Keeping scientific materials alive and reusable by supporting collaboration

Aleix Garrido[1], Jose Manuel Gomez-Perez[1], and Boris Villazon-Terrazas[1]

iSOCO, Intelligent Software Components S.A., Av. del Partenon, 16-18, 1-7, 28042, Madrid, Spain,
{agarrido,jmgomez,bvillazon}@isoco.com

Abstract. Data-intensive scientific knowledge faces many issues such as the decrease of quality of materials and the difficulties on finding relevant information. Those aspects need to be addressed in order to enhance the future of scientific development. This article presents a set of tools focused on reusability of scientific information based on two different aspects: collaboration and search. We present a visual metaphor called Collaboration Spheres which allow search through exploration of scientific social networks based on a combination of people and information objects. This search paradigm is built around the use of customizable contexts that act as intuitive examples conducted by the users. This tool eases collaboration, expert finding and access to relevant information.

1 Introduction

Introduction related to data quality

The application aims to help journal editors, conference and workshop chairs on identifying suitable reviewers for a particular paper, by providing a nice visualization of related authors and papers. To this end, the pilot relies on an innovative social network visualization, namely Collaboration Spheres.

In a nutshell the Collaboration Spheres provide a mechanism to improve, share and reuse research papers and author experiences information based on the exploitation of semantic descriptions, relations and similarities between research papers and authors.

2 Reviewer finder

The Reviewer finder is illustrated by a web application that covers the process of finding a good reviewer for a specific article or publication. The process is driven by actions performed by the user while creating different contexts towards his/her final objective. Each action allows the modification of the search parameters in an intuitive way and offers summaries of information that may help on understanding the available resources.

2.1 Interface Description

The user interface has been designed in order to keep a minimalist layout that makes user experience smooth and simple. The screen is easily divided in four different parts:

Spheres The main part displays a set of concentric circles. These circles have two different functionalities. On the one hand, the center of the circles contains the main article for which we are looking for reviewers. Its adjacent circle is the place where search queries are defined through dropping elements from the lists of items(see Lists) for creating contexts. On the other hand the two external circles offer the results of the reviewer finder. The results are small icons that use different colors depending on the relative quality of the recommendation (green for the best results, yellow for medium, and red for weak recommendations). The icon may have a warning sign associated if a conflict of interest (such as previous co-authors or same organization) is detected for the recommended reviewer.

Lists A total of six lists are loaded on the right side of the screen. They are divided in two columns: authors and articles. Each column has three lists. The lists on the first column are: list of authors of the main article, list of previous co-authors of the main authors and a list of authors which have worked on related topics to the article to be reviewed. The lists on the second column are: articles of the main authors, articles of previous co-authors and topic related articles. All the items of the lists are draggable and can be dropped on the Spheres in order to customize the context.

Tag Cloud A tag cloud that gathers the most representative topics of the contexts that are created by the user is placed beneath the Spheres. The tag cloud allows the understanding of the search and results at a glance.

Information box Each element of the lists and the spheres is clickable. If they are clicked, a box of information placed on the right-bottom corner of the screen loads a brief summary of information, together with a link to its uri and a list of relevant topics for the current context.

2.2 Basic Architecture

The architecture of the system consists of three modules. Each module provides the needed information to the others creating a complete system:

Data Module The developed platform uses data provided by the APA VIVO, which is an RDF dataset holding authors, publications and additional data from the psychology domain. The base dataset is enriched with topics from the titles and abstracts that have been extracted using the TextRazor[1] API and KT (Knowledge Tagger) working together with a Thesaurus. Post-processing calculations are used in order to add a weight value for each topic

[1] http://www.textrazor.com/ TextRazor provides topic extraction and additional Wikipedia URIs for different categories.

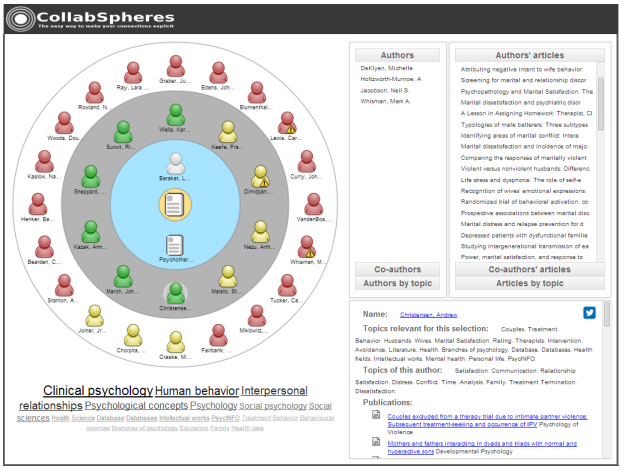

Fig. 1. User interface screenshot

in every article and to assess an expertise value to the authors. The data is stored in a Virtuoso Triplestore accessible via HTTP queries.

