The Smart Book Recommender: An Ontology-Driven Application for Recommending Editorial Products

Thiviyan Thanapalasingam[1], Francesco Osborne[1], Aliaksandr Birukou[2], Enrico Motta[1]

[1] Knowledge Media Institute, The Open University, MK7 6AA, Milton Keynes, UK
{ thiviyan.thanapalasingam, francesco.osborne, enrico.motta }@open.ac.uk
[2]Springer-Verlag GmbH, Tiergartenstrasse 17, 69121 Heidelberg, Germany
aliaksandr.birukou@springer.com

Abstract. Promoting books and journals to the relevant research communities is an important task for major academic publishers. Unfortunately, identifying which are the best editorial products to market at a certain academic venue is a time-consuming and error-prone process. Here we present the Smart Book Recommender (SBR), an ontology-based recommender that supports the Springer Nature editorial team in selecting the editorial products to market at specific venues. SBR provides an interactive visualisation for analysing the topics characterizing conference series and books. It builds on a dataset of 27K books, journals, and conference proceedings annotated with topics from the Computer Science Ontology, a large-scale ontology of research areas. A user study showed that SBR is able to produce useful recommendations for both editors and researchers.

Keywords: Recommender System, Scholarly Data, Scholarly Ontologies, Data Mining, Conference Proceedings, Metadata.

1 Introduction

Major academic publishers, such as Springer Nature, often attend academic conferences and present their most recent books and journals relevant to the conference topics. The purpose of such stands is twofold: the conference participants can browse the most recent research and the authors get additional exposure for their work.

Typically, the selection of products is performed by publishing editors who possess domain specific knowledge and years of experience in publishing books and journals on the conference topics. The process usually requires browsing through a vast catalogue of editorial products on multiple platforms [1,2] . It is a time-consuming and error-prone practice since it is easy to miss some important publications. Furthermore, the constant emergence of new research areas over time poses the challenge of keeping up to date with the research dynamics.

The ongoing collaboration between Springer Nature (SN) and the Knowledge Media Institute (KMi) of The Open University has given rise to a number of semantic

[1] https://link.springer.com
[2] https://www.springer.com

technologies to support publishing activities [1]. These solutions include the Smart Topic Miner[3], a tool for classifying proceedings according to a large-scale Computer Science Ontology, and the Smart Topic API, a web service for assigning semantic topics to research papers [2].

The most recent product of this collaboration is the Smart Book Recommender (SBR) [3], an ontology-based recommender system designed for suggesting books, journals, and proceedings for specific Computer Science (CS) venues. This demo paper is complementary to the one accepted in the ISWC 2018 In-Use track [3] and focuses on the main functionalities of the system. A demo of SBR is available at http://rexplore.kmi.open.ac.uk/SBR_demo/.

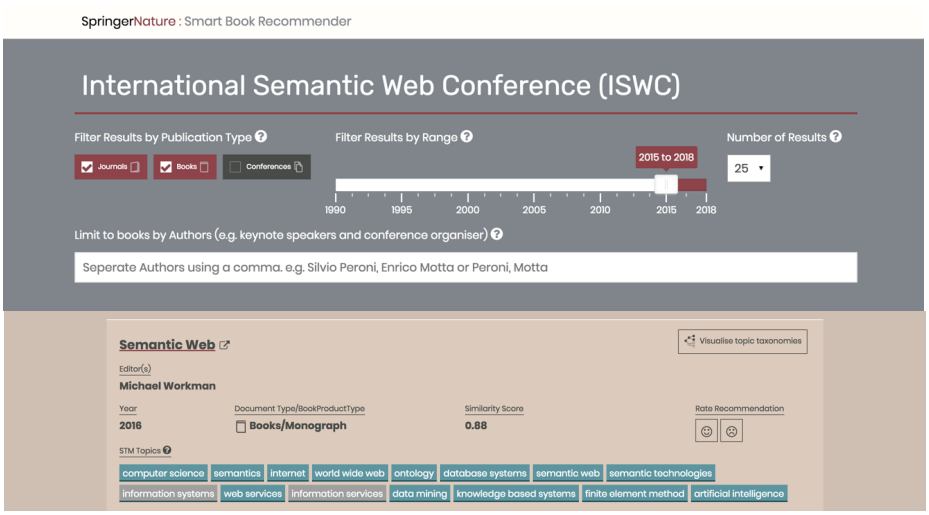

Figure 1: SBR interface showing the search tools (top) and one of the recommended items for the selected conference (bottom)

2 Smart Book Recommender

The Smart Book Recommender (SBR) is a web application that takes as input a conference series (e.g., "International Semantic Web Conference") and returns the books, journals, and conference proceedings that are characterized by a set of similar research topics (e.g., the "Handbook of Semantic Web Technologies"). The frontend was implemented using HTML5 and JavaScript, while the backend was developed in Python and queries a MariaDB database.

SBR builds on a dataset of 27K books, journals, and conference proceedings published by Springer Nature that we annotated with topics from the Computer Science Ontology (CSO)[4] using the Smart Topic API [2]. CSO is a large-scale ontology of research topics which includes about 26K research topics and 226K relationships [4]. It was automatically generated using the Klink-2 algorithm [5] on a dataset of 16 million publications, mainly in the field of Computer Science. SBR produces

[3] A demo of STM is available at http://rexplore.kmi.open.ac.uk/STM_demo/
[4] https://w3id.org/cso/

recommendations for a specific conference series by suggesting the items associated with a similar distribution of semantic topics. It does so by computing the cosine similarity between the vectors of topics associated with the conference proceedings and all the other editorial products. For a more comprehensive discussion of the SBR implementation, please refer to [3].

