26th International Conference on Science and Technology Indicators | STI 2022
"From Global Indicators to Local Applications"
7-9 September 2022 | Granada, Spain
#STI22GRX

SPECIAL TRACK: Present and future of research metadata: where do we want to go from here?

Alberto Martín-Martín[*], Benjamín Vargas-Quesada[*], Zaida Chinchilla-Rodríguez[**] and Manuel Jesús Cobo[***]

[*] albertomartin@ugr.es
[*] benjamin@ugr.es
Facultad de Comunicación y Documentación, Universidad de Granada (Spain)

[**] zaida.chinchilla@csic.es
Instituto de Políticas y Bienes Públicos (IPP), Consejo Superior de Investigaciones Científicas (CSIC), Albasanz 26-28, Madrid, 28037 (Spain)

[***] mjcobo@ugr.es
Departamento de Lenguajes y Sistemas Informáticos, Universidad de Granada (Spain)

To steer the scientific system towards specific goals, it is first necessary to develop an effective understanding of all phases and aspects of the research workflow. Research metadata, as the collective record of traces that are generated when scientific activities take place, serves as evidence of these activities. Therefore, the availability of authoritative research metadata is essential for science-related decision-making at various levels.

In the past, large-scale research metadata collections mostly dealt with items in the public record, such as bibliographic metadata about academic publications. There used to be few of these large-scale metadata collections, and they were often provided by commercial actors, those which invested the necessary resources to compile and process disperse public information with the goal of turning it into usable services.

As the capabilities of available technologies increase, each day more sectors of the scientific system are becoming aware of how their activities could benefit from updating their workflows, a process often referred to as digital transformation. Thus, a plethora of tools and standards are being developed to streamline processes, increase interoperability, and in general overcome the limitations of the paper era. This is having a large effect in the quantity and quality of research metadata that is now being recorded.

A clear example of the above is the case of bibliographic metadata. Currently, an increasing number of organizations, spurred by the decreasing barriers to collecting and processing large amounts of bibliographic metadata, are already providing services and datasets that rival the offerings of the traditional commercial providers. Some of these new datasets, provided under open licenses that allow unrestricted reuse and redistribution, have boosted innovation by allowing the development of downstream applications that rely on these metadata collections.

However, as scientific activities in general and scientific communication in particular are increasingly moving to the digital space, traditional bibliographic metadata is no longer the

only kind of research metadata that is being collected and processed at a large scale to inform decisions. Social network platforms now capture a portion of academic-related conversations and other kinds of interactions. Processes such as peer review that were previously carried out behind closed doors are now being opened, generating their own public trace. Publishing platforms are implementing increasingly sophisticated methods to track and mine user actions for their benefit.

All these recent developments call for a discussion on the role of research metadata in the scientific system going forward. This discussion should be open to a large variety of stakeholders, including data providers, scientometricians, academic librarians, higher education institutions, policy managers, and developers of downstream applications.

The topics of the contributions to this special track can include:

- Analyses of the suitability of research metadata sources for specific use cases
- Sustainability and governance of research metadata
- Innovations in research metadata
- Downstream applications of open research metadata
- Surveillance through research metadata

Contributions to this special track would be open to everyone interested and peer-reviewed. The format of the session would be 15-20 minutes per presentation, with time for questions after each presentation.