# SAVE-SD 2016 Submission: Supplementary Material

- Semantic Publishing
- Semantic Publishing
- Zeeva
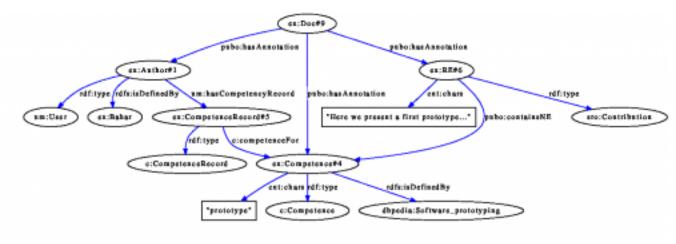
toc_collapse=0; Table of Contents

# 1. Overview

This page provides supplementary material for our submission to the SAVE-SD 2016 workshop on *Semantics, Analytics, Visualisation: Enhancing Scholarly Data*.

# 2. Supplements



**Semantic User Profiles Example**

The data sets used in the experiments (*Evaluation* and *Application* sections), were collected from a total of 47 scientific articles. They cover the 10 different computer scientists from Concordia University and the University of Jena (including the co-authors) that participated in the evaluation.

## 2.1. Source Code

*Coming soon...!*

## 2.2. Knowledge Base

We processed the publications dataset with our Semantic Profiling pipeline and populated a TDB-backed knowledge base. You can download the complete (zip-compressed) knowledge base in N-Quads format (840,476 triples).

Note that the authors' names and paper identifiers have been anonymized.

# 3. Evaluation

## 3.1. User Profiles

We processed the full-text of the corpora described above and populated a knowledge base with the researchers' profiles. We also developed a Java command-line tool that queries the knowledge base and generates LaTeX documents to provide for a human-readable format of the researchers' profiles that lists their top-50 competence topics sorted by the number of occurrence in the user's publications. Subsequently, we asked the researchers to review their profiles across two dimensions: *(i)* the relevance of the extracted competences, and *(ii)* their level of expertise for each extracted competence. You can find the anonymized scans of the user profiles in anonymized-profiles.zip.

## 3.2. Evaluation Results

To evaluate the effectiveness of our system, we utilize one of the most popular ranked retrieval evaluation methods, namely the Mean Average Precision (MAP). MAP indicates how precise an algorithm or system ranks its top-N results, assuming that the entries listed on top are more relevant for the information seeker than the lower ranked results. Please refer to Table 2 in the paper for an overview of the evaluation results. You can download and check out the evaluation results of the full-text profiles and re-only-profiles spreadsheets.

# 4. Application

## 4.1. Knowledge Base Queries

**Finding all competences of a user:** By querying the populated knowledge base with the researchers' profiles, we can find all topics that a user is competent in:

```
1.  PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

2.  PREFIX um: <http://intelleo.eu/ontologies/user-model/ns/>

3.  PREFIX c: <http://www.intelleo.eu/ontologies/competences/ns/>

4.  PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

5.  PREFIX pubo: <http://lod.semanticsoftware.info/pubo/pubo#>

6.  PREFIX sro: <http://salt.semanticauthoring.org/ontologies/sro#>

7.
```

```
8.  SELECT DISTINCT ?uri (COUNT(?uri) AS ?count) WHERE {

9.    ?creator rdf:type um:User .

10.    ?creator rdfs:isDefinedBy <http://semanticsoftware.info/lodexporter/creator/R1>
       .

11.    ?creator um:hasCompetencyRecord ?competenceRecord .

12.    ?competenceRecord c:competenceFor ?competence .

13.    ?competence rdfs:isDefinedBy ?uri .

14.    ?rhetoricalEntity rdf:type sro:RhetoricalElement .

15.    ?rhetoricalEntity pubo:containsNE ?competence .

16.  } GROUP BY ?uri ORDER BY DESC(?count)
```

**Ranking papers based on a user's competences:** We showed in [1] that retrieving papers based on their LOD entities is more effective than conventional keyword-based methods. However, the results were not presented in order of their *interestingness* for the end-user. Here, we integrate the semantic user profiles to rank the results, based on the common topics in both the papers and a given user's profile:

```
1.  PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

2.  PREFIX um: <http://intelleo.eu/ontologies/user-model/ns/>

3.  PREFIX c: <http://www.intelleo.eu/ontologies/competences/ns/>

4.  PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

5.  PREFIX pubo: <http://lod.semanticsoftware.info/pubo/pubo#>

6.  PREFIX sro: <http://salt.semanticauthoring.org/ontologies/sro#>

7.

8.  SELECT (COUNT(DISTINCT ?uri) as ?rank) WHERE {

9.    <http://example.com/example_paper.xml> pubo:hasAnnotation ?topic .

10.    ?topic rdf:type pubo:LinkedNamedEntity .

11.    ?topic rdfs:isDefinedBy ?uri .

12.    FILTER EXISTS {

13.      ?creator rdfs:isDefinedBy
       <http://semanticsoftware.info/lodexporter/creator/R8> .

14.      ?creator um:hasCompetencyRecord ?competenceRecord .
```

```
15.        ?competenceRecord c:competenceFor ?competence .

16.        ?competence rdfs:isDefinedBy ?uri .

17.     }

18. }
```

**Finding users with related competences:** Given the semantic user profiles and a topic in form of an LOD URI, we can find all users in the knowledge base that have related competences. By virtue of traversing the LOD cloud, we can find topic URIs that are (semantically) related to a given competence topic and match against users' profiles to find competent authors:

```
 1.  PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

 2.  PREFIX um: <http://intelleo.eu/ontologies/user-model/ns/>

 3.  PREFIX c: <http://www.intelleo.eu/ontologies/competences/ns/>

 4.  PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

 5.  PREFIX pubo: <http://lod.semanticsoftware.info/pubo/pubo#>

 6.  PREFIX sro: <http://salt.semanticauthoring.org/ontologies/sro#>

 7.  PREFIX dcterms: <http://purl.org/dc/terms/>

 8.  PREFIX dbpedia: <http://dbpedia.org/resource/>

 9.

10. SELECT ?author_uri WHERE {

11.    SERVICE <http://dbpedia.org/sparql> {

12.       dbpedia:Ontology_(information_science) dcterms:subject ?category .

13.       ?subject  dcterms:subject ?category .

14.    }

15.    ?author rdf:type um:User .

16.    ?creator rdfs:isDefinedBy ?author_uri .

17.    ?creator um:hasCompetencyRecord ?competenceRecord.

18.    ?competenceRecord c:competenceFor ?competence.

19.    ?competence rdfs:isDefinedBy ?subject.

20.    ?rhetoricalEntity pubo:containsNE ?competence.
```

```
21.    ?rhetoricalEntity rdf:type sro:RhetoricalElement.

22.  }
```

## References

1. Sumner, T. (Eds.),  Sateli, B., and R. Witte,  "Semantic representation of scientific literature: bringing claims, contributions and named entities onto the Linked Open Data cloud",  *PeerJ Computer Science*, vol. 1, no. e37 PeerJ, 12/2015.

| Attachment | Size |
|---|---|
| anon-profiles.zip | 6.63 MB |
| profiles_evaluation.ods | 48.88 KB |
| profiles_evaluation_re.ods | 57.37 KB |

**Source URL (retrieved on *2025-12-22 09:44*):** https://www.semanticsoftware.info/save-sd2016