Web Services The communications between the Data Module and the Front End are performed by the Web Services. It is a web project that triggers different SPARQL queries by using JENA under user request. The results are processed and returned via REST-JSON to the Front End.

Front End A web application with a browsable interface that is directly available for the users. The interface reacts to user actions and throws different Ajax queries to the REST Web Services. It has been developed in HTML5, using also javascript, jQuery and CSS.

2.3 Scenario

In this scenario the editor in chief of the Journal of Personality and Social Psychology[2] is trying to select reviewers for a submission they have just received. The first step will be to open the collaboration spheres for that Article:

- The selected Article is called "Attachment change processes in the early years of marriage" [3].
- Then we can launch the Collaboration Spheres for the Article [4].

As we can see, at the tag cloud shows the most relevant topics for the Article placed in the center of the Collaboration Spheres and we get a preliminary set of recommended reviewers for it.

[2] http://www.apa.org/pubs/journals/psp/
[3] http://vivo.apa.org/individual/n61963
[4] http://apaproject.isoco.com/index.html?id=http://vivo.apa.org/individual/n61963

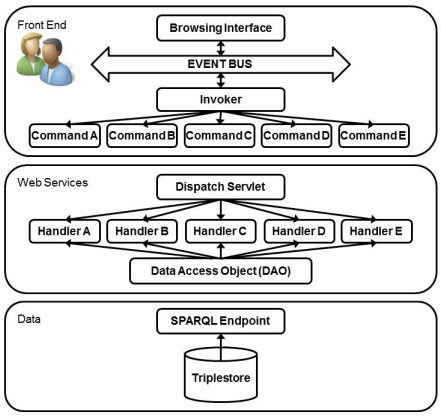

Fig. 2. Basic architecture

Since we want to define a bit better the topics for our context, i.e. the special issue, and, in order to get better recommendations, we will add an article of our interest (can be found at the Articles by Topic section), which is very related to the topics of the special issue. We select the following article in the Articles by topic panel on the right and drag and drop it into the blue circle of the context of interest:

- The added article is "Adult attachment and the transition to parenthood" [5].

Now the editor in chief will include some of the members of the Editorial Board (Interpersonal Relations and Group Processes Section[6]), who will serve as exemplars of the knowledge required to properly evaluate the submissions which are received. Suitable reviewers will therefore be knowledgeable of (part of) those topics. This way we increase the alignment between the related topics in the expertise of our editors and the recommended experts for reviewing the article. In this case, the editor in chief adds:

- Consulting Editors:
  - Bolger, Niall[7]
  - Algoe, Sara B.[8]
- Associate Editors:
  - Finkel, Eli J.[9]
  - Gable, Shelly L.[10]

[5] http://vivo.apa.org/individual/n5724
[6] http://www.apa.org/pubs/journals/psp/edboard-irgp.aspx
[7] https://vivo.apa.org/display/n41433
[8] https://vivo.apa.org/display/n59044
[9] https://vivo.apa.org/display/n36079
[10] https://vivo.apa.org/display/n37385

We can see at every addition how the main topics change in the tag cloud and how the recommendations are adjusted to the context. When the editor in chief adds the last of the associate editors, the relevance of the different related topics seems to be uniformly distributed. The editor in chief then may want to reconsider the composition of the board and replace some of the editors (e.g. Gable with another expert with knowledge about other topics more, which the editor in chief finds especially relevant for the special issue). In this case, the editor in chief selects an author who brings in expertise related to topics like alcohol abuse:

– Zywiak, William H.[11]

In addition to the summary that we provide for each author and article, the system can also launch a Twitter search that covers his/her/its intersection with the most relevant topics of the context of interest.

3 Dataset description

Vocabularies

1 VIVO Ontology	http://vivoweb.org/ontology/core
2 BIBO Ontology	http://purl.org/ontology/bibo
3 FAO Geopolotical	http://aims.fao.org/aos/geopolitical.owl
4 Vitro Application Ontology	http://vitro.mannlib.cornell.edu/ns/vitro/0.7
5 FOAF	http://xmlns.com/foaf/0.1
6 SKOS	http://www.w3.org/2004/02/skos/core
7 iSOCO vocab	http://vocab.isoco.net
8 Event Ontology	http://purl.org/NET/c4dm/event.owl

Some stats

Number of triples	2605589
Number of links	123606
Number of authors	12090
Number of papers	17386
Number of topics	465706

4 Conclusions and Future Work

Acknowledgments

This work was supported by SmartContent project.

[11] https://vivo.apa.org/display/n8844