Personal Research Assistants
for Young Researchers

Bahar Sateli and René Witte

Semantic Software Lab
Department of Computer Science and Software Engineering
Concordia University, Montréal, QC, Canada
{sateli,witte@semanticsoftware.info}

We present the concept of *Personal Research Agents* as virtual entities capable of helping researchers who have to deal with the overwhelming amount of scientific literature to carry out their daily research tasks. These personal agents offer a variety of services to user groups, including graduate students, junior scientists, and learners, assisting them in knowledge-intensive tasks, such as writing scholarly documents.

While existing spell and grammar checking tools can help students and young authors improve the readability of their manuscripts by spotting syntactical defects like typographical errors or passive voice, to the best of our knowledge no tools exists that can automatically analyze the *rhetorical* structure of a document and help the students in formulating scientific argumentations. Rhetorical analysis of scientific texts requires an automatic *understanding* of the article's meaning. We argued in [3] that this can be achieved using a confluence of techniques from the Natural Language Processing and Semantic Web domains. The idea is to transform scholarly documents (e.g., theses, journal articles) from textual documents in isolation to a semantically-rich, queryable and machine-readable format [2]. We realized our ideas in form of virtual research agents that can extract the structural and rhetorical elements of scholarly documents using modern text mining techniques. Based entirely on open source software, we demonstrate how the resulting knowledge base, analogous to the agent's *brain*, can be exploited to give students feedback, for instance, when no rhetorical elements (e.g., claims or contributions) are found within their manuscripts. In addition, we present how the agent's services can be *personalized* for users through automatic profiling of their background knowledge, interests and tasks using their reading or writing history [1]. Additional information, as well as our open source workflow [4] is available on http://www.semanticsoftware.info/from-papers-to-triples.

References

1. Bahar Sateli, Felicitas Löffler, Birgitta König-Ries, and René Witte. Semantic User Profiles: Learning Scholars' Competences by Analyzing their Publications. In *Semantics, Analytics, Visualisation: Enhancing Scholarly Data (SAVE-SD 2016)*. Springer, 2016.
2. Bahar Sateli and René Witte. Semantic representation of scientific literature: bringing claims, contributions and named entities onto the Linked Open Data cloud. *PeerJ Computer Science*, 1(e37), 2015.
3. Bahar Sateli and René Witte. What's in this paper? Combining Rhetorical Entities with Linked Open Data for Semantic Literature Querying. In *Semantics, Analytics, Visualisation: Enhancing Scholarly Data (SAVE-SD 2015)*, page 1023–1028, Florence, Italy, 2015. ACM. Received "Best Paper Award" at SAVE-SD 2015.
4. Bahar Sateli and René Witte. From papers to triples: An open source workflow for semantic publishing experiments. In *Semantics, Analytics, Visualisation: Enhancing Scholarly Data (SAVE-SD 2016)*, Montréal, QC, Canada, 2016. Springer.