

International Register of Academic Book Publishers (IRAP): overview, current state and future challenges

Elea Giménez-Toledo¹, Gunnar Sivertsen² and Jorge Mañana-Rodríguez³

¹*elea.gimenez@cchs.csic.es*

Research Group on Scholarly Books (ILIA), Institute of Philosophy (IFS). Spanish National Research Council (CSIC). Madrid, Spain

²*gunnar.sivertsen@nifu.no*

Nordic Institute for Studies in Innovation, Research and Education. Oslo, Norway

³*jorge.mannana@cchs.csic.es*

Research Group on Scholarly Books (ILIA), Institute of Philosophy (IFS). Spanish National Research Council (CSIC). Madrid, Spain

Abstract

In this contribution, the authors present the background, main underlying concept, data sources, current state and technical and scientific challenges of the International Register of Book Publishers (IRAP) project. This project aims at the creation of a register of scholarly book publishers collecting, normalizing and aggregating different data sources used for evaluation at the national or supra-national level, both for basic research on scholarly books as publication channel and for the provision of aggregated information to all stakeholders involved with scholarly publishing.

Introduction

Looking at the current international trends in research evaluation in the social sciences and humanities (SSH), we observe tendencies to: a) consider full data sources and different types of research results, not only traditional journal publications and metrics; 2) emphasize the societal impact of research; 3) diversify the sources for analyzing outputs and for obtaining indicators (REF, DORA, Leiden, Metric Tide, ENRESSH Manifesto, etc.). All these tendencies have the potential of recognizing the importance of books for scholarly communication in SSH, as well as the diversity of publishers in which humanists and social scientists publish to reach their relevant audiences. The closeness of SSH to the societies and cultures being studied may often imply publishing in local or national languages and with book publishers from the region or country. The diversity of publishers used in SSH can even be seen as requirement for societal impact and for responsible research and innovation. The diversity of scholarly publishers in a country is needed, not only because they publish scientific knowledge which other publishers with a more international profile would not publish, but because they help fulfil the aims of the research itself by communicating with society. These scholarly publishers have an important role in the broader national book market. Given the concentration in the international book market, where a few editorial groups and imprints amass a large portion of the market (*Coufal, 2017*), it seems necessary to facilitate an adequate safeguard of the existing diversity. In this sense, the defense of the publication of scholarly books in the national framework for the aforementioned reasons should also be accompanied by recognition of those publishers and their contribution to research evaluation processes as the select and improve manuscripts before publishing.

The International Register of Academic Book Publishers (IRAP) is being developed as a response to this situation. The register is an initiative within the COST action ENRESSH which has the more general aim of improving the basis and methods for research evaluation in SSH. IRAP in particular is intended to provide structured, precise and quality information on scholarly book publishers, mainly in Europe, but also in United States, Canada and Latin America with the objective of facilitating scientific research, showing the editorial diversity of the different countries. This research in progress paper aims to present the state of the art as well as the methodological challenges involved in the development of this project.

International sources for academic books

The two main commercial products providing indicators for scholarly book publishers and individual books are Book Citation Index (Clarivate Analytics) and Scopus Title Expansion Project (Elsevier). The analysis of aspects such as diversity of languages of publication and countries of origin of the publishers revealed some strong biases towards the inclusion of publishers from English-Speaking countries, particularly those that specialize in STEM fields. (Leydesdorff, L., & Felt, U., 2012; Torres Salinas et al., 2014). These findings are consistent with previous analyses of the bibliometric products for journals from the same companies (Moed 2005; Oppenheim and Summers 2008).

Given the relevance of national languages for the SSH, those biases diminish the suitability of the databases for evaluation purposes at the national level in many European countries. As a result, several initiatives have been developed in recent years in several European countries in order to create information systems that allow the provision of reliable information for evaluation purposes. In example, Norway, Finland, Denmark or Belgium (Flanders region) count with Current Research Information Systems (CRIS) which integrate the whole research publication output of the country, thus providing complete data (Sivertsen, 2016). The publication channels (journals or publishers, in example) are then rated in terms of quality by expert panels (Sīle, L. et al. 2017; Sīle, L. et al. 2018)

