

Assessing Scientific Conferences through Knowledge Graphs

Simone Angioni¹, Angelo Salatino², Francesco Osborne², Aliaksandr Birukou³,
Diego Reforgiato Recupero¹, and Enrico Motta²

¹ Department of Mathematics and Computer Science, University of Cagliari (Italy)
{simone.angioni, diego.reforgiato}@unica.it

² Knowledge Media Institute, The Open University, Milton Keynes (UK)
{angelo.salatino, francesco.osborne, enrico.motta}@open.ac.uk

³ Springer-Verlag GmbH, Tiergartenstrasse 17, 69121 Heidelberg, Germany
aliaksandr.birukou@springer.com

Abstract. Springer Nature is the main publisher of scientific conferences in Computer Science and produces several well-known series of proceedings books, such as LNCS. The editorial team needs to take critical decisions about which conferences to publish as well as actively scan the horizon for identifying emerging ones. In this short paper, we present the Conference Dashboard, a new web application based on a large knowledge graph of scholarly data (1.3B triples) for assessing scientific conferences and informing editorial decisions.

Keywords: Scholarly Data · Knowledge Graphs · Scholarly Ontologies

1 Analysing Conferences with Semantic Technologies

Scientific conferences are of paramount importance in the field of Computer Science. They are the main spaces used by researchers for coordinating common initiatives, exchanging new ideas, disseminating research outputs, bootstrapping collaborations, and defining the paradigms in the field.

Springer Nature (SN) is the main publisher of scientific conferences in Computer Science and produces several well-known series of proceedings books, including LNCS, LNAI, IFIP-AICT, CCIS, and LNBIP, for a total of about 800 volumes per year. The SN editorial team needs to take critical decisions about which conferences to publish and in which series, as well as actively scan the horizon for identifying new emerging conferences. This requires an excellent understanding of how a conference ranks in a certain fields and where it appears to be heading in the following years. It is also important to determine which industrial sectors are mostly interested in a conference, to promote the resulting proceedings books. However, current tools only provide a limited support for assessing research conferences. Therefore, this complex and time-consuming task has traditionally been tackled manually by a few experienced senior publishing editors, resulting in bottlenecks, delays, and a costly editorial process.



Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In this short paper, we present the *Conference Dashboard*, a web application which combines semantic technologies and visual analytics for assisting editors in assessing conferences and informing editorial and marketing decisions. This application was developed in the context of the collaboration between Springer Nature and The Open University, which, since 2015, has produced an array of applications based on semantic technologies, leading to several improvement to the editorial pipeline and significant cost reductions [2]. The Conference Dashboard offers an interacting interface that allows the user to navigate several analytics based on the novel *Academia/Industry DynAmics Knowledge Graph* (AIDA) [1]. This knowledge graph includes 1.3B RDF triples and was produced by integrating information about 21M papers, 8M patents, and 5K conferences from Microsoft Academic Graph, Dimensions, DBpedia, and the Global Research Identifier Database. We also classified articles and conferences according to 14K research topics from the Computer Science Ontology (CSO, <https://cso.kmi.open.ac.uk/>) and 66 industrial sectors from the Industrial Sectors Ontology (INDUSO, <http://w3id.org/aida/downloads/induso.ttl>).

The Conference Dashboard introduces three main novel features that are not supported by current tools, but are important for addressing our user cases. First, it offers an interactive interface to compare and rank conferences within specific fields and time-frames. Second, it allows users to analyze the research areas of a conference across time, in order to assess how the focus of the conference is shifting. Finally, it offers several analytics for assessing the involvement of commercial organizations. For instance, it allows to visualize the ratio of articles from academia-industry collaborations or from specific industrial sectors (e.g., automotive, financial, energy). A demo version of the dashboard (covering the 2000-2020 period) is available at <http://w3id.org/aida/dashboard/>. The AIDA knowledge graph is stored using Virtuoso and can be queried via SPARQL Endpoint at <http://w3id.org/aida>.

2 Business Value

Our experiments showed that the Conference Dashboard has an excellent usability (87.5/100 according to the SUS questionnaire) and halves the time needed for analysing a conference. It also allows less experienced members of the team to undertake this task, saving the time of senior members. This solution can thus significantly improve the efficiency of the process and reduce the editorial costs.

We are now working on extending our analytical solutions to journals and countries, to support further queries from other SN departments. We are also producing an open version of the dashboard that will be made publicly available.

References

1. Angioni, S., Salatino, A., Osborne, F., Reforgiato Recupero, D., Motta, E.: Integrating knowledge graphs for analysing academia and industry dynamics. In: ADBIS, TPD and EDA 2020 Common Workshops and Doctoral Consortium. Cham (2020)
2. Salatino, A.A., Osborne, F., Birukou, A., Motta, E.: Improving editorial workflow and metadata quality at springer nature. In: The Semantic Web – ISWC 2019. pp. 507–525. Springer International Publishing, Cham (2019)