

WS-06

A primer to using open scholarly metadata for research integrity

Evans Atoni¹, Madhura Amdekar²

¹Crossref, Kenya

²Crossref, Netherlands

This workshop is for researchers, research integrity officers, members of research funding organisations, and others, to introduce them to scholarly metadata and the tools and techniques they can use to retrieve metadata for assessing the integrity of research outputs.

The scholarly record consists of research outputs such as journal articles, book chapters, datasets, software, preprints, and more. Metadata about each of these outputs also forms a part of the scholarly record and conveys important information about these items. Crossref is a community-governed, not-for-profit membership organisation that enables its members to register metadata and persistent identifiers (DOIs) for the work that they publish. Made up of over 165 million research outputs from over 21,000 members, the metadata registered with Crossref is a rich and open source of data for researchers and several downstream services. In addition to basic metadata such as titles, publication dates, names of authors, journal titles, members also deposit metadata on funding information, post-publication updates (retractions and corrections), relationships between research outputs (e.g. “is review of”, “is preprint of”) and more. This data is made openly available via Crossref’s APIs and therefore can be a valuable resource for assessing the integrity of published work – providing context for individual outputs, as well as for conducting research on research at scale.

In this session, we will introduce the participants to the metadata that is deposited by the scholarly community with Crossref, and how this can act as trust signals about scholarly works. We will provide an overview of the metadata elements that are of most relevance in the context of research integrity, such as affiliation metadata, funding metadata, retraction metadata (including retractions and corrections from the Retraction Watch database which has been acquired by Crossref and integrated into our REST API), and how they are connected to each other. We will demonstrate how to construct topical API queries to retrieve this metadata to answer questions such as:

Which works have been funded by specific funders?

Which of these funded works have been retracted?

Which works have been retracted?

How to find works associated with your institution?

How to find works published by a specific research organisation that have been retracted?

During the remainder of the session, the participants will have the opportunity to try their hands at building API queries to retrieve metadata of their interest and ask any questions. We hope that with information from this session, not only will the participants learn how to use open metadata to track the outputs of their research institutions, but they will also be able to see the value of metadata as a tool available to them for assessing the integrity of work produced by the scholarly community.