a DataCite e construíram um fluxo de trabalho para gerar DOIs que são conectados ao ORCID, sendo recuperados em ferramentas de busca Web.
- É uma ferramenta recomendada por diversas instituições brasileiras de ensino e fomento à pesquisa.

	DMPTOOL	DMP IBICT	DMPONLINE	ARGOS
Personalização com logo	não	sim	não	não
Versionamento	iniciado	iniciado	iniciado	sim
Desenvolvimento relacionado com maDMP (acionável por máquina)	iniciado	iniciado	iniciado	sim
Research Organization Registry (ROR) IDs - para identificar afiliações de organizações de pesquisa	iniciado	previsto	iniciado	não identificado
IDs do Crossref Funder Registry para identificar financiadores de pesquisa	iniciado	previsto	iniciado	iniciado
RDA Common Standard JSON usado para trocar metadados DMP entre ferramentas (easyDMP e Data	iniciado	previsto	iniciado	iniciado

	DMPTOOL	DMP IBICT	DMPONLINE	ARGOS
Stewardship Wizard)				
API DataCite GraphQL - pode expor todas as conexões referentes à pesquisa no PGD	iniciado	previsto	iniciado	não identificado
Taxonomia: Contributor Roles Taxonomy (CRediT)	iniciado	previsto	iniciado	não identificado
Fluxo de trabalho de aprovação de patrocinadores e financiadores em que podem revisar, comentar e aprovar DMPs enviados	iniciado	previsto	não identificado	não identificado
Integração com FAIRsharing	iniciado	previsto	iniciado	iniciado
interoperabilidade entre ferramentas PGD (trocas de metadados)	iniciado	previsto	iniciado	iniciado
Integração com Dataverse	iniciado	previsto	não identificado	iniciado

	DMPTOOL	DMP IBICT	DMPONLINE	ARGOS
Integração com ferramentas de gestão de dados	iniciado	previsto	não identificado	sim
Criação de templates (orientações)	sim	sim	sim	sim
Identificador persistente (DOI)	sim (DMP ID)	DMP ID e DOI	não	
autenticação CAFe	sim	sim	sim	não
GOV.BR (www.gov.br)		previsto		
autenticação e-mail pessoal	sim	sim	sim	sim
ID são automaticamente vinculados ao registro ORCID	sim	iniciado	não	não identificado
Seleção de modelos de PGDs por instituição	não	iniciada	não	iniciada
Inclusão de colaboradores da equipe de pesquisa	sim	sim	sim	sim

	DMPTOOL	DMP IBICT	DMPONLINE	ARGOS
Inclusão de identificador (ORCID) dos colaboradores	iniciado	sim	sim	sim
inclusão de informações dos resultados da pesquisa	sim	sim	não	sim
lista de repositórios	sim (utiliza o Registry of Research Data Repositories - re3data registration)	sim (utiliza o Registry of Research Data Repositories - re3data registration)	sim (utiliza o Registry of Research Data Repositories - re3data registration)	sim
lista de padrões de metadados	sim	sim	RDA Metadata Standards Catalog (opens in a new window)	sim
Seletor de licenças	sim (SPDX)	sim (SPDX) implementando EUDAT	não	sim
modelo de referência/citação para cada PGD	sim	sim	não	

	DMPTOOL	DMP IBICT	DMPONLINE	ARGOS
inclui os outputs da pesquisa	sim	sim	não	sim
Configurações detalhadas de exportação	sim	sim	sim	sim
Definição da visibilidade do plano	sim	sim	sim	sim
código fonte	disponível com licenças abertas	disponível com licenças abertas	disponível com licenças abertas	disponível
Exemplos de DMP disponíveis publicamente	sim	previsto	sim	sim
Comentários	sim	sim	sim	não identificado
Relatórios para as agências de fomento	não identificado	previsto	não identificado	não identificado
Integração com ambiente de preservação digital	não identificado	previsto	não identificado	não identificado

Tabela 2. Análise das funcionalidades das ferramentas de planes de gestão de dados. Fonte: Elaborada pelos autores (2022).

