# Boosting QAKiS with multimedia answer visualization

Elena Cabrio[1,2], Vivek Sachidananda[2], and Raphael Troncy[2]

[1] INRIA Sophia Antipolis, France
`firstname.lastname@inria.fr`
[2] EURECOM, France
`firstname.lastname@eurecom.fr`

**Abstract.** We present an extension of QAKiS, a system for Question Answering over DBpedia language specific chapters, that allows to complement textual answers with multimedia content from the linked data, to provide a richer and more complete answer to the user. For the demo, English, French and German DBpedia chapters are the RDF data sets to be queried using a natural language interface. Beside the textual answer, QAKiS output embeds *i)* pictures from Wikipedia Infoboxes, *ii)* OpenStreetMap, to visualize maps for questions asking about a place, and *iii)* YouTube, to visualize pertinent videos (e.g. movie trailers).

## 1 Multimedia Question Answering

The goal of a Question Answering (QA) system 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 the linked data, is a wide and emergent research area that still needs an in-depth study to benefit from the rich linked data resources on the web. Up to now, QA research has largely focused on text, mainly targeting factual and list questions (for an overview on ontology-based Question Answering systems, see [5]).[1] However, a huge amount of increasingly multimedia contents are now available on the web on almost any topic, and it would be extremely interesting to consider them in the QA scenario, in which the best answers may be a combination of text and other media answers [4].

This demonstration presents an extension of QAKiS [2], a Question Answering system over DBpedia [1], that allows to exploit the structured data and metadata describing multimedia content on the linked data to provide a richer and more complete answer to the user, combining textual information with other media content. A first step in this direction consists in determining the best sources and media (image, audio, video, or a hybrid) to answer a query. For this reason, we have carried out an analysis of the questions provided by the Question Answering over Linked Data (QALD) challenge, and we have categorized them

---

[1] See the Question Answering over Linked Data (QALD) challenges `http://greententacle.techfak.uni-bielefeld.de/~cunger/qald/`

according to the possible improved multimedia answer visualization. Then, we have extended QAKiS output to include *i)* pictures from Wikipedia Infoboxes, for instance to visualize images of people or places (for questions as *Who is the President of the United States?*); *ii)* OpenStreetMap, to visualize maps for questions asking about a place (e.g. *What is the largest city in Australia?*) and *iii)* YouTube, to visualize videos related to the answer (e.g. a trailer of a movie, for questions like *Which films starring Clint Eastwood did he direct himself?*).

## 2 Extending QAKiS to visualize multimedia answers

***QAKiS system description.*** QAKiS (Question Answering wiKiFramework-based System) [2] addresses the task of QA over structured knowledge-bases (e.g. DBpedia) [3], where the relevant information is expressed also in unstructured forms (e.g. Wikipedia pages). It implements a relation-based match for question interpretation, to convert the user question into a query language (e.g. SPARQL). More specifically, it makes use of relational patterns (automatically extracted from Wikipedia and collected in the WikiFramework repository [2]), that capture different ways to express a certain relation in a given language. QAKiS is composed of four main modules (Fig. 1): *i)* the *query generator* takes the user question as input, generates the typed questions, and then generates the SPARQL queries from the retrieved patterns; *ii)* the *pattern matcher* takes as input a typed question, and retrieves the patterns (among those in the repository) matching it with the highest similarity; *iii)* the *sparql package* handles the queries to DBpedia; and *iv)* a *Named Entity (NE) Recognizer*.
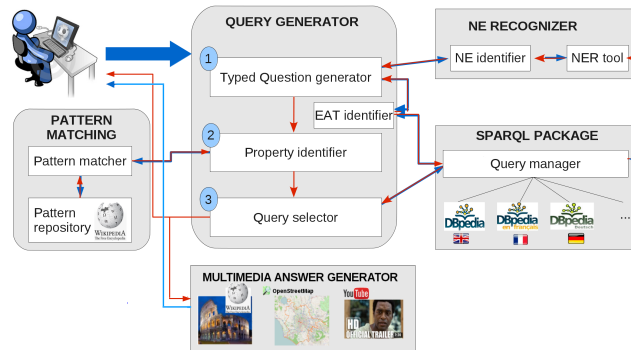


**Fig. 1.** QAKiS workflow [2]

The actual version of QAKiS targets questions containing a Named Entity related to the answer through one property of the ontology, as *Which river does the Brooklyn Bridge cross?*. Such questions match a single pattern (i.e. one relation).

