

Construção do perfil de metadados do repositório brasileiro de dados da pesquisa em ciência do solo (SoilData)

Construction of the metadata profile for the Brazilian Soil Science Research Data Repository (SoilData)

Gabriele Vitória Silva Serena¹, Aline Mari Huf dos Reis², Débora Liriel Kerber Kempner³, Taciara Zborowski Horst⁴, Alessandro Samuel-Rosa⁵

RESUMO

O gerenciamento eficaz de conjuntos de dados em repositórios é essencial para promover a ciência do solo no Brasil. Os metadados desempenham um papel crucial ao facilitar a organização e recuperação desses dados, fomentando a pesquisa e o reuso. O repositório brasileiro de dados da pesquisa sobre o solo (SoilData), tem trabalhado incansavelmente para aprimorar sua plataforma de busca, resultando em maior reuso de dados e no avanço da ciência do solo. Com um total de treze metadados obrigatórios e cinco recomendados, a busca por arquivos tornou-se mais eficiente. Agora, é possível pesquisar conjuntos de dados com base na localização, data, estado e até pela classificação do solo. Esses esforços refletem o compromisso contínuo em aperfeiçoar o acesso e a qualidade dos recursos relacionados ao solo. A plataforma SoilData está capacitando pesquisadores, estudantes e profissionais para realizar o depósito de seus conjuntos de dados de forma correta, impulsionando o progresso da ciência do solo e sua aplicação em desafios ambientais e agrícolas.

PALAVRAS-CHAVE: Ciência aberta, Dataverse, Metadados.

ABSTRACT

The effective management of data sets in repositories is essential for advancing soil science in Brazil. Metadata plays a crucial role in facilitating the organization and retrieval of these data, promoting research and reuse. The Brazilian soil research data repository (SoilData) has been tirelessly working to enhance its search platform. This will result in increased data reuse and the advancement of soil science. With a total of ten mandatory and eleven recommended metadata fields, the search for files has become more efficient. Now, it is possible to search for data sets based on location, date, state, and even soil categorization. These efforts reflect the ongoing commitment to improving access and the quality of soil-related resources. The SoilData platform is empowering researchers and professionals to conduct more detailed analyses, driving progress in soil science and its application in environmental and agricultural challenges. The continuous pursuit of excellence in this repository is making a significant contribution to the scientific community and sustainable development.

KEYWORDS: Open science, Dataverse, Metadata.

INTRODUÇÃO

O Brasil possui uma vasta diversidade de solos em seu território. Os dados de solo são ferramentas importantes capazes de fornecer uma análise confiável, segura e de

¹ Iniciação científica voluntária. Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil. E-mail: gabiserena16@outlook.com. ID Lattes: 6313962630892604.

² Pesquisadora de Pós-Doutorado. Universidade Federal de Goiás, Goiânia, Goiás, Brasil. E-mail: huf.aline@gmail.com. ID Lattes: 1570834132484121.

³ Bolsista da Fundação Araucária. Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil. E-mail: deboraliriel@alunos.utfpr.edu.br. ID Lattes: 2569410943832682.

⁴ Docente do curso de bacharelado em Agronomia e Engenharia Florestal da Universidade Tecnológica Federal do Paraná, Dois Vizinhos, Paraná, Brasil. E-mail: tacihorst@gmail.com. ID Lattes: 6763043931071514.

⁵ Docente do curso de bacharelado em Agronomia da Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil. E-mail: alessandrosamuelrosa@gmail.com. ID Lattes: 1609751519717461.

qualidade com embasamento técnico e científico. Ter um repositório de dados de solo é crucial para garantir que as informações armazenadas sejam facilmente acessíveis, usadas e compartilhadas por pesquisadores e outros interessados de forma eficiente (ANJOS, 2022).

Um repositório de dados é onde se armazenam e gerenciam dados importantes de forma eficaz, disponibilizando-os para reuso por indivíduos ou organizações, promovendo a colaboração e o compartilhamento dessas informações (NAEEM, 2020). Preservar dados de pesquisa a longo prazo requer uma infraestrutura técnica e administrativa abrangente que cubra todo o ciclo de vida dos dados (NASCIMENTO; QUEIROZ; ARAUJO, 2019).

Metadados de métodos descrevem como os dados foram gerados, incluindo amostragem, procedimentos laboratoriais e correções aplicadas antes do uso, essenciais para definir se dados de diferentes origens podem ser combinados e o método mais adequado para fazê-lo. Esses metadados são essenciais para garantir a utilização segura dos dados da pesquisa no futuro, para evitar que pessoas interessadas enfrentem trabalho adicional, ineficaz ou impossível devido à falta de informações adicionais dos produtores, ou, na pior das hipóteses, devido ao falecimento destes.