4.2 Proposta de modelo conceitual da ferramenta brasileira de elaboração de PGDs acionáveis por máquina (maPGD)

Definido o caminho na customização do DMPTool, o próximo passo foi delinear o modelo conceitual de funcionamento da ferramenta considerando os atores envolvidos no ecossistema científico brasileiro. Destaca-se o papel fundamental do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em iniciativas de incentivo ao desenvolvimento científico brasileiro alinhadas aos preceitos da Ciência Aberta, em especial a condução e desenvolvimento deste estudo que favorece ações de gestão de dados de pesquisa em todo o país, além de possibilitar a sistematização e cruzamento de informações em Ciência, Tecnologia e Inovação (CTI), fundamentais para as agências de fomento e instituições de pesquisa.

A ferramenta brasileira de elaboração de PGDs está representada na figura 2 como "PGD BR (IBICT)", localizada ao centro propositalmente, uma vez que a ideia é que ela atue como um centralizador de informações referentes às pesquisas científicas brasileiras. Na parte superior do fluxo, tem-se as informações referentes aos pesquisadores e custos envolvidos no projeto, incluindo o acesso à plataforma PGD BR com as credenciais previamente definidas. Tais insumos serão provenientes das seguintes fontes informacionais:

- Comunidade Acadêmica Federada (CAFe), a primeira federação acadêmica do país, sendo hoje uma das cinco maiores do mundo em números de clientes (RNP, 2022).
- ORCID (Open Researcher and Contributor Identifier), que além de fornecer um identificador único ao pesquisador, vincula-o com suas informações profissionais integradas (afiliações, bolsas, publicações científicas, revisões por pares previamente cadastradas em uma espécie de currículo de alcance internacional). Estando interligado com a ferramenta PGD, o ORCID poderá oferecer todas estas informações de maneira sistematizada automaticamente.
- Currículo Lattes, principal plataforma de gestão de currículos brasileira.
- GovBr, serviço governamental brasileiro de identificação em meios digitais. Permite identificação com segurança na hora de acessar serviços digitais governamentais (Ministério da Gestão e da Inovação em Serviços, 2022).

No lado esquerdo da figura 2, tem-se as agências de fomento como fornecedoras de informações referentes às chamadas de financiamento, projetos financiados e seus respectivos vínculos com a equipe de pesquisadores. Destaca-se aqui a possibilidade de registro de um identificador único das agências de fomento via recurso [Crossref Founder Registry](#), favorecendo a parametrização destas informações. Destaca-se aqui a bidirecionalidade da relação entre o PGD BR e os sistemas das agências de fomento uma vez que, além de fornecerem informações, estes agentes também podem fazer uso das informações provenientes dos PGDs

especialmente para geração de indicadores e estatísticas de financiamento.