2.1 User Interface

Figure 1 shows the interface of the SBR web application. The users can select a conference series by using either the full conference name or its abbreviated form (e.g., "International Semantic Web Conference" or "ISWC"). After the conference series has been selected, a list of recommended editorial products is loaded via an AJAX request. The results are displayed as cards and sorted in descending order of similarity. Each card summarises an editorial product using its title, publication year, fifteen most significant topics, and the overall similarity score with the input conference. SBR also provides hyperlinks to the corresponding Springer Link catalogue pages, where more information can be found.

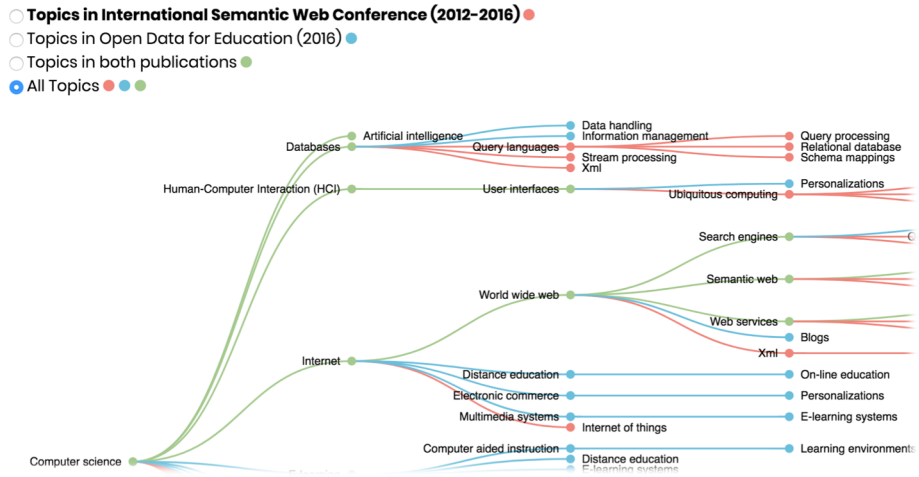

Figure 2: Subset of an interactive graph showing the similarities and dissimilarities in the topic coverage of ISWC (2012 – 2016 proceedings) and Open Data for Education (book).

SBR allows users to filter recommendations according to i) publication types (books, journals, or proceedings), ii) publication years, and iii) authors and editors of the books. The publishing editors can also inspect the topic taxonomy of the suggested products, rate the results, and export the list of recommendation, either as a JSON or CSV file. This document is reviewed and sent to the Exhibit Department which takes care of dispatching the publications to the conference.

Figure 2 shows the interactive visualisation offered by SBR for comparing the topics of the input conference series with the ones of a recommended item. This interface intends to provide an intuitive explanation for the similarity score by representing items as taxonomies of topics. It displays in red and blue the topics that appear only in one of the two items and in green the shared ones. The users can also change the granularity

of the representation by opting to show only topics that appear in a minimum number of chapters.

2.2 User Study

We conducted a user study for evaluating SBR involving 7 SN publishing editors and 7 researchers[5]. These users rated as relevant 76.1% of the first ten recommended items (i.e., precision@10). They also reported that SBR was "intuitive" and "easy to pick up" and that its interface was "simple" and "well-organised". In particular, editors and the researchers yielded respectively an average SUS score[6] of 77.1±15.2 and 80.3±11.3, which converts in a percentile rank of about 75%.

The users also suggested some additional functionalities that we are working to implement, such as the ability to change the topic-based representation of the input conference by adding and removing significant research areas, and the ability to take download statistics into consideration.

3 Conclusion

In this demo paper we presented our work on the Smart Book Recommender, an ontology-driven recommendation system designed for assisting the Springer Nature publishing editors in selecting books to be marketed at conferences. The semantic representation of research conferences used by SBR was considered very helpful both by editors and researchers, since it allows the users to understand why a certain item was recommended and to compare different products.

As future work, we plan to cover other scientific domains, such as Engineering and Life Science.

References

1. Osborne, F., Salatino, A., Birukou, A., Thanapalasingam, T., Motta, E.: Supporting Springer Nature Editors by means of Semantic Technologies. In International Semantic Web Conference 2017 Industry Track. (2017)
2. Osborne, F., Salatino, A., Birukou, A., Motta, E.: Automatic Classification of Springer Nature Proceedings with Smart Topic Miner. In International Semantic Web Conference 2016 Application Track. (2016)
3. Thanapalasingam, T., Osborne, F., Birukou, A., Motta, E.: Ontology-Based Recommendation of Editorial Products. In International Semantic Web Conference 2018 In-use Track. (2018). Pre-print available on https://goo.gl/osLBQz.
4. Salatino, A.A., Thanapalasingam, T., Mannocci, A., Osborne, F. and Motta, E.: The Computer Science Ontology: A Large-Scale Taxonomy of Research Areas. In International Semantic Web Conference 2018 Resource Track. (2018). Pre-print available on https://goo.gl/oZ1eT4.
5. Osborne, F., Motta, E.: Klink-2: integrating multiple web sources to generate semantic topic networks. The Semantic Web-ISWC 2015, pp. 408-424. Springer. (2015)

[5] The study is reported in [3]. The relevant data are available at https://figshare.com/articles/Smart_Book_Recommender_Evaluation_Data/6087032/2

[6] https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html