



## Curadoria Digital de Dados Agropecuários

Hugo Henrique Teixeira **Sanches**<sup>1</sup>; Debora Pignatari **Drucker**<sup>2</sup> Carla Geovana do Nascimento  
**Macario**<sup>3</sup>  
**Nº 23605**

**RESUMO** – A publicação de dados de pesquisa em repositórios confiáveis vem se tornando cada vez mais relevante e estratégico para as instituições. Este estudo apresenta desafios da curadoria digital de dados agropecuários a partir da experiência do Repositório de Dados da Embrapa - Redape. O estudo utilizou os princípios F do FAIR<sup>4</sup> (Findable (Encontrável), Accessible (Acessível), Interoperable (Interoperável), Reusable (Reutilizável)) como base metodológica para analisar os dados publicados no Redape. A aderência a cada subprincípio associado ao princípio F foi analisada e foram identificados aqueles que estão bem atendidos e aqueles que precisam ser aprimorados. O Redape contém 14 campos obrigatórios, porém, os autores não preenchem corretamente todos os campos: existem erros comuns entre os campos preenchidos, sendo a imprecisão do preenchimento das palavras-chaves a falha mais com. Os resultados revelaram a aderência do Redape aos subprincípios de F1 (Os metadados e dados são identificados por identificador digital persistente, de forma única e global), F3 (Os metadados incluem identificadores dos dados por eles descritos) e F4 (Os metadados são indexados por sistemas de pesquisa), os quais são fornecidos pela própria ferramenta. O subprincípio F2 (os dados são descritos com metadados ricos) depende da curadoria para ser devidamente satisfeito, uma vez que, na maioria das vezes, os autores não inserem as palavras-chave com base nos vocabulários controlados que constam nas orientações. Essa análise ressalta a importância da curadoria na gestão de dados e metadados de pesquisa, considerando o subprincípio F2. Conclui-se que a curadoria desempenha um papel relevante na gestão de dados agropecuários, mas há oportunidades para aprimorar a padronização dos metadados e facilitar a integração e reutilização de conjuntos de dados agropecuários.

**Palavras-chave:** Curadoria digital; Encontrabilidade; Acessibilidade; Interoperabilidade; Padronização; Metadados.

<sup>1</sup>Autor, Bolsista CNPq (PIBIC): Graduação em Engenharia Agrícola, UNICAMP, Campinas-SP; hhtsanches01@gmail.com

<sup>2</sup>Analista da Embrapa Agricultura Digital, Campinas-SP, debora.drucker@embrapa.br .

<sup>3</sup>Orientadora: Pesquisadora da Embrapa Agricultura Digital, Campinas-SP; carla.macario@embrapa.br.

<sup>4</sup>Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>



**ABSTRACT** – *The publication of research data in reliable repositories is becoming increasingly relevant and strategic for institutions. This study presents challenges in the digital curation of agricultural data based on the experience of the Embrapa Data Repository - Redape. The study used the FAIR principles (Findable, Accessible, Interoperable, Reusable) as a methodological basis to analyze the data published in Redape. Adherence to each subprinciple associated with principle F was analyzed, identifying those that were well-addressed and those that need improvement. Redape contains 14 mandatory fields; however, the authors do not correctly fill in all the fields. There are common errors among the filled fields, with imprecise keyword entries being the most common failure. The results revealed the adherence of Redape to the F1 sub-principle (Metadata and data are identified by a persistent digital identifier, uniquely and globally), F3 sub-principle (Metadata includes identifiers for the data they describe), and F4 sub-principle (Metadata is indexed by research systems), which are provided by the tool itself. The F2 sub-principle (Data is described with rich metadata) depends on curation to be properly satisfied, as authors often do not enter keywords based on the controlled vocabularies listed in the guidelines. This analysis highlights the importance of curation in the management of data and research metadata, considering the F2 subprinciple. It is concluded that curation plays a relevant role in the management of agricultural data, but there are opportunities to improve metadata standardization and facilitate the integration and reuse of agricultural datasets.*

**Keywords:** *Digital curation; Findability; Accessibility; Interoperability; Standardization; Metadata.*

**Observação:** *A responsabilidade pela revisão ortográfica do resumo simples é dos autores.*