Before running the *pattern matcher* component, the question target is identified combining the output of Stanford NE Recognizer, with a set of strategies

that compare it with the instances labels in the DBpedia ontology. Then a *typed question* is generated by replacing the question keywords (e.g. who, where) and the NE by the types and supertypes. A Word Overlap algorithm is then applied to match such typed questions with the patterns for each relation. A similarity score is provided for each match: the highest represents the most likely relation. A set of patterns is retrieved by the pattern matcher component for each typed question, and sorted by decreasing matching score. For each of them, a set of SPARQL queries is generated and then sent to the endpoints of language specific DBpedia chapters that the user has selected. If no results are found the next pattern is considered, and so on. Currently, the results of a SPARQL query on the different language specific DBpedia chapters are aggregated by the set union.

***QAKiS multimedia.*** While providing the textual answer to the user, the *multimedia answer generator module* queries again DBpedia to retrieve additional information about the entity contained in the answer. To display the images, it extracts the properties `foaf:depiction` and `dbpedia-owl:thumbnail`, and their value (i.e. the image) is shown as output. To display the maps (e.g. when the answer is a place), it retrieves the GPS co-ordinates from DBpedia (properties `geo:geometry`, `geo:lat` and `geo:long`), and it injects them dynamically into OpenStreetMap[2] to display the map. Given the fact that DBpedia data can be inconsistent or incomplete, we define a set of heuristics to extract the co-ordinates: in case there are several values for the latitude and longitude, *i)* we give priorities to negative values (indicating the southern hemisphere[3]), and *ii)* we take the value with the highest number of decimal values, assuming it is the most precise. Finally, to embed YouTube[4] videos, first the Freebase[5] ID of the entity is retrieved through the DBpedia property `owl:sameAs`. Then, such ID is used via YouTube search API (v3) (i.e. it is included in the embed code style <`iframe`>, that allows users to view the embedded video in either Flash or HTML5 players, depending on their viewing environment and preferences). Moreover, since we want to have pertinent videos (i.e. showing content related to the answer in the context of the question only), we remove stopwords from the input question, and we send the remaining words as search parameters. For instance, for the question *Give me the actors starring in Batman Begins*, the words "actors", "starring", "Batman Begins" are concatenated and used as search parameters, so that the videos extracted for such actors are connected to the topic of the question (i.e. the actors in their respective roles in Batman Begins).

## 3   QAKiS demonstrator

Figure 2 shows QAKiS demo interface (`http://qakis.org/`). The user can select the DBpedia chapter she wants to query besides English (that must be selected as

---

[2] `www.openstreetmap.org`

[3] We verified that when both a positive and a negative value are proposed, the negative is the correct one (the letter S, i.e. South, is not correctly processed.)

[4] `www.youtube.com/`

[5] `www.freebase.com`

it is needed for NER), i.e. French or German DBpedia. Then the user can either write a question or select among a list of examples, and click on *Get Answers!*. As output, in the tab *Results* QAKiS provides: *i)* the textual answer (linked to its DBpedia page), *ii)* the DBpedia source, *iii)* the associate image in Wikipedia Infobox, *iv)* a *more details* button. Clicking on that button, both the entity abstract in Wikipedia, the map and the retrieved videos (if pertinent) are shown. In the tab *Technical details*, QAKiS provides *i)* the user question (the recognized NE is linked to its DBpedia page), *ii)* the generated typed question, *iii)* the pattern matched, and *iv)* the SPARQL query sent to the DBpedia SPARQL endpoint. The demo we will present follows these stages for a variety of queries, described in the next section.
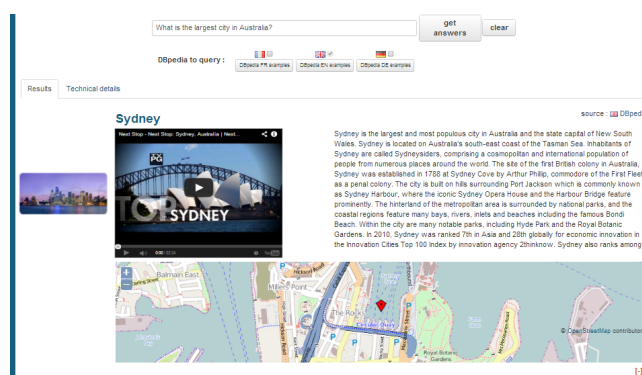