O repositório brasileiro de dados da pesquisa em ciência do solo, até então conhecido como FEBR, recentemente passou por uma remodelação completa da sua infraestrutura tecnológica (ANJOS, 2022). O objetivo do presente trabalho foi construir o perfil de metadados do repositório de dados, que agora é designado SoilData ([SoilData](#)) e integra o rol de produtos e serviços oferecidos pela Rede MapBiomias.

MATERIAL E MÉTODOS

O Projeto Dataverse está comprometido em usar metadados compatíveis com o padrão para assegurar que os metadados de uma instalação do Dataverse possam ser facilmente mapeados para esquemas de metadados padrão e exportados para o formato JSON (XML para metadados de arquivo tabular) para preservação e interoperabilidade.

Os repositórios de dados são plataformas autônomas, principalmente devido à necessidade de enriquecer os metadados, que possibilitam o compartilhamento com outros repositórios (SALES, 2014). A organização de informações sobre metadados na Web envolve a criação de esquemas de metadados para representar as características e atributos de objetos reais, a fim de estabelecer um modelo que facilite a recuperação posterior (ALVES, 2010). Isso contribui para melhorar a eficácia e a eficiência dos serviços de informação, tornando essencial o uso de metadados semânticos, que empregam várias formas e estruturas de metadados em diferentes serviços e softwares. A padronização é um elemento chave para a utilização eficaz dos metadados (SANTOS; SIMIONATO; ARAKAKI, 2014).

O perfil de metadados do SoilData foi construído utilizando o rol de metadados de citação e geoespaciais disponíveis no Dataverse. Metadados de citação são os metadados referentes ao conjunto de dados, já os geoespaciais são metadados referentes a geolocalização sobre a área de estudo. Os metadados com requisitos de obrigatoriedade são sinalizados por um asterisco (*) vermelho, os metadados opcionais não são obrigatórios serem informados (*Figura 1*).

Figura 1. Exemplo de metadados obrigatórios (indicados com asterisco) e opcionais do Dataverse .

Metadados de citações ^

Título * ?

Digite o título ...

Adicionar "Dados de replicação para" ao título

Autor * ?

(Nome) * ?

FamilyName, GivenName ou Organizatio

Afiliação ?

[en]



Esquema de Identificadores ?

Selecione ...

Identificador ?

[en]

Entre em contato * ?

Nome ?

FamilyName, GivenName ou Organizatio

Afiliação ?

[en]



E-mail * ?

[en]

Fonte: Elaborado pelos autores (2023).

A decisão sobre quais dos metadados do Dataverse manter como obrigatórios ou opcionais foi realizada com base em três critérios: 1) as configurações internas do Dataverse, 2) a experiência pretérita das equipes de gestão do repositório e curadoria de dados no antigo FEBR e 3) as políticas do SoilData.

RESULTADOS

METADADOS OBRIGATÓRIOS

Os metadados obrigatórios referem-se a dados que devem ser registrados, são cinco metadados requeridos pelo Dataverse, e outros oito metadados definidos pela equipe que seriam obrigatórios. Desta forma, atuam de forma fundamental na organização, gestão e controle de informações, tornando-os mais compreensíveis e acessíveis (*Tabela 1*).

Tabela 1. Metadados de uso obrigatório no SoilData.

Tipo	Metadado	Descrição
Citação	Título*	Nome do trabalho ou nome do conjunto de dados.
Citação	Nome do autor*	Responsável pelo conteúdo, pode conter um ou mais.
Citação	Email de contato*	Comunicação com o usuário que está cadastrando o conjunto de dados.
Citação	Texto de descrição dos dados*	Auxilia na interpretação e compreensão dos dados.
Citação	Assunto/tema*	Área do conhecimento que representa o conjunto de dados.
Citação	Depositante	Nome do usuário que realizou o depósito do conjunto de dados.
Citação	Data do depósito	Data do depósito do conjunto de dados no SoilData.
Citação	Palavras-chave	Termos-chave que descrevem aspectos importantes do conjunto de dados.
Citação	Classificação de tópicos	Ajuda a identificar o escopo do trabalho (área da ciência do solo).
Citação	Período coberto	Tempo ao qual os dados se referem.
Geoespacial	País	País de origem dos dados.
Geoespaciais	Estado	Estado de origem dos dados.
Geoespaciais	Cidade	Cidade de origem dos dados.

* Metadado requerido pelo Dataverse.
Fonte: Elaborado pelos autores (2023).

