Persistent conference identifiers in CRIS – first implementations with VIVO and URIS

Sabina Auhunas, Stephanie Hagemann-Wilholt, Christian Hauschke, Benjamin Kampe, Tetiana Rudakova, Viktor Shapovalov, Yevhenii Shapovalov

Introduction

Academic events are an essential part of scholarly communication: They do not only facilitate the dissemination of scientific results but also foster networking and exchange within the scientific community. Despite their important role in advancing academic endeavors, the significance of these events is not adequately reflected in the quality of the metadata that is generally available to describe them. For example, a survey of Ukrainian researchers and conference organisers revealed that there are frequent errors in metadata and a need for improvement in terms of better data management using persistent identifiers such as ORCID (~30% of organizers and ~36% of researchers) and DOI (~30% of organizers and ~2% of researchers) (Auhunas, 2024).

The ConfIDent project aims to remedy this situation by creating a registry for high-quality conference metadata with persistent identifiers (PIDs). In a first funding phase RWTH Aachen and TIB set up a prototype of a registry for academic event metadata.[1] ConfIDent provides structured metadata on conferences, workshops and similar formats of scientific exchange. Its features include to browse, archive and announce events. It is based on Semantic MediaWiki (SMW) and enables collaborative data curation by members of scientific communities and professional data curators while offering transparency on provenance of each entry. Furthermore, it allows users to assign subject-specific tags to events which are systematically mapped to a subject classification. ConfIDent considers itself as an open research information infrastructure and seeks to make event metadata findable and reusable in accordance with the FAIR principles and to ensure interoperability and accessibility: Accordingly, it offers both an open API and metadata under a free license (CC0). ConfIDent minted approx. 5.000 DOIs for events and event series in a first test corpus, e.g. linking events to hosting organizations, and related events. The metadata schema is aligned with the DataCite metadata schema as well as schema.org and is based on an elaborate data model for capturing event entities.

Using Conference PIDs in Research Information Systems (RIS)

One of the main purposes of RIS is usually research evaluation, whose traditional implementation with simple indicators such as the listing of citation counts has been subject to strong criticism for several years. In recent years, there have been increasing calls for diversification both in terms of methodology and in terms of the activities and outcomes to be assessed (Leiden Manifesto, COARA etc.). For example, in the Hong Kong Principles for the Assessment of Researchers (Moher et al. 2020), one of the five principles is to recognise the essential tasks of the research enterprise - and organizing and participating in academic events is certainly one of them. This means that the clear identification of events and links is becoming increasingly important.

Case 1: Implementation in the Forschungsatlas

The Forschungsatlas[2] is a prototype for a subject based RIS for the communities of researchers in architecture, civil engineering and urbanism. Technically, it was implemented

[1] https://www.confident-conference.org, cf. Franken et al. (2022), Hagemann-Wilholt et al. (2019)
[2] https://forschungsatlas.fid-bau.de/, cf. Arndt et al. (2022)

with the linked-data-based open source software VIVO (Conlon et al. 2019). The Forschungsatlas was customized to fit the needs of the target community. One is that it queries the ConfIDent API to get Events that are of interest for the community and another that it displays the geographic location of these events on a map. Retrieving and transforming these data is fully automated via the VIVO-Harvester[3]. The process includes calling the ConfIDent API, parsing the JSON response into RDF-XML, using XSLT to map the data to the VIVO Ontology, calculating the delta to the previous harvests and finally ingesting the new elements into the VIVO triplestore. Each entry includes details like the academic field, event title, series, dates, mode (online, on-site, hybrid), venue, city, region, country, official website, and DOI. The events span from 2020 to 2024, covering themes like construction, engineering, urbanism, building design, and computational architecture, among others.

Case 2: Implementation in the Ukraine Research Information System (URIS)

The Ukrainian Research Information System (URIS) is a service that collects, manages, stores, and analyzes information about research being carried out by Ukrainian researchers (Kaliuzhna & Auhunas 2022). Since 2022, as part of URIS, a module for registering and monitoring providers and academic events has been developed. A key element is the development of persistent identifiers for academic events.

In 2023, two stages of the module development were successfully implemented: research of the general architecture of the academic events module, including the development of the architecture of the State register of academic events, and the development and selection of the model of the academic events platform (Shapovalov 2023). In 2024, it is planned to implement stages 3 and 4, which include the analysis of functional features and technical capabilities for a specific list of sources, APIs, data formats, and access conditions, as well as the analysis of the methodology for importing and exporting data from external information systems and databases ConfIDent, which forms the basis for further integration and automation of filling the academic events module in URIS with international academic events. Development of algorithms for the procedure for state registration of academic events (conferences). Research was conducted to develop basic metrics for academic events.

Outlook

It is very likely that academic events will play a greater role as an independent entity in the field of RIS. The use of PIDs for events is an opportunity here to avoid problems with the disambiguation and deduplication of events. The systems presented are two examples of regional and subject-orientated RIS; institutional RIS will follow.

References

Arndt, S. et al. (2022): FID Civil Engineering, Architecture and Urbanism digital - A platform for science (BAUdigital). In: Research Ideas and Outcomes (Vol. 8). https://doi.org/10.3897/rio.8.e82563

Auhunas, S. (2024, January 25). Evaluation and impact of descriptive metadata on academic event management in Ukraine: A quantitative study. https://doi.org/10.31235/osf.io/fxb74

[3] https://github.com/vivo-community/VIVO-Harvester

Conlon, M. et al. (2019): VIVO: a system for research discovery. In: JOSS 4 (39), p. 1182. https://doi.org/10.21105/joss.01182

Franken, J. et al. (2022): Persistent Identification for Conferences. In: Data Science Journal (Vol. 21). Ubiquity Press, Ltd. https://doi.org/10.5334/dsj-2022-011

Hagemann-Wilholt, S. et al. (2019): ConfIDent – An Open Platform for FAIR Conference Metadata. Twentieth-First International Conference on Grey Literature “Open Science Encompasses New Forms of Grey Literature”, Hannover, Germany, October 22-23, 2019, S. 47-51 (GL Conference Series ; 21)

Hauschke, C. et al. (2021): Roadmap to FAIR Research Information in Open Infrastructures. In: Journal of Library Metadata, 21(1–2), 45–61. https://doi.org/10.1080/19386389.2021.1999156

Kaliuzhna, N., and Auhunas, S. (2022). Research Information Infrastructure in Ukraine: First steps toward building a national CRIS. In: Procedia Computer Science, 211, 230-237. https://doi.org/10.1016/j.procs.2022.10.196.

Moher, D. et al. (2020): The Hong Kong Principles for assessing researchers: Fostering research integrity. In: PLoS biology 18 (7), e3000737. https://doi.org/10.1371/journal.pbio.3000737

Shapovalov, Y. B. et al. (2023): An academic events sub-system of the URIS and its ontology representation to improve scientific usability and motivation of scientists in terms of European integration. In: 3rd Edge Computing Workshop 2023, 130–140.