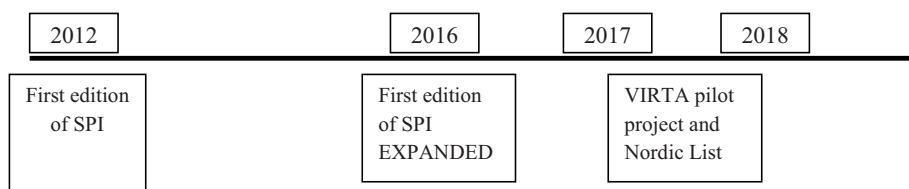
Several other initiatives and approaches have been developed in different countries (Giménez –Toledo et al., 2017; Giménez-Toledo et al., 2018). In Spain, ILIA Research Group (Research Group on Scholarly Books, Spanish National Research Council) has developed *Scholarly Publishers Indicators* (SPI) (Giménez-Toledo et al. 2012); it provides information on the perceived prestige of Spanish and non-Spanish scholarly book publishers, information concerning the manuscript selection processes used by the publishers, the thematic specialization and the presence or absence of a given publisher in five information systems (SPI Expanded, from 2016 onwards). These developments are considered as a reference by the main research evaluation agency in Spain (ANECA).

The development of SPI Expanded has allowed the verification of the potential interest of a register of publishers at the European level, which reflects the presence or absence of the different publishers in various information systems used for evaluation purposes, together with attached information on the quality level that the publisher has in each information system. The Nordic countries have developed a ‘Nordic List’ merging the respective lists of publication channels from their respective countries through a project funded by NORDFORSK. Finally, there is also a clear interest at the European level in the creation of common infrastructures allowing the convergence in terms of information for research evaluation, as Puuska et al. 2018 (p.1) point out.

International Register of Academic Book Publishers (IRAP)

IRAP is a research proposal and technical development for improving the evaluation of scholarly books while preserving national book industries. Taking into account the described context, the relevance of books in SSH, the need of considering diversity in publishing channels and also the growing research on academic book publishing, a working group of ENRESSH COST action¹ is working on the development of the International Register of Academic Book Publishers. It aims at: a) listing all relevant academic book publishers, it is to say, actually used or to be used by researchers b) providing basic bibliographical information c) offering relevant information for research evaluation purposes such us manuscript selection processes or other useful and transparent information to evaluation agencies, academic institutions, etc.

Figure 1. IRAP history timeline



Methodological issues and current state of the Register

SPI Expanded (http://ilia.cchs.csic.es/SPI/expanded_index.html) is the starting point for building up the register. The following table (1) reflects the key features of the Register at its current stage.

Table 1. Key features of the European Register of Scholarly Publishers at its current state.

<i>Feature</i>	<i>Values</i>
Number of distinct publishers	5917
Number of different countries of publishers in the Register	110
Number and percentage of University Presses	766 (12.94%)
Number and percentage of publishers with set of ISBN prefixes	2420 (40.89%)

Table 2. Number of distinct publishers in each source.

<i>Database</i>	<i>Number of different publishers</i>
BFI (Denmark)	1371
Publication Forum (Finland)	2747
NSD (Norway)	2891
SPI (Spain)	1097
VABB-SHW (Flanders)	134
Book Citation Index	467
Scopus Title Expansion Program	341

Current sources included in the register are: Book Citation Index (2018); Scopus Title Expansion Program (2018); Norwegian NSD (2018) lists; Finnish Publication Forum lists (2018); Danish BFI lists (2018); Spanish Scholarly Publishers Lists (2014); Flemish VABB-SHW lists (2016). The expansion of the register is presently based on the aggregation of information from other existing lists of book publishers at the national level (linked to databases for evaluation purposes). After these expansions, the project has now come to a stage where the scholarly publishers themselves, represented by their international organizations, can be invited to take part in the project. The project is also in contact with ERIH PLUS, the European register of journals in the SSH, which has for a long time planned to include a register of scholarly book publishers.

The construction of the register implies in the first place a work of presentation of the project as well as establishing contact with those responsible of the national information systems (CRIS or equivalent sources), with national organizations in charge of scientific evaluation and with associations of scholarly publishers.

The inclusion of publishers to IRAP through various authorized methods for data importation and crawling implies several phases of technical treatment of the information, but also some challenges from the point of view of research.