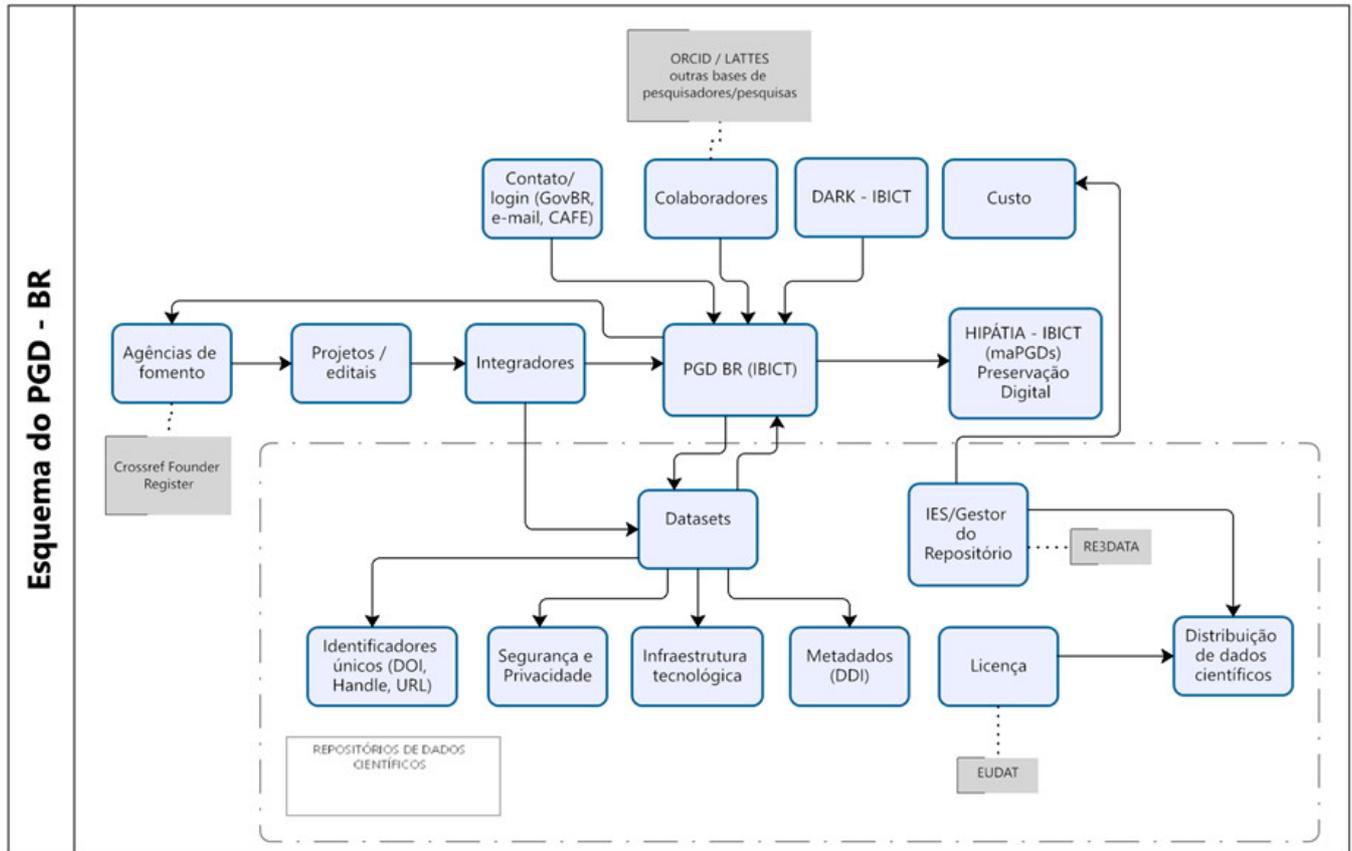


Figura 2. Modelo conceitual do PGD IBICT.

Fonte: dos autores, baseada em Miksa, Walk e Neish (2020).

No lado inferior da figura, tem-se a interligação dos PGDs com as Instituições de ensino e pesquisa que poderão atuar como gestoras de repositórios de publicações e dados científicos. Estes sistemas de armazenamento são essenciais no contexto das pesquisas científicas uma vez que os pesquisadores podem depositar seus dados e demais publicações favorecendo o compartilhamento e disponibilização, tendo a Ciência Aberta como pano de fundo de todo processo de abertura do fazer científico.

Destaca-se a complexidade e diversidade no processo descritivo dos dados e demais publicações nestes repositórios por meio de metadados que obedecem padrões internacionais em conformidade com princípios de descoberta e reuso. Tais repositórios abrangem identificadores persistentes, licenças de acesso e uso (que podem ser obtidas de bases de dados automaticamente oferecendo as opções que melhor se adequem às situações dos projetos de pesquisa), requisitos de segurança, confiabilidade dos dados, controle de versionamento, buscas multifacetadas, dentre outros recursos que enriquecem os dados. Destaca-se aqui também a possibilidade de bidirecionalidade da relação entre os PGDs e os

repositórios uma vez que além de fornecerem informações, estes agentes também as recebem dos PGDs de maneira automática alimentando os metadados dos conjuntos de dados (*datasets*). Um exemplo desta troca de informações são os DOI (identificadores persistentes) dos conjuntos de dados (atribuídos pelos repositórios), quando o pesquisador informa no PGD BR qual será o repositório escolhido para futuro armazenamento. Assim, será possível alocar espaço de armazenamento neste repositório ainda na etapa de planejamento da pesquisa.

Finalizando a descrição do modelo conceitual, tem-se o IBICT ao lado direito da figura como órgão que disponibiliza a solução tecnológica de gestão e elaboração de PGDs. Destaca-se que este órgão também proverá, uma solução de preservação digital do PGD BR, visando garantir o acesso ao longo do tempo, utilizando o modelo Hipátia.

