**Social Structured and Semantic Search**

Raphaël Bonaque, Bogdan Cautis, François Goasdoué, Ioana Manolescu

---

**General Framework**

A model, $S3$, that we introduce, to include:
- Structured documents, such as JSON and XML
- with Semantic annotations, in RDF
- within a Social network with usual interactions: relations between user, posting, commenting and tagging

Several goals to achieve in this framework:
- Top-k search on this model: find the k best documents for a given user query
- Efficient and practical implementation of the search
- Proof of correctness

---

**State of the Art**

- Top-k in social networks, unstructured document, no semantics:
- Semantics, structured or unstructured documents, without social aspects:
- Structured documents, without semantic or social aspects:

---

**Contributions**

- Formalised data and query model
- Implemented top-k search algorithm
- Proof of the correctness of the top-k search algorithm for a wide range of general scores
- Evaluation on datasets using real world data: Twitter (a subset of 2.8M nodes), Vodkaster, and Yelp!

---

**Our Top-k Approach**

We propose a top-k algorithm working on customisable scores: the score of a document for a query must depends on the distance from the user making the query to the sources of relevant keywords for the query.
- Relevant keywords are derived from the query keywords by inference in RDF
- Keyword sources are users posting documents or tags containing them
- The distance between users depends on every social path between them: paths following social interactions and going through ancestor relations in documents

---

**Results**

Our implementation was tested on datasets created from several social networks and knowledge bases with a score function generalising standard social and structural scores.

Compared to a state of the art algorithm working on social data, we:
- capture ~ 34% more results from the social interactions
- capture ~ 15% more results from the semantic inheritance

---

**Perspectives**

- Personalizing query results based on per user ontology developed from the social links
- Developing a comprehensive way to query heterogeneous data models, not only social and structured, through extended RDF queries
- Accepted for publication at EDBT 2016 as "Structured, Social and Semantic Search"