Technical challenges

The aggregation of publishers from different sources requires data cleansing and normalization of publishers' names following the common procedures to that effect. This first normalization is followed by a second, manual stage which has the objective of disambiguation of those cases not identified in the first step. In this second step, the official publisher's names available at the Global Register of Publishers (International ISBN Agency) are used.

The depuration of the data from the different source has presented a series of challenges to its reliability

a) Single name and de-duplication

A single name for each publisher is a desirable condition for a register of publishers. Nevertheless, disambiguation and de-duplication of individual names has been one of the main sources of concern in the process of unifying the different sources. On the one hand, the way in which a given publisher is written can take several forms depending on the inclusion of acronyms and the use of common abbreviations for company type. Also, changes in the names of publishers keeping their activity intact are a source of error in the de-duplication process. In those cases, the main source of information for de-duplicating the names have been the use of the Global Register of Publishers (GRP; <https://grp.isbn-international.org/>), the largest authoritative list of publishers, developed and updated by the International ISBN Agency. It is a challenge for the development of the Register the identification of an optimal process of disambiguation.

b) Imprints and publishing groups

A second type of error sources are those related to the imprints. Many publishing groups have acquired smaller, independent publishers during the course of their

business history. Generally, those previously independent publishers (with their set of ISBN prefixes) are included within the publishing group as imprints. With a single name and a series of ISBN prefixes it is possible to find independent publishers, up to a given date and imprints, from the point in time when the publisher was incorporated or merged into a publishing group. The treatment of such instances requires a case by case review of the publishers' history and, apart from time-consuming, the results are not always clear.

c) Co-editions

Co-editions are a further source of error: the co-editing publishers can be kept as independent publishers, the publisher associated to the ISBN prefix can be kept and the rest discarded or co-publications can be discarded beforehand. It remains a challenge to determine which option would be optimal taking into account the different pros and cons of each approach.

Some research challenges

The CRIS systems used in the Nordic countries and Flanders count with complete data on publication in each country. In the case of Norway, Finland and Denmark, the processes and criteria for the classification of publishers are similar. On the other hand, SPI counts with a completely different approach, based on a survey to Spanish scholars on the prestige of both Spanish and foreign publishers. In the case of Flanders, the GRPC provides a source of recognition of individual books but its interpretation in terms of quality or prestige of the publisher is different from the previous ones. Book Citation Index and Scopus provide a completely different set of indicators, based on citation counts and several other information systems include publishers without indicators allowing the categorization or classification of the publishers. It seems clear that the levels, quality labels, citation counts or prestige of the publishers are not comparable. Nevertheless, the presence or absence of the publishers in the databases is driven by an intentional selection process or, in the case of the CRIS-based systems the publishers are rated in a scale according to their quality level and, on the other, the presence of a given publisher in the highest positions in all sources or, by the opposite, in the lower positions can be understood as a potentially useful information (at least in the extreme cases). The option taken until now is to attach the information on the quality of the publisher to its name when available so that, given the proper conditions for its aggregation or further use; there is the possibility of counting with such aggregated data. Furthermore, as pointed out in Mañana-Rodríguez & Pölönen, 2018, the data on the quality 'national' publishers can be imputed into the judgment-based evaluation systems of other countries which do not count with specific information on the quality of foreign, maybe linguistically or culturally distant publishers.

Other relevant point in the development of the register is how to tackle with the information regarding on manuscript selection processes within publishing houses, a critical issue for research evaluation purposes (Giménez-Toledo, Sivertsen & Mañana-Rodríguez). Opening the debate on standards in academic publishing of books entails, among other issues, questioning of peer review and distinguishing the role book editors from the journals editors. A text is not the same published by a publisher or other. The editors provide quality, correction, style or rigor. A scholarly book is part of a publisher and the 'brand' it prints on it (Calasso, 2015). Pointing out the differences between book publishing and journal publishing is a way of breaking down some inertia in research evaluation and science policy. This topic deserves some research for showing

different selection practices within publishing houses –apart from peer review- and their relationship with quality or academic recognition. Results from this research might have positive effects in the development of IRAP.

Future expansion of the Register

Different research projects under development are going to offer results on most relevant academic publishers in Colombia (Giménez-Toledo, 2018) and Brazil (Borges de Oliveira, 2018) and other countries in Latin America. It is foreseen to take these results into account for providing information to the Register.

