

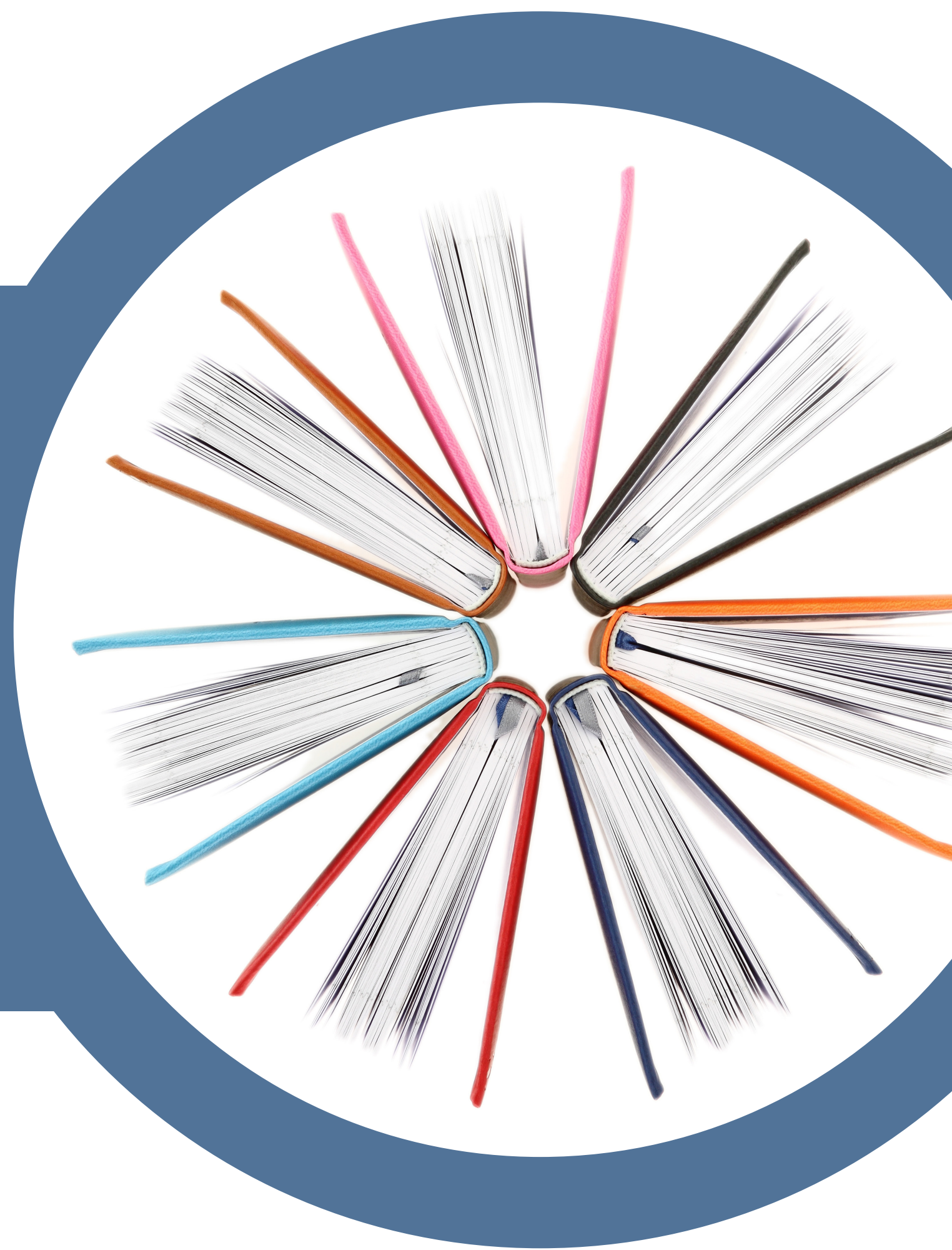
OPTIMIZING METADATA ENHANCEMENT WORKFLOW OF CURRENT RESEARCH INFORMATION SYSTEM VIA CROSSREF

Summary

This pilot study investigated the potential for reducing human intervention in maintaining metadata quality within a Current Research Information System (CRIS) by utilizing CrossRef metadata ingestion. 100 research output records were randomly selected from Scopus, focusing on five key Dublin Core elements: **title, creator, description, publisher** and **date**.

The findings revealed high accuracy in the title and creator fields, with minimal discrepancies, primarily related to complex publications like books and group authorships. However, significant issues were identified in the description, publisher, and date fields.

The description field often lacked content or contained machine-readable data that impeded human accessibility due to schema incompatibilities. Discrepancies in the publisher field arose from inconsistent naming practices between publishers and CRIS administrators, potentially affecting open access compliance reporting.



The date field exhibited substantial inconsistencies, notably the omission of online publication dates critical for compliance with research evaluation frameworks like the UK's REF. Additionally, some records were entirely absent from CrossRef, highlighting limitations in relying solely on this source for comprehensive metadata harvesting.

Conclusion

The pilot study concludes that, while automation is feasible for certain metadata elements, human oversight remains essential to ensure accuracy and full compliance. Future research should explore the metadata quality from other third-party sources such as Scopus, Web of Science, and Dimensions to further assess opportunities for reducing manual intervention in the metadata quality assurance process.