Gathering Photos from Social Networks using Semantic Technologies

M.B. Alves[1,2], C. V. Damásio[1], N. Correia[3]

CENTRIA, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
ESTG, Instituto Politécnico de Viana do Castelo 4900-348 Viana do Castelo.
CITI, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
mba@estg.ipvc.pt, cd@fct.unl.pt,nmc@fct.unl.pt

Abstract. We present a system based on a Semantic Web approach to retrieve photos from events, organizations or even related to semantic concepts using context and social information, without require extra annotation work by the publishers. We make use of the knowledge of the system, represented in ontologies and semantic rules, to allow users search using its own terminology.

1   Introduction

The motivation of our work is moving towards an easy searching and browsing of photos collections as is required by the huge amount of digital photo collections made available on the web by the photographers in different repositories, like Facebook, Picasa or Flickr. The gap between the photo information and the users desire is addressed by us through a Semantic Web approach, namely by a) associating photo metadata with precisely defined semantics, represented through ontologies [4], and b) through reasoning over this information. In this work, we describe a system that searches for photos on the web, that can be tailored to a specific domain like events, organisations, sports. We combine information that we have in ontologies with the information about the photos, either photo metadata or contextual information, to retrieve photos of a given event, from a given person, or in a given place. This system acts as a semantic mashup of photos collection, in other words, as a personal collection, but in reality these photos are spread in the web, social networks and photo databases. Our system requires a knowledge model of the domain, represented by a domain ontology and by rules. This information model gives to the system knowledge about the domain, allowing queries in a terminology recognised by the user. As a result, we are dealing with the semantic gap that exists in multimedia content [9], the lack of coincidence between the information that can be extracted from the photo and the meaning of that photo to the human.

This document is organised as follows. In section 2 we present our approach and we explain how a semantic web approach can meet our purposes in photo retrieval. In section 3 we detail our system architecture and we explain each component of the architecture and we finish with the conclusions where we analyze our work and we present some benchmarking in section 4.

2 Semantic Web approach for multimedia retrieval

In our Semantic Web approach we use ontologies to formally describe the domain. Upper-level ontologies, describing very general concepts that are the same across all domains, are used to modelling concepts such as Events, Photos, Time, etc. In this work we have used: Ontology for Media Resources [5], to describe media resources; LODE [7], ontology for Linking Open Descriptions of Events; Time ontology [1], FOAF [2] for describing persons. Each domain, for which we want to implement a photo retrieval engine with the approach described in this work, requires a particular domain ontology to represent the domain-specificities. We will model the domain ontology using OWL 2 [3]. To overcome some of the OWL limitations [6], we also use semantic rules to add expressivity and expertise to our model, giving support to some object properties, which link individuals to individuals.

Now, we will illustrate how our Semantic Web approach is used to support photos retrieval, giving meaning to the content. To simplify reading we will use the well known prefixes of the used ontologies and the prefix *don* is our domain ontology. In our domain ontology, we define an **objectProperty** *isPhotoOfEvent* having as domain *ma:Image* and as range *lode:Event* to make the relationship between photos and events instances. We define some rules to deduce, in our system, when a given *Photo* can belong to a given *Event*. These are general rules independent of the specific domain. **Rule 1** is one of these rules, present using the syntax `Conclusion <- Premises(body atoms)`.

Rule 1: `(?Photo don:isPhotoOfEvent ?Event) <-`

`(?Photo don:wasTakenAtTimeOf ?Event), (?Photo don:wasTakenInSamePlace ?Event),`

`(?Person don:isTaggedIn ?Photo), (?Person don:participates ?Event).`

With **Rule 1**, we declare that a given *Photo* belongs to a given *Event* if: *s1*) the *Photo* was taken at the same time of the *Event*; *s2*) that *Photo* was taken in the same place of the *Event*; *s3*) there is at least one *Person* that was tagged in the *Photo*; *s4*) who participated in the *Event*. The first statement, *s1*, makes use of the **objectProperty** *wasTakenAtTimeOf* which is supported by rules to define if a given photo was taken while the event occurred. The *isTaggedIn* and *participates* are also object properties that define if a *Person* is tagged in a *Photo* and if a *Person* participates in a *Event*. Notice that **Rule 1** cannot be captured by OWL 2 role inclusion chain axioms.

In our system, we allow to give a confidence to a relation (real number between 0.0 and 1.0). The RDF reification vocabulary is extended to associate a confidence to each inferred triple used in search dimensions/values, according to the following pattern stating that a given *Relation* between a *Subject* and an *Object* has a given *Confidence*:

`?bn rdf:subject Subject .    ?bn rdf:predicate Relation .`

`?bn rdf:object Object .    ?bn don:confidence Confidence .`

As we can define queries with confidence we can define weaker relations, for instance between photos and events, and give a confidence to them. This confidence will be used in ranking the photos.