O modelo Hipátia, desenvolvido pelo IBICT, é uma iniciativa que cria uma camada de barramento tecnológico interoperável para automatizar o processo de preservação de objetos digitais. Além disso, existe compatibilidade com o modelo referencial **OAIS** de preservação digital incorporando a estruturação de Repositórios Arquivísticos Digitais Confiáveis, os quais fazem parte da etapa final de armazenamento, garantindo a integridade, autenticidade, confiabilidade e segurança dos dados armazenados em um sistema, neste caso especificamente, o PGD BR. Um dos softwares compatíveis para estes repositórios no final da cadeia de custódia é o **Archivematica**. Essa solução tecnológica pode se adaptar a qualquer sistema que faça gestão de processos e objetos digitais.

5 Considerações Finais

O Plano de Gestão de Dados acionável por máquina (maDMP/ maPGD) definido conceitualmente para o cenário brasileiro tem como premissa básica ser centralizador de informações entre sistemas dos diversos agentes envolvidos no ecossistema de Ciência, considerando as infraestruturas tecnológicas disponíveis e as possibilidades de utilização dos recursos por parte de pesquisadores de diferentes áreas do conhecimento, instituições de ensino e pesquisa, agências de fomento, dentre outras. Optou-se por utilizar uma ferramenta livre já existente que conta com uma rede robusta de colaboradores e suporte, o DMPTool. Entende-se que tal decisão poderá otimizar e agilizar o processo de desenvolvimento rumo à uma ferramenta acionável por máquina, contribuir para o amadurecimento de padrões e definições deste tipo de solução tecnológica ao nível internacional.

No contexto brasileiro, a rede federada de instituições de ensino e pesquisa conferem uma utilização consorciada inovadora a este tipo de ferramenta de elaboração de PGDs permitindo avanços no amadurecimento do desenvolvimento e uso da ferramenta por parte dos diversos agentes envolvidos. Outro destaque é a possibilidade de descoberta das funcionalidades e

necessidades de cada agente envolvido no ecossistema brasileiro de Ciência, o que pode conferir possibilidades de uso e compartilhamento de informações de maneira enriquecedora. Ressalta-se enfim que uma ferramenta maPGD, customizada com base neste modelo conceitual, estará ajustada para interligar agentes, produtos, serviços e demais entidades envolvidas no ecossistema científico por meio de padrões, diretrizes e melhores práticas difundidas em nível internacional, como os princípios FAIR, esquemas de metadados previamente definidos, identificadores únicos, dentre outros.

Bibliografia

Dresch, Aline; Lacerda, Daniel Pacheco; Antunes, José Antônio Valle (2015). "Chapter 4. Design science research". En: *Design Science Research: a method for science and technology advanced*. Switzerland: Springer, p. 67–102.

DOAJ. *Directory of Open Access Journals* (2022). "Periódicos por países dos editores: Brasil". <https://doaj.org/search/journals?source=%7B%22query%22%3A%7B%22query_string%22%3A%7B%22query%22%3A%22Brasil%22%2C%22default_operator%22%3A%22AND%22%7D%7D%2C%22size%22%3A50%2C%22sort%22%3A%5B%7B%22created_date%22%3A%7B%22order%22%3A%22desc%22%7D%7D%5D%2C%22track_total_hits%22%3Atrue%7D>. [Consulta: 28/09/2022].

Hevner, Alan; March, Salvatore T.; Park, Jinsoo; Ram, Sudha (2004). "Design Science in Information Systems Research". *Management Information Systems Quarterly*, vol. 28, no. 1, p. 75–105.

Miksa, Tomasz; Walk, Paul; Neish, Peter (2020). *RDA DMP Common Standard for Machine-actionable Data Management Plans*. <https://doi.org/10.15497/rda00039>.

Miksa, Tomasz; Simms, Stephanie; Mietchen, Daniel; Jones, Sarah (2019). "Ten principles for machine-actionable data management plans". *PLoS Computational Biology*, vol. 15, no. 3, e1006750. <https://doi.org/10.1371/journal.pcbi.1006750>.

