Semantic Annotation of Images on Flickr⋆

Pierre Andrews, Sergey Kanshin, Juan Pane, and Ilya Zaihrayeu

Department of Information and Communication Technology
University of Trento Italy
{andrews,kanshin,pane,ilya}@disi.unitn.it

Abstract. In this paper we introduce an application that allows its users to have an explicit control on the meaning of tags they use when uploading photos on Flickr. In fact, this application provides to the users an improved interface with which they can add concepts to photos instead of simple free-text tags. They can thus directly provide semantic tags for their photos that can then be used to improve services such as search.

1 Introduction

The task of discovering the semantics of photos is still very difficult and hence, automatic annotation with concepts "is widely recognized as an extremely difficult issue" [1]. It is thus preferable to ask the creators of the photos to annotate them directly when they share them. As was demonstrated in [2], simple free-text annotations are not sufficient for performing good indexing and leveraging semantic search.

In this paper we discuss an extension to a popular open source photo uploader for the Flickr[1] website that allows the annotation of image files with semantic annotations without extra involvement from the user. One of the feature of this tool is the bootstrapping of semantic annotations by extraction of the intrinsic semantics contained in the context in which the images reside on the local computer of the user before uploading them to Flickr by using the technology described in [3]. The users can also manually provide semantic annotations through an extended interface. These semantic annotations are linked to their meaning in a knowledge organisation system such as Wordnet[2].

The source code for the tools described in this paper is available freely at https://sourceforge.net/projects/insemtives/.

2 Semantic Annotation Platform

The Semantic Annotation tool is built as an application on top of the INSEMTIVES platform. Interested readers are referred to [4] for details on this platform.

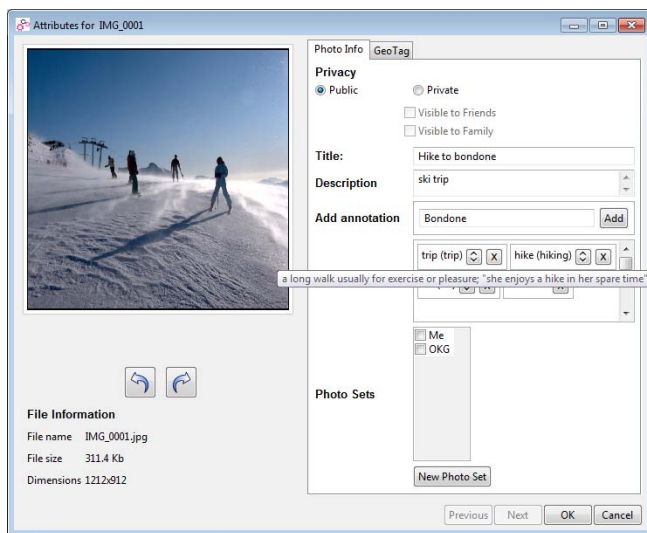[1] http://www.flickr.com
[2] http://wordnet.princeton.edu/

Fig. 1. Image properties with added concepts

The platform services are exposed through the INSEMTIVES platform API's, which can then be used by third party applications. The API's are based on a communication framework which represents a highly modular client-server architecture implemented using communication layer over JMS[3] messaging protocol as well as six other protocols, including REST[4] and Web Service Notification. The platform is divided in two main components: *Structured Knowledge* and *User and Communities*. The Structured Knowledge component stores all artifacts in RDF following the semantic annotation model defined in [5]. The semantic store relies on OWLIM[5] while the initial lexical resource is based on Wordnet and DBPedia. The User and Communities component contains user profiles and is responsible for maintaining provenance information of the annotations and resources in the platform.

3   Semantic Photo Annotation

The semantic annotation application has been developed as an extension to the JUploadr[6] uploading tool for Flickr. This tool has been chosen as it is open source and developed in Java which is the main development language for the INSEMTIVES platform and makes the tool easily multi-platform.

---

[3] Java Messaging Service, see `http://java.sun.com/products/jms/`

[4] `http://en.wikipedia.org/wiki/Representational_State_Transfer`

[5] `http://www.ontotext.com/owlim/`

[6] `http://www.juploadr.org/`

jUploadr allows the user to queue photos for uploading on Flickr and set their properties (see Figure 1). In particular, the tags, description and title, which are the main metadata stored and displayed by the Flickr website.

3.1 Semi-automatic Semantic Annotation

Concept Tagging The INSEMTIVES Uploadr tool will try to understand where the users have stored their photos and automatically add tags to them. For instance, if the users have stored their photos in the folder: *.../Pictures/Trips/ Trentino/Hike to Bondone-10-01-10/IMG_0001.jpg* then the software will understand that these were photos of a trip in the Trentino's mountains and will automatically propose relevant *concepts* (see Figure 1).

When a new photo is added to the upload batch, the following processing takes place:

1. clean and split the photo path in *phrases*,
2. send the ordered set of *phrases* to the INSEMTIVES platform,
3. receive a set of *tokens* for each *phrase* with the corresponding *concepts*,
4. attach each identified *concept* to the photo to be uploaded as a semantic annotation.

