RESEARCH PROJECT: Bridging Natural Language Processing and the Web of Data for Multimedia Question Answering

Research Axis: Data Centric and Networking

JOINT PROPONENTS: EURECOM (Multimedia Communications Department), INRIA (Wimmics Research Team)

What is now called the web of data[1] can been seen as the first wave of the deployment of the Semantic Web (SW). It aims at transforming the access to information by adding machine-readable linked data and the semantics of their schema to the already visible content, in order to facilitate automated processing and integration of the vast amount of information available on the web. To formalize the description of the data objects and relations among them, the web of data relies on the Resource Description Framework (RDF) data model, whose independence from the data instantiation makes it suitable to describe the heterogeneity of the resources present in the web, being textual data or multimedia content (e.g. pictures, graphics, audio, speech, video). Initiatives like the Linked Open Data[2] support the idea that data should be freely available to everyone to use and republish without restrictions, boosting *i)* the creation of new data sets, automatically connecting pieces of data, information, and knowledge using standard models for web data interchange, and *ii)* the possibility to query interconnected data, distributed by nature, constantly evolving and expanding with the web itself (as shown by the Linked Data cloud[3]). This allows interoperability, re-usability and potentially unforeseen applications of opened data. To allow non-expert users to benefit from this new framework, the use of NLP techniques is the way to go to bridge semi-structured and formal resources, both for extracting knowledge from textual resources and for satisfying information needs using natural language queries.

The research project we propose here considers an open domain Question Answering (QA) scenario on the web of data, and focuses in particular on multimedia content. Given for instance a natural language question like "*Which aircraft were used during World War II?*", the goal is to mine the web to be able to provide the user with the pieces of information required, combining both textual information with other media content. To achieve this goal, it is necessary to:

- *i)* interpret the user question in order to understand its target (i.e. the piece of information required, in this case the models and illustrations of the aircrafts fighting in World War II);
- *ii)* mine the web of data to detect and collect such information, combining the results obtained from heterogeneous sources (e.g. encyclopedia knowledge on DBpedia, aircraft pictures on MIRFLICKR[4], videos of World War II on YouTube), and that can change dynamically over the time (e.g. the list of the museum exhibitions in which such planes are temporary exhibited);
- *iii)* provide to the user the list of the war aircrafts together with their images, ranked for instance in a chronological order, or geo-localized on a map according to their manufacturing country.

To reach our goal, the following research questions should therefore be addressed:

1. how to map natural language expressions (e.g. user queries) with concepts and relations in a structured knowledge base of multimedia content?
2. how to handle the heterogeneity and dynamics of linked data in a QA scenario?
3. how to present the results to the users in the form of natural language answers and/or semantically enriched multimedia presentations?

[1] `linkeddatabook.com/`
[2] `http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData`
[3] `http://linkeddata.org/`
[4] `http://press.liacs.nl/mirflickr/`

Differently from search engines, the goal of QA is to return precise answers to users' natural language questions, extracting information from both documentary text and advanced media content. The research area on QA, and especially on scaling up QA to linked data, is a wide and emergent research area that still needs an in-depth study to benefit from the rich linked data resources available on the Web [2]. Up to now, QA research has largely focused on text, mainly targeting factual and list questions. The goal of the proposed project is instead to exploit the structured data and metadata describing multimedia content on the LOD to: *i)* provide a richer and more complete answer to the user, combining textual information with other media content; and *ii)* address types of questions that have been less investigated in the literature (e.g. *how* to and *why* questions), and for which multimedia answers seem to be more intuitive and appropriate [3]. Moreover, a natural language answer should be generated and presented to the user in a narrative form for an easy consumption, supported by multimedia elements [1].

The project is jointly proposed by the Multimedia Communications Department at EURECOM, and the Wimmics team at INRIA. Given the multidisciplinary research topic, the ideal candidate would have a strong background in both Natural Language Processing and Semantic Web research fields and a strong interest in multimedia presentation authoring and generation supported by discourse and storytelling techniques.

References

[1] L. De Vocht, S. Coppens, R. Verborgh, M. Vander Sande, E. Mannens, R. Van de Walle (2013): Discovering Meaningful Connections between Resources in the Web of Data, Procs of the WWW2013 Workshop on Linked Data on the Web, Rio de Janeiro, Brazil.

[2] V. Lopez, V.S. Uren, M. Sabou, E. Motta: Is Question Answering fit for the Semantic Web?: A survey. Semantic Web 2(2): 125-155 (2011)

[3] R. Hong, M. Wang, G. Li, L. Nie, Z.J. Zha, T.S. Chua: Multimedia Question Answering. IEEE MultiMedia 19(4): 72-78 (2012)