

Enlarging the spectrum. Implementing a local extension of ROR as identification instrument for additional actors in Flemish (SSH) research

Peter Aspeslagh

peter.aspeslagh@uantwerpen.be

Centre for R&D Monitoring, University of Antwerp, Middelheimlaan 1, 2020 Antwerp (Belgium)

Abstract

International organization databases, like the Research Organization Registry (ROR), are essential tools to identify unique organizations for a wide range of purposes. Our experience with the structural enrichment of author affiliation data for SSH publications pointed to opportunities that a local extension of an international database can provide. Applying a step-by-step approach, we are steadily extending the original set of Flemish/Belgian ROR organizations, building a dataset that contains unique organization identifiers covering a broader spectrum of research, educational and governmental organizations, linking multiple datasets in a national environment. In this poster, we will present the context and methodology related to the creation of this new instrument, the Flemish Organization Registry.

Introduction

In 2019, a parameter measuring international collaboration was added to the Flemish performance-based research funding system (PRFS). Therefore, for SSH disciplines, author affiliation data had to be provided for the full set of publications included in the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB). This dataset is a compilation of approved publications authored by all researchers affiliated to a SSH faculty at a Flemish university. However, author affiliation data was not available in VABB. As only less than half of the VABB publications are included in the Web of Science (Aspeslagh & Guns, 2023), we launched a large and multifaceted author affiliation data collection operation. Browsing alternative databases like Scopus or OpenAlex

only resulted in data for a limited percentage of publications; most of the cases required manual intervention for the retrieval of the actual affiliation data.

During the subsequent coding we assigned unique (ROR-)identifiers to each author affiliation. However, the coding process showed that only two thirds of the affiliated organizations were covered by ROR. In order to allow the completion of the project – and to comply with the new parameter in the Flemish PRFS – new organizations were added to a local copy of ROR and coded according to the ROR data scheme. For the 2012 to 2022 time window, 34.7% of the unique organizations found in author affiliation data of the non-Web of Science publications did not have an identifier in ROR (n=2,348).

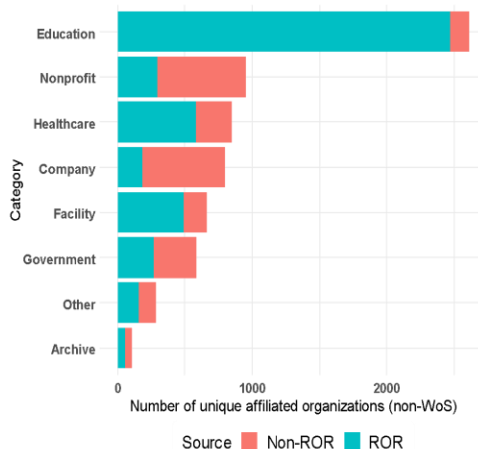


Figure 1. Number of unique affiliated organizations for non-WoS publications (ROR vs non-ROR organizations).

When taking into account the (ROR) type variable as categorization, ROR very much covers the Education category (Figure 1), but

less so for Company, Nonprofit and Government.

National extension

The addition of non-ROR organizations for a complete coverage of author affiliation data demonstrated that there is room for a national extension of an international database. Such an expansion provided a new opportunity. The novel unique identifiers and related metadata could be valorised in a broader context: not only in function of a PRFS parameter, but also as tool to map the full spectrum of organizations that are involved in research in Flanders, independent of their inclusion in common international databases. Too often, organization data is scattered over different lists with distinct identifiers: lists with organizations only having a EU Participant Identification Code (PIC), selections of the entities included in the Belgian commercial register (KBO) etc. With the development of a new instrument, the Flemish Organization Registry (FOR), an integration of different lists is envisaged.

Compilation

Due to the recurrent author affiliation data collection, a platform for hosting, managing and enriching organization data was already available. This allowed us to shift from a technical to a content-related perspective, focussing on the addition of an extensive set of Flemish/Belgian organizations not included in the current dataset.

During a first phase, we made an inventory of the different commonly used organization lists in a broad research, governmental and educational context in Flanders. Deployed for diverse purposes, these lists often contained equal organizations but were not (entirely) interoperable.

Secondly, the selected new datasets, as well as the extended ROR database, were consolidated both by matching via the available metadata or by manual intervention. Often, this completed types of organizations of which our original database contained some, but not all. For example, 20 municipal administrations, found as author affiliations during the affiliation data collection project, were added as these organizations (local governments) were not available in ROR. The consolidation phase provided the remaining

561 Belgian municipal administrations. When future registering of an author affiliated to the 21st municipal administration, the organization will be available in the database. In a third step, which is ongoing, we are uploading the newly consolidated organizations to the main FOR database. Relevant metadata that was available in the original lists is also being added to the database, which is compatible with the ROR data model. Often, this results in a single organization ID now containing a PIC, KBO, FOR and, if available, a ROR identifier. It will allow the continued addition of metadata and enabling, among other features, customized categorization.

Governance and future developments

As FOR can serve diverse purposes in multiple entities in the Flemish governmental and research landscape, a governance framework is being established to ensure the extended use of this project.

Even if FOR places an emphasis on Flemish/Belgian organizations, the relevant part of the new additions and related metadata can be transferred to international databases.

The database allows the continued addition of new variables and categories, which can be used to add new dimensions to future bibliometric studies from an organizational perspective.

Conclusion

With this extended organization database, the efforts invested in the implementation of a modification in the Flemish PRFS are being extra valorised. The format created for the registration of additional organizations during the author affiliation data collection is now being developed into an instrument to be used for multiple purposes. The step-by-step approach to compile and consolidate a diverse set of organization lists results in a Flemish Organization Registry, enabling the unique identification of a broader spectrum of actors in Flemish (SSH) research. While the focus is currently on Flemish/Belgian organizations, the data scheme is compatible with ROR with the possibility to transfer relevant data to external databases.

Reference

Aspeslagh, P. & Guns, R. (2023). How international is co-authorship outside the Web of Science? The case of social sciences and humanities in Flanders, Belgium. Proceedings of ISSI 2023 – the

9th International Conference of the International Society for Scientometrics and Informetrics, 2, 29-36.

<https://doi.org/10.5281/zenodo.8350379>