Evaluating Publisher-Generated Metadata Quality in Spanish Law Journals: A Comparative Study Using Crossref, OpenAlex and Dialnet

Ángel Borrego[1,*,†], Cristóbal Urbano[1,†]

[1]*Universitat de Barcelona, Departament de Biblioteconomia, Documentació i Comunicació Audiovisual*

Abstract

This ongoing project aims to evaluate the quality and completeness of publisher-generated metadata in Spanish law journals by comparing records from Crossref and OpenAlex with those curated by Dialnet, a major Spanish bibliographic platform. Using a mixed-methods approach, we analyze metadata elements such as authorship, abstracts, references and ORCID identifiers across a corpus of journals. The expected findings may help identify inconsistencies and gaps, particularly in metadata standardization and completeness, which hinder the discoverability and interoperability of legal scholarship.

Keywords

Crossref, Dialnet, legal publishing, publisher-generated metadata, metadata quality, OpenAlex, scholarly communication Spanish law journals

Metadata quality plays a critical role in the visibility, discoverability and interoperability of scholarly outputs. This study assesses the quantity and quality of metadata created by Spanish law journal publishers when they assign DOI identifiers to their articles. By examining the metadata available in Crossref (https://search.crossref.org) and OpenAlex (https://openalex.org), and comparing it with metadata curated by Dialnet (https://dialnet.unirioja.es), a major bibliographic platform in Spain, this research identifies key gaps, inconsistencies and best practices in metadata provision by Spanish legal publishers. The findings highlight challenges related to standardization and completeness, especially within legal scholarship, and propose recommendations to enhance metadata quality for improved research dissemination and academic visibility.

Legal scholarship in Spain, often published in national journals with limited international reach, faces specific challenges regarding metadata quality and interoperability. There are also persistent problems in tracking Spanish legal literature through bibliometric networks based on references and citations. Despite the growing emphasis on open access and digital discoverability, legal journals frequently exhibit poor metadata quality, or lack metadata entirely,

limiting their visibility in global indexing systems. This issue is particularly relevant for Crossref-registered publishers, whose metadata feeds into systems like OpenAlex, an open scholarly index increasingly used in research analytics and discovery tools.

This study has three main objectives:

- To analyze the quantity and completeness of metadata deposited by Spanish law journals in Crossref and made accessible through OpenAlex;
- To evaluate the quality of this metadata in terms of standard elements (title, authorship, affiliations, abstracts, subject keywords, bibliographic references, ORCIDs, funding information, etc.);
- To compare the metadata from Crossref/OpenAlex with the corresponding records in Dialnet in order to assess differences in richness, accuracy and consistency.

To carry out this study, we first constructed a corpus of Spanish law journals that were indexed both in Dialnet and in Crossref, with metadata available through Crossref (https://api.crossref.org) and OpenAlex (https://docs.openalex.org/how-to-use-the-api/api-overview) APIs. For each journal, we extracted article-level metadata and systematically analyzed key fields such as title, authors, publication date, ORCID identifiers, abstracts, references, language and funding information. We evaluated the presence, completeness and consistency of these elements, applying a comparative framework to assess discrepancies between the metadata provided through Crossref/OpenAlex and that available in Dialnet. Quantitative measures, such as metadata completeness scores, were combined with qualitative observations regarding field normalization and inconsistencies. In addition, we selected a small number of journals as case studies to illustrate typical metadata patterns, as well as exemplary and problematic practices. This mixed-methods approach allowed us to identify both systemic issues and journal-specific variations in metadata quality and coverage.

The study provides empirical evidence on the current state of metadata practices in Spanish legal publishing, highlighting how these practices affect the visibility and discoverability of scholarly outputs. It offers a detailed analysis of common metadata deficiencies, categorizing them and discussing their potential impact on research dissemination and academic indexing. By comparing metadata from Crossref/OpenAlex with that curated by Dialnet, the study reveals the added value and limitations of aggregator platforms in enriching and standardizing metadata. The findings will inform a set of practical recommendations aimed at publishers, journal editors and repository managers to improve metadata creation and management workflows. Ultimately, this research will contribute to ongoing efforts to strengthen the quality and interoperability of scholarly metadata in the legal domain and beyond.

References

[1] J. Beel, B. Gipp, E. Wilde, Academic search engine optimization (¨ASEO), Journal of Scholarly Publishing 41 (2010) 176–190. doi:`10.3138/jsp.41.2.176`.

[2] T. Bergstrom, O. Y. Rieger, R. C. Schonfeld, The Second Digital Transformation of Scholarly Publishing: Strategic Context and Shared Infrastructure, Technical Report, Ithaka S+R, 2024. doi:`10.18665/sr.320210`.

[3] V. Craigle, Adopting DOI in legal citation: A roadmap for the legal academy, https://papers.ssrn.com/abstract=3800465, 2021. SSRN Scholarly Paper No. 3800465.

[4] L. Delgado-Quirós, J. L. Ortega, Completeness degree of publication metadata in eight free-access scholarly databases, Quantitative Science Studies 5 (2024) 31–49. doi:10.1162/qss_a_00286.

[5] E. T. Hart, Indexing open access law journals... or maybe not, International Journal of Legal Information 38 (2010) 19–42. doi:10.1017/S0731126500005539.

[6] B. J. Keele, Improving digital publishing of legal scholarship, Legal Reference Services Quarterly 34 (2015) 119–137. doi:10.1080/0270319X.2015.1038143.

[7] T. Marks, A. Le, Increasing article findability online: The four Cs of search engine optimization, Law Library Journal 109 (2017) 83–100. doi:10.2139/ssrn.3204550.

[8] A. Nishikawa-Pacher, Law reviews, open metadata and RSS feeds current interests, Legal Information Management 22 (2022) 156–165. doi:10.1017/S1472669622000305.

[9] A. Retteen, M. Hall-Retteen, Persistent identifiers and the next generation of legal scholarship, Law Library Journal 116 (2024) 133–158. URL: https://scholarship.law.tamu.edu/facscholar/1994.