3 System Architecture

In figure 1 is represented the architecture of our tool. Our system uses the APIs of Facebook, Flickr and Picasa to retrieve the photos, the information associated with the photo and the context information. We simply use the information that we can obtain with the photo and the normal operations that the users usually perform: tagging photos, with text tags or people, indicate the place where the photos were taken, or give a name to an album. The information is extracted from the multimedia databases and is kept in a knowledge base to support the inference engine. The photos are available to the users through a web interface that retrieve the photos that answer to the users queries.

ETL process - The Extract, Transform and Load process (ETL), acts like a web crawler. The metadata that we can get from these new photos is extracted and saved using the vocabulary specified by the Ontology for Media Resources. The transform task can make use of the semantic model and the rules to do inference and materialize some of the consequences in the pre-processing task due to performance issues.

Fig. 1. System architecture of the photos retrieval engine

Semantic Model and rules - Each domain has its own vocabulary, the terms that are recognised by the users. Therefore, a domain ontology must be developed to represent the application domain, providing the knowledge of the domain that allows us to achieve a higher precision in the retrieval process. Without a domain ontology, we couldn't use these terms as a search dimension/facet.

The query process and the interface - The users make queries through the Web interface. The system returns the links to the photos that are related to the queries. These queries are answered using the information kept in the knowledge base, making inferences from the semantic model and the rules. We have a semantic model to support the querying. This semantic model is a meta-model within which the classes of search dimensions are defined in the web interface. This meta-model, that must be tailored to each different domain, allows the system to know what must be searched. With this approach, we do not need any change in code in a new system. Everything is knowledge provided to the system and everything is kept outside in simple configuration files. As we can give a confidence to a relation, we can rank the images using the confidence value of inferred triples.

System details - Our implementation is based on Jena framework[1], which offers an "all-in-one" solution for Java, OWL reasoning, inference and rule engine. We use Pellet [8] together with Jena to make OWL 2 inference. The TDB component of Jena is used for RDF storage and query.
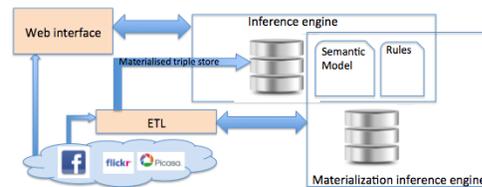
---

[1] https://jena.apache.org

4 Conclusions

In this work, we presented a system to retrieve photos using a Semantic Web approach. Our system uses the context information of the photo or annotations done by the user and other metadata that can be retrieved from social networks to combine with the knowledge of the system to classify the photos. The user can search for photos using its own terminology, creating dynamically its owns personal collections despite these photos are distributed along the web. . We can perform this because we have knowledge of the domain and, in this way, we try to overcome the problem of the semantic gap between the photos information and the means of that photo to the user.

The precision and the recall of our system it is an open question, and depends of each new implemented domain and the requirements. If we want to give relevance to the precision, we define rules that represent exactly one relationship or, at most, with a high probability of occurrence. If we want to give relevance to the recall, we define rules to represent how relationships may happen, even if those rules bring some fake positive results. As we can introduce a confidence factor in our rules, we can give relevance to the recall but improving the F1-Score, a measure that combines precision and recall. Even though precision and recall be an open question, we performed some benchmarking with a small example using a Facebook account of a swimming club with 3599 friends. The system focused on 49 albums and 1148 photos. It were retrieved 431 photos, 425 of which were distributed over 9 albums. We had 100% precision in the retrieved photos. We inspected the photos of the swimming club events that were published but that were not retrieved, achieving a recall of 95,5%. An implementation in a real situation is on-going work that can be tried at `http://www.estg.ipvc.pt/~mba/SemanticPhotosSearch/`.

The system presented in this work was tailored to be enhanced at production time. The semantic model and the rules are kept in configuration files. Any change to these files only requires an re-initialisation of the system. Thus, we can readily improve our semantic model or our rules, adding new knowledge to our system or refining the existing one.

References

1. Time ontology in OWL. Electronic, September 2005.
2. D. Brickley and L. Miller. FOAF Vocab. Spec. 0.97. Namespace doc., 2010.
3. Cuenca et al. OWL2: The next step for OWL. *Web Semantics*, 6(4):309–322, 2008.
4. T. Hofweber. Logic and ontology. In *The Stanford Ency. of Phil.* Spring, 2009.
5. Lee et al. Ontology for Media Resources 1.0, Feb. 2012. W3C Recommendation, http://www.w3.org/TR/mediaont-10/.
6. B. Parsia et al. Cautiously approaching SWRL, 2005.
7. R. Shaw et al. Lode: Linking open descriptions of events. In *Proc. of the 4th Asian Conf. on The Semantic Web*, ASWC '09, pages 153–167, Berlin, Heidelberg, 2009.
8. E. Sirin et al. Pellet: A practical OWL-DL reasoner. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):51–53, jun 2007.
9. Smeulders et al. Content-based image retrieval at the end of the early years, 2000.