Ministério da Gestão e da Inovação em Serviços (2022). *Governo digital: Gov.br*. <<https://www.gov.br/governodigital/pt-br/conta-gov-br>>. [Consulta: 30/09/2022].

Pinto, Carlos Sousa; Costa, Joaquim Luís (2018). "Capítulo 5. Padrões de comunicação em diferentes comunidades científicas". En: Costa, Sely Maria de Souza; Leite, Fernando César Lima; Tavares, Rosemeire Barbpsa (eds.). *Comunicação da informação, gestão da informação e gestão do conhecimento*. Brasília: IBICT. p. 145–159. <https://doi.org/10.18225/9788570131485>

RNP (2022). *Comunidade acadêmica federada: rede CAFe*. <<https://www.rnp.br/servicos/cafe>>. [Consulta: 30/09/2022].

UNESCO (2021). *Recomendações da UNESCO sobre ciência aberta*. <https://unesdoc.unesco.org/ark:/48223/pf0000379949_por>. [Consulta: 27/09/2022].

Wieringa, Roel (2009). "Design science as nested problem solving". *DESRIST '09: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*. Philadelphia, May 2009. Article no. 8.
<https://doi.org/10.1145/1555619.1555630>

Articles similars a BiD

- **Competències, coneixements i carta de serveis del personal bibliotecari de dades**. Serrat i Brustenga, Marta. (2023)
- **Les universitats espanyoles i la ciència oberta : un estudi sobre barreres i elements afavoridors**. Abad García, María Francisca; González Teruel, Aurora; Abadal, Ernest; Ollé i Castellà, Candela. (2022)
- **Govern obert i accés a la informació : un estudi de cas sobre l'impacte en l'economia local**. Ferrer Sapena, Antonia; Calabuig, Jose M.; Sánchez Pérez, Enrique A.; Vidal Cabo, Christian. (2020)
- **L'educació oberta a Europa : avenços, integració amb la ciència oberta i rol bibliotecari**. Santos Hermosa, Gema. (2019)

Articles similars a Temària

- **Gestión documental de un grupo de investigación biomédica**. Romero Cuevas, Miguel; Reviriego Barquilla, Fernando; Marín Martínez, María; Rodríguez de Fonseca, Fernando. (2010)
- **La gestión documental en procesos estratégicos de mejora : un agente de oportunidad en una institución pública de educación superior : el caso de CIDE**. Revah Lacouture, Magdalena. (2009)
- **El sistema de gestió documental de la Universitat de Catalunya : una funció de suport a la creació de patrimoni intangible?**. Andreu i Daufí, Jordi. (2009)
- **La información para empresa y el sistema de I+D+I : el caso de la región valenciana**. Muñoz Cañavate, Antonio. (2009)
- **Un nuevo cometido para los bibliotecarios académicos : data curation**. Martínez Uribe, Luis; Macdonald, Stuart. (2008)

Articles del mateix autor a Temària

Rezende, Laura Vilela Rodrigues, Monteiro, Elizabete Cristina de Souza de Aguiar, Stueber, Ketlen, Silva, Fabiano Couto Corrêa da, Grácio, Alexandre Faria de Oliveira, Olivera, Alexandre Faria de

[[més informació](#)]



Creative Commons licence (**Attribution-Non-Commercial-No Derivative works**).

They may be consulted and distributed freely provided that the author and publisher are quoted (in accordance with the "Recommended citation" section in

each of the articles). However, no derivative works (translation, change of format, etc.) may be made without the publisher's permission. Therefore, it meets the definition of open access form the Budapest Open Access Initiative declaration. **The journal allows the author(s) to hold the copyright without restrictions and to retain publishing rights** without restrictions.



Facultat d'Informació i Mitjans Audiovisuals (UB)

Melcior de Palau, 140, 08014-

Barcelona (revistabid@ub.edu)

Graus

Màsters

Doctorat

Altres estudis

Recerca

Estudis de Ciències de la Informació i de la Comunicació (UOC)

Rambla del Poble Nou, 156, 08018-

Barcelona (publicacions@uoc.edu)

Oferta formativa

Coneix la UOC

Recerca i innovació

Estudia a la UOC



Entrada