METADADOS RECOMENDADOS

O Dataverse não prevê a configuração de metadados recomendados. No entanto, considerando a experiência do pretérito das equipes de gestão do repositório e curadoria de dados no antigo FEBR, juntamente com as políticas do SoilData, identificamos um rol de metadados opcionais de grande importância. Embora sejam metadados opcionais, após a submissão dos dados, a equipe entra em contato para recomendar o registro, a fim de aumentar a visibilidade das instituições que produzem, financiam os autores e as publicações relacionadas (*Tabela 2*).



Tabela 2. Metadados de uso recomendado no SoilData.

Tipo	Metadado	Descrição
Citação	Afiliação	Organização o qual o contato é afiliado.
Citação	Esquema de identificadores	Identificadores como (ORCID, ISNI, DAI)
Citação	Identificador	Identificação exclusiva de um autor ou organização.
Citação	Publicação relacionada	Publicações que utilizam o conjunto destes dados.
Citação	Informação de financiamento	Valorização das fontes dos recursos financeiros utilizados

Fonte: Elaborado pelos autores (2023).

METADADOS OPCIONAIS

Metadados opcionais são informações adicionais que não são estritamente necessárias para identificação, mas a inclusão deles pode ser útil para aprofundar a compreensão e o contexto dos dados. Por exemplo, o metadado da *publicação relacionada* ajuda a utilizar o conjunto destes dados, aumentando as chances de serem localizados e utilizados por pessoas interessadas. A decisão de registrar ou não esses metadados fica a cargo dos autores dos dados. Assim como os metadados recomendados, alguns metadados opcionais podem se tornar obrigatórios ou recomendados.

CONSIDERAÇÕES FINAIS

Com o passar do tempo, alguns metadados recomendados podem se tornar obrigatórios, como, por exemplo, a afiliação. Atualmente, esse metadado não é obrigatório, uma vez que pode não haver informações acessíveis sobre a organização à qual o autor dos dados estava associado à época da realização do trabalho. Para avançar no tópico em questão, consideramos uma oportunidade de expandir os esforços iniciais deste estudo por meio de pesquisas que abordem a necessidade de tornar os dados facilmente acessíveis a humanos, máquinas e reutilizáveis. Esses assuntos podem beneficiar no desenvolvimento e na gestão do SoilData. Além disso, outros estudos que podem se beneficiar com o desenvolvimento do SoilData incluem políticas de preservação de dados e análises das demandas públicas e privadas em relação aos dados.

Agradecimentos

À UTFPR, ao Instituto Arapyaú e ao CNPq pelo financiamento do projeto, à Fundação Araucária, FUNTEF, FUNAPE e IAMAP pelo suporte. Em especial, agradecemos o Vitor Peruzzi de Almeida, e Marcos Alexandre dos Anjos, que ficaram fora da lista de co-autores em função do limite imposto pela plataforma.

Conflito de interesse

Não há conflito de interesse.

REFERÊNCIAS

ALVES, R. C. V. [UNESP]. **Metadados como elementos do processo de catalogação**. Faculdade de Filosofia e Ciências: Universidade Estadual Paulista, 15 mar. 2010.

ANJOS, M. A. **REPOSITÓRIO DE DADOS DO SOLO BRASILEIRO: CONCEPÇÃO E IMPLEMENTAÇÃO**. Santa Helena - Paraná: Universidade Tecnológica Federal do Paraná, 2022.

NAEEM, T. **Repositório de dados: importância, desafios e práticas recomendadas**. **Astera**, 7 nov. 2020. Disponível em: <<https://www.astera.com/pt/tipo/blog/reposit%C3%B3rio-de-dados/>>. Acesso em: 19 set. 2023

NASCIMENTO, A. G. DO; QUEIROZ, C. F. DE; ARAUJO, L. D. DE. **A importância dos Repositórios e da Preservação Digital para a pesquisa científica**. Universidade Federal Fluminense, 2019. Disponível em: <<https://www.arca.fiocruz.br/handle/icict/38900>>. Acesso em: 19 set. 2023.

SALES, L. F. **Integração semântica de publicações científicas e dados de pesquisa: proposta de modelo de publicação ampliada para a área de Ciências Nucleares**. Programa de Pós-Graduação em Ciência da Informação: Instituto Brasileiro de Informação em Ciência e Tecnologia/Universidade Federal do Rio de Janeiro, 23 jul. 2014.

SANTOS, P. L. V. A. D. C.; SIMIONATO, A. C.; ARAKAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. **Informação & Informação**, v. 19, n. 1, p. 146, 25 fev. 2014.