

ESTRATÉGIAS DE ORGANIZAÇÃO E ENRIQUECIMENTO DE DADOS CURRICULARES PARA ANÁLISE DE MOBILIDADE CIENTÍFICA

STRATEGIES FOR ORGANIZING AND ENRICHING CURRICULUM VITAE DATA TO ANALYZE SCIENTIFIC MOBILITY

Higor Alexandre Duarte Mascarenhas

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

E-mail: higoralexandre1996@gmail.com

ORCID: <https://orcid.org/0000-0001-6057-3888>

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>

RESUMO

A análise da mobilidade científica depende da disponibilidade de bases de dados estruturadas capazes de representar trajetórias acadêmicas de pesquisadores ao longo do tempo. No Brasil, a Plataforma Lattes constitui uma das principais fontes de informação sobre formação e produção científica. Entretanto, os dados disponíveis nos currículos apresentam heterogeneidade estrutural, exigindo processos de organização e enriquecimento para viabilizar análises em larga escala. Este artigo apresenta estratégias para organização, tratamento e enriquecimento de dados curriculares extraídos da Plataforma Lattes, com o objetivo de possibilitar estudos sobre mobilidade científica. A partir da extração de currículos de indivíduos com doutorado concluído, foram aplicados procedimentos de filtragem, normalização de dados, geolocalização de instituições e integração com bases externas. Os resultados demonstram que a estruturação desses dados permite reconstruir trajetórias acadêmicas e identificar padrões de mobilidade entre instituições, regiões e países. As estratégias propostas contribuem para ampliar o uso de dados curriculares em estudos cientométricos e na geração de indicadores sobre mobilidade acadêmica.

Palavras-chave: Plataforma Lattes, Organização de Dados, Análise de dados, Mobilidade Científica.

ABSTRACT

The analysis of scientific mobility depends on the availability of structured databases capable of representing researchers' academic trajectories over time. In Brazil, the Lattes Platform is one of the main sources of information on academic training and scientific output. However, the data available in curricula vitae are structurally heterogeneous, requiring organization and enrichment processes to enable large-scale analysis. This article presents strategies for organizing, processing, and enriching curriculum vitae data extracted from the Lattes Platform, with the aim of enabling studies on scientific mobility. After extracting the curricula vitae of individuals who had completed a doctorate, procedures were applied for filtering, data normalization, institutional geolocation, and integration with external databases. The results demonstrate that structuring these data makes it possible to reconstruct academic trajectories and identify mobility patterns across institutions, regions, and countries. The proposed strategies contribute to expanding the use of curriculum vitae data in scientometric studies and in the production of indicators of academic mobility.

Keywords: Lattes Platform, Data Organization, Data Analysis, Scientific Mobility.

REFERÊNCIAS

- ADAMS, W. **The brain drain**. New York: Macmillan, 1968.
- DIAS, T. M. R. LattesDataXplorer: uma abordagem para análise de redes científicas. 2016.
- HAYASHI, M. C. P. I. Bibliometria e cientometria: aplicações e indicadores científicos. 2012.
- LOMBAS, M. Mobilidade acadêmica e circulação de pesquisadores. 2017.
- MEYER, J. B. Network approach versus brain drain: lessons from the diaspora. International Migration, 2001.
- VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual. **Ciência da Informação**, 2002.
- WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.