

QUALIDADE E GOVERNANÇA DE DADOS CIENTÍFICOS

desafios estruturais no ecossistema BrCris

QUALITY AND GOVERNANCE OF SCIENTIFIC DATA

Structural Challenges in the BrCris Ecosystem

Thiago Magela Rodrigues Dias

Centro Federal de Educação Tecnológica de Minas Gerais, CEFET-MG

E-mail: thiagomagela@gmail.com

ORCID: <https://orcid.org/0000-0001-5057-9936>

Washington Luis R. de Carvalho Segundo

Instituto Brasileiro de Informação em Ciência e Tecnologia, IBICT

E-mail: washingtonsegundo@ibict.br

Marcel Garcia de Souza

Instituto Brasileiro de Informação em Ciência e Tecnologia, IBICT

E-mail: marcelsouza@ibict.br

Priscila Machado Borges Sena

Instituto Brasileiro de Informação em Ciência e Tecnologia, IBICT

E-mail: priscilasena@ibict.br

RESUMO

A consolidação da Ciência Aberta depende não apenas da ampliação do acesso aos dados científicos, mas sobretudo da garantia de sua qualidade, integridade e confiabilidade. Este estudo analisa os desafios associados à gestão e governança de dados no ecossistema BrCris, uma infraestrutura nacional de integração de informações científicas brasileiras. A partir da análise de dados provenientes de múltiplas fontes, como Plataforma Lattes, OpenAlex e Oasisbr, são identificados problemas recorrentes relacionados à duplicidade de registros, inconsistência de metadados, ausência de identificadores persistentes e desatualização das informações. Esses fatores impactam diretamente a precisão das análises bibliométricas e cientométricas, comprometendo diagnósticos e decisões estratégicas. Como resposta, são discutidas estratégias tecnológicas e metodológicas voltadas à qualificação dos dados, incluindo deduplicação automatizada, desambiguação de entidades, validação de metadados e adoção de padrões internacionais. O estudo contribui para o fortalecimento da governança da informação científica, destacando o

papel dos sistemas CRIS na promoção da integridade informacional e no enfrentamento da desinformação científica.

Palavras-chave: Qualidade de dados, Governança da informação, Ciência Aberta, CRIS, BrCris, Bibliometria.

ABSTRACT

The consolidation of Open Science depends not only on expanding access to scientific data, but above all on ensuring their quality, integrity, and reliability. This study analyzes the challenges associated with data management and governance in the BrCris ecosystem, a national infrastructure for integrating Brazilian scientific information. Through the analysis of data from multiple sources, such as the Lattes Platform, OpenAlex, and Oasisbr, recurring problems are identified involving duplicate records, inconsistent metadata, the absence of persistent identifiers, and outdated information. These factors directly affect the accuracy of bibliometric and scientometric analyses, compromising assessments and strategic decisions. In response, technological and methodological strategies for improving data quality are discussed, including automated deduplication, entity disambiguation, metadata validation, and the adoption of international standards. The study contributes to strengthening the governance of scientific information, highlighting the role of CRIS systems in promoting information integrity and addressing scientific disinformation.

Keywords: Data quality, Information governance, Open Science, CRIS, BrCris, Bibliometrics.

REFERÊNCIAS

COLLAZO-REYES, F. Growth of the number of indexed journals of Latin America and the Caribbean: The effect on the impact of each country. **Scientometrics**, v. 98, p. 197–209, 2014. DOI: <https://doi.org/10.1007/s11192-013-1036-2>.

SARAVIA, N. G.; MIRANDA, J. F. Plumbing the brain drain. **Bulletin of the World Health Organization**, v. 82, p. 608–615, 2004. Disponível em: <https://www.scielosp.org/article/bwho/2004.v82n8/608-615/>.