**Fig. 2.** QAKiS demo interface

### 3.1 Queries and datasets for demonstration

In order to determine the best sources and media (image, audio, video, or a hybrid) to answer a query, we have carried out an analysis on a subset of the questions provided by the QALD-3 challenge.[6] The goal was to categorize them according to the possible improved multimedia answer visualization, and to extract some heuristics to be exploited by QAKiS to provide the most complete answer to a certain question. In this analysis, we discarded the questions for which no additional multimedia content would be pertinent, e.g. questions whose answer is a number (e.g. *How many students does the Free University in Amsterdam have?*), or boolean questions (e.g. *Did Tesla win a nobel prize in physics?*). In future work we could provide multimedia content on the entity in the question, but in the current work we are focusing on boosting the answer visualization only. Table 1 shows the categories of multimedia content for answer visualization on which we are focusing, together with an example of question for which such kind of multimedia content would be appropriate.

---

[6] `http://doi.org/10.4119/unibi/citec.2013.6`

| Multimedia | Example question |
|---|---|
| Picture | *Give me all female Russian astronauts.* |
| Picture + video | *Give me all movies directed by Francis Ford Coppola.* |
| Picture + map | *In which country does the Nile start?* |
| Map + barchart | *Give me all world heritage sites designated within the past 5 years.* |
| Statcharts | *What is the total amount of men and women serving in the FDNY?* |
| Timelines | *When was Alberta admitted as province?* |

**Table 1.** QALD-3 questions improved answer visualization

## 4 Future perspectives

The work we present in this demonstration is ongoing, and represents a first step in the direction of dealing with the huge potential amount of available multimedia data. As a short-term improvement, we are planning to add other sources of images. For instance, Ookaboo RDF Data contains pictures with topics derived from Freebase and DBpedia, and can therefore be coupled with the output of QAKiS, to provide additional images describing the answer. For other available datasets, metadata could be RDF-ized (e.g. MIRFLICKR, IMAGENET[7]) and the interlinking of such structured sources with DBpedia can be explored to provide to the user semantically enriched multimedia presentations. As a long-term improvement, we plan to 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 [4]. 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 [6].

## References

1. C. Bizer et al. DBpedia - a crystallization point for the web of data. *Web Semant.*, 7(3):154–165, Sept. 2009.
2. E. Cabrio, J. Cojan, A. P. Aprosio, B. Magnini, A. Lavelli, and F. Gandon. QAKiS: an open domain qa system based on relational patterns. In *Proceedings of the ISWC 2012 Posters and Demonstrations Track*, Boston, US, November 2012.
3. E. Cabrio, J. Cojan, F. Gandon, and A. Hallili. Querying multilingual dbpedia with qakis. In *ESWC (Satellite Events)*, pages 194–198, 2013.
4. R. Hong, M. Wang, G. Li, L. Nie, Z.-J. Zha, and T.-S. Chua. Multimedia question answering. *IEEE MultiMedia*, 19(4):72–78, 2012.
5. V. Lopez, V. S. Uren, M. Sabou, and E. Motta. Is question answering fit for the semantic web?: A survey. *Semantic Web*, 2(2):125–155, 2011.
6. L. D. Vocht, S. Coppens, R. Verborgh, M. V. Sande, E. Mannens, and R. V. de Walle. Discovering meaningful connections between resources in the web of data. In *LDOW*, 2013.

---

[7] `http://press.liacs.nl/mirflickr/`; `https://www.image.net/`