# Collaboration Spheres: Finding Reviewers for Scholarly Publications

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.** Science is an incremental process. New breakthroughs are built on top of previous results which will themselves be the basis for future developments once validated and accepted for publication. However, identifying suitable reviewers for a particular article is not an easy task. Not only do journal editors need to have a thorough, personal knowledge of their journals and the members of the community who can be suitable reviewers but also they must have an eye for spotting potential conflicts of interest that may raise biased reviews. Thus, means are required that help editors in order to identify suitable reviewers for an article. In this paper we present an expert finding system based on the application of the Collaboration Spheres search-by-example visual interface, to the case of the American Psychologic Association, a publishing body with over 70 journals and 134,000 members. We show how we enriched APA's existing RDF data model and produced an extended dataset with additional relations and information. Based on this dataset we describe a service layer which feeds the Collaboration Spheres and enables the expert finding capabilities.

## 1 Introduction and motivation

In this article we present a system description for 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.

The presented 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 Expert finder

The Collaboration Spheres are illustrated by the APA prototype that we have developed. It consists on a web application that covers the process of finding a good reviewer for a specific article or publication. The whole process is driven by the actions performed by the user while creating different contexts towards his/her final objective which is finding an expert for a specific article. A context is a group of articles and authors surrounding the main article that are used as search parameters. Each action allows the modification of the context of the Collaboration Spheres 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. This simplicity does not renounce to provide the content needed by te users in the aim of getting the desired results. The main part of the screen displays a set of concentric circles that serve both as a playground to customize the search and as a front of the most relevant results of the search. These circles are the Collaboration Spheres metaphor, where the importance of the elements placed in the circles decreases as we move away from the center. This metaphor is supported not only by proximity to the center, but also by colors and iconography as we describe below. The center of the circles contains the article for which we are looking for reviewers. The immediate adjacent circle is the place where the user can define different contexts (each context created is equivalent to a search) by adding and removing authors and articles. The two external circles collect the most relevant reviewers that the system has found in order to tackle the main article together with the context that surrounds it. Other aspect that is covered by the Collaboration Spheres is the usage of a color code and warning icons when showing the results on the circles. The colors follow the traffic-light metaphor, where green represents the best reviewers, yellow is used for other good reviewers and red is used for the less (but still) recommended reviewers. As we evaluate possible conflicts between the reviewers and the article to be reviewed, we are able to warn the users by adding a warning icon for those recommendations where conflict has been detected (author, previous co-author or same organization are examples of conflicts in our scenario). On the right of the screen there are two different columns that gather three distinct lists each. The first column focuses on authors at different levels:

1. The first level shows the main authors of the article.
2. The second level shows other authors that had collaborated with the main authors in previous publications.
3. The third level presents a set of relevant authors whose work share one or more topics with the main article.

A parallel approach has been taken for the second column which focuses on articles at three different levels:

1. The first level shows other articles authored by the main authors.
2. The second level shows a set of articles authored by previous co-authors.
3. The third level presents a set of relevant articles that have related topics with the main article.

The preselection of elements enhances the user experience because the user has to deal with a set of related items instead of a huge amount of unlinked information. Every element from the lists is draggable and can be dropped at the circles. The drag-and-drop action allows the user to create their customized context in the Collaboration Spheres. Every time that a new element is added to the context, a tag cloud that is placed under the circles gets updated. This tag cloud represents the key topics for the created contexts and provides an informative abstraction of the search at a simple glance. Some of the elements at the tag cloud are a link to their corresponding Wikipedia and APA Thesaurus URIs. Apart from that, it is worth to mention that every element is clickable in order to get a summary of its contents, together with the link to the resource itself in the VIVO platform enhancing the users exploration. A twitter search for the most relevant topics of every element is provided with the intention of adding a nice feature of pointing topic appearance in social networks.

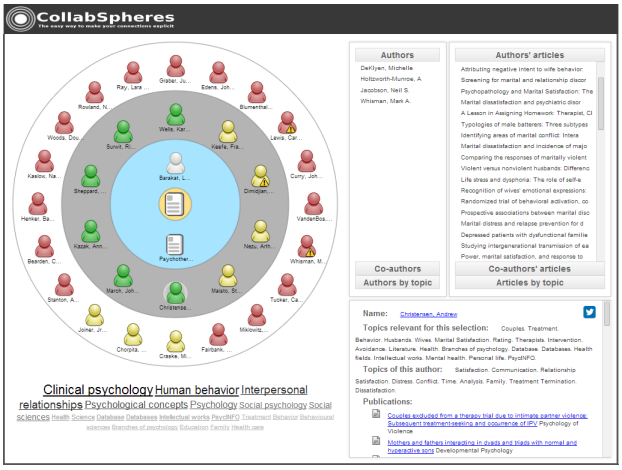


**Fig. 1.** User interface screenshot

### 2.2 Architecture

The architecture and implementation of the system can be easily divided into three independent modules. Each module provides the needed information to the others creating a complete system.

**Data Module:** The first module is the data module. The base data used in our scenario is provided by the VIVO APA[1] platform. It consists on an RDF dataset that contains authors and publications that belong to the Psychology domain, together with additional information for each element and the relations between them. One of our tasks related to this dataset has been an information enrichment. Topics form the abstracts and titles have been extracted through NLP techniques and tools for each publication. One of the tools used to perform this task has been TextRazor[2] through its API. The second tool that has been used is KT (Knowledge Tagger) working together with a Thesaurus for the psychology domain provided by the APA. Post-processing calculations are used in order to add a weight value for each topic in every article. These topic weights together with the knowledge of the authors that provides the dataset allows the assessment of expertise values for each author. These weighed values are the a key part of data when it comes to recommending experts. The dataset together with the new relations and the added values are stored in a Virtuoso Triplestore. The new dataset offers an explicit representation which allows not only a better understanding of the data but also an improvement when building queries and when retrieving data. The Triplestore is accessible via HTTP SPARQL queries that are posted by the next module:the Web Services.

**Web Services:** The Web Services is in charge of communicating the data form the Data Module to the Front End. It is a web project that triggers different SPARQL queries by using JENA under user request. The results of the queries are processed and returned via REST-JSON web services that respond to the Front End. Some filtering of the data and improvements are done in the Java project, but they are kept as minimum as possible. In this way we try to reduce the impact of this module and use it as a communication channel between the other modules. The power of the Web Services resides on the complex SPARQL queries and the ability to handle user requests.

**Front End** Finally, the Front End Module offers a web application with a browsable interface that is directly available for the users. It has been developed in HTML5, using also javascript, jQuery and CSS. The interface reacts to user actions and throws different Ajax queries to the REST Web Services. Once the JSON response is obtained it is parsed in order to show the results to the users and update the interface as expected. The final result offers a smooth experience that retrieves data through recommendations under user demand.

[1] https://vivo.apa.org/

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

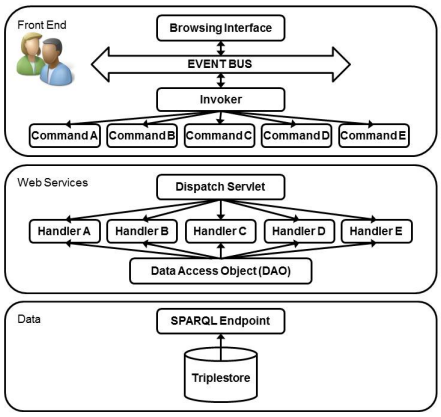


**Fig. 2.** Basic architecture

## 3 Dataset Information

In this section we provide information related with the vocabularies that are used in the dataset together with additional stats about the amount of information managed by the application.

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.