Also, an initial exploration of the information systems used for evaluation in Croatia, Slovenia, Slovakia and the Czech Republic allows concluding that it would be possible to count with structured sets of scholarly publishers from these information systems. On the other hand, further research on the information systems used for evaluation in other European countries such as the UK, France, Italy, Austria or Germany would provide an opportunity for the broadening of the scope of the Register.

References

- Borges de Oliveira, Aline (2018). *Los libros en Ciencias Sociales y Humanidades en Brasil: un estudio a partir de los investigadores y las editoriales*. Doctoral Dissertation. Universidad Complutense de Madrid.
- Calasso, R. *The Art of the Publisher*. Penguin, 2015.
- Coufal, J. (ed.) (2017). *Global Ranking of the publishing industry 2017*. Rüdiger Wischenbart Content and Consulting.
- DORA. *San Francisco Declaration on Research Assessment* (2013) <http://crln.acrl.org/index.php/crlnews/article/view/9104/9996>
- ENRESSH (2015). *Memorandum of understanding for the implementation of the COST action European Network for Research Evaluation in the SSH CA15137*. <http://enressh.eu/wp-content/uploads/2016/10/CA15137-e.pdf>
- Giménez-Toledo, E. (2018). *Recognition of academic books in Spanish*. CSIC Research project 201810E125.
- Giménez-Toledo, E., Sivertsen, G., & Mañana-Rodríguez, J. (2017). Peer review as a delineation criterion in data sources for the assessment and measurement of scholarly book publishing in social sciences and humanities. In *16th International conference on scientometrics and informetrics*. Wuhan.
- Giménez-Toledo, Elea; Mañana-Rodríguez, Jorge & Sivertsen, Gunnar (2017). Scholarly book publishing: Its information sources for evaluation in the social sciences and humanities, *Research Evaluation*, 26, 2, 91–101, <https://doi.org/10.1093/reseval>
- Leydesdorff, L., & Felt, U. (2012). Edited volumes, monographs, and book chapters in the Book Citation Index (BKCI) and Science Citation Index (SCI, SoSCI, A&HCI). *arXiv preprint arXiv:1204.3717*.
- Mañana-Rodríguez, J. & Pölonen, J. (2018). Scholarly book publishers' ratings and lists in Finland and Spain: Comparison and assessment of the evaluative potential of merged lists. *ASLIB. Journal of Information Management*, 70, 6, pp. 643-659 <https://www.emeraldinsight.com/doi/full/10.1108/AJIM-05-2018-0111>
- Moed, H. F. (2006). *Citation analysis in research evaluation* (Vol. 9). Springer Science & Business Media.
- Oppenheim, C., & Summers, M. A. (2008). Citation counts and the Research Assessment Exercise, part VI: Unit of assessment 67 (music). *Information Research: An International Electronic Journal*, 13(2).
- Research Council of Norway (2017). Evaluation of the Humanities in Norway. Report. Oslo: *The Research Council of Norway*. <https://www.forskingsradet.no/en/Publications/1220788265688>

Research Excellence Framework 2014. *Main panel D criteria*

http://www.ref.ac.uk/2014/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2D.pdf

Torres-Salinas, D., Robinson-Garcia, N., Miguel Campanario, J., & Delgado Lopez-Cozar, E. (2014). Coverage, field specialisation and the impact of scientific publishers indexed in the Book Citation Index. *Online Information Review*, 38(1), 24-42.

Wilsdon, J.; Allen, L.; Belfiore, E.; Campbell, P.; Curry, S.; Hill, S.; Jones, R.; Kain, R.; Kerridge, S.; Thelwall, M.; Tinkler, J.; Viney, I.; Hill, J. Wouters, P. and Johnson, B. (2015). *The Metric Tide: The Independent Review of the Role of Metrics in Research Assessment and Management*. doi:10.13140/RG.2.1.4929.1363.

ⁱ IRAP working group: Gunnar Sivertsen, Vidar Røeggen, Janne Pölonen, Emanuel Kulczycki, Tim Engels, Raf Guns, Elea Giménez, Jorge Mañana, Alesia Zuccala & Kasper Bruun.