The cleaning of the path is performed to remove generic parts of the path that might not contain interesting semantic information for bootstrapping. The cleaning is specific to that particular type of media and the one applied for image files is the following: 1. if the path starts with the name of the user's home directory, drop the beginning. 2. if the path starts with /media or /mnt, drop the beginning. 3. if the path contains the word "picture(s)", "Media", "Photos" drop anything on the left of that word. Including that word. 4. treat _ as spaces.

Each section of the path (between path separators) is treated as a *phrase* – i.e., a short sentence made of related tokens. The set of phrases in the path is sent in an ordered way to the NLP service to be used as a context for disambiguation. For instance, "Hike to Bondone" will be treated in the context of "Trip" and thus "hike" will be disambiguated to "a long walk usually for exercise or pleasure" instead of "the amount a salary is increased"[7](see Figure 1).

These special tags have a "meaning" (or "semantics") attached to them, a summary of that meaning is always displayed next to the concept and if the user hovers the mouse over one of them, it's definition is displayed (see Figure 1).

Location Tagging. In the "geo" tab of the properties window for a photo, the user can specify where the photo was taken. Once a location is specified for a photo, the user can press the "Find concepts" button to ask the INSEMTIVES tool to find automatically new concepts that relate to the location where the photo was taken by looking at popular terms already used for this location on Flickr. These concepts will then be added to the list of semantic tags in the "Photo Info" tab, where the user can remove the ones that might not fit the photo.
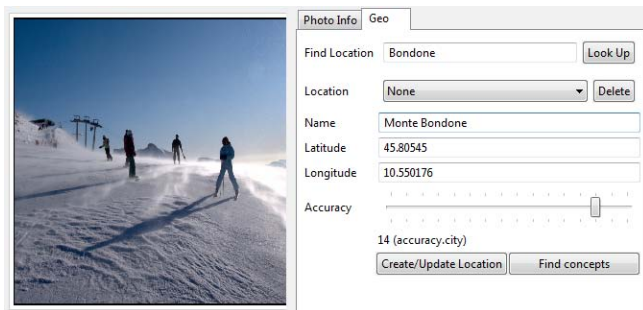
---

[7] These senses are taken from WordNet.

Fig. 2. Specifying Location of an Image

3.2   Manual Semantic Annotation

The user can also provide manual semantic annotations. When typing a free-text tag, the user is given a list of possible concepts from the platform knowledge base. Each concept is disambiguated by a summary word (see Figure 3.2a) and hovering it provides its definition.

a) Concept Completion   b) Concept Disambiguation
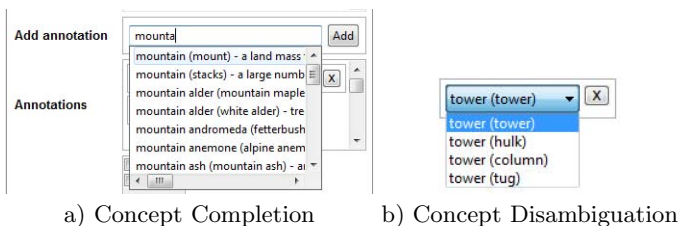
For the concepts that were automatically proposed by the services described earlier, the user can correct the disambiguation by selecting the right sense from a drop-down list. Here also, a summary is shown and the user can display the definition of the concept (see Figure 3.2b)) .

4   Semantic Search

On the Flickr website, photos can only be searched by keywords, hence, if a photo is tagged with "cat" and another one with "dog", these photos will not be found when searching for "animal".

However, if the photos were uploaded with the INSEMTIVES Uploadr and were assigned the *concepts* "cat" and "dog", then they can be retrieved with a semantic search for "animal". To do this, the INSEMTIVES Uploadr provides a specific search interface as shown in Figure 3.

Fig. 3. Example of a Semantic Search. A search for "journey" found photos about a "trip".

5 Demonstration

The goal of the demonstration will be to show the extension of the standard open source jUploadr application with the specific semantic annotation tools. The demonstration will be split in two main scenarios:

Annotation of Photos the new tools for semantic annotation of images will be demonstrated in a scenario showing the use of automatic concept extraction from the local context, recommendation of concepts from the location and manual input of concepts to annotate a photo.

Semantic Search for Photos once some photos have been annotated and uploaded on Flickr, we will demonstrate the semantic search tool that is able to retrieve photos, no only on the tags attached to them, but also on the concepts used for annotation, thus finding related photos by mapping synonymous terms to the same concept and by reasoning about the subsumption relationship between concepts.

The visitors will also be shown how to download and install the tool for their own use with their Flickr accounts so that they can upload ESWC'11 photos with semantic annotations.

References

1. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv. 40, 5:1–5:60 (2008)
2. Andrews, P., Pane, J., Zaihrayeu, I.: Semantic disambiguation in folksonomy: a case study. In: Bernardi, R., Chambers, S., Gottfried, B., Segond, F., Zaihrayeu, I. (eds.) Advanced Language Technologies for Digital Libraries. LNCS Hot Topic subline. Springer, Heidelberg (2011)
3. Zaihrayeu, I., Sun, L., Giunchiglia, F., Pan, W., Ju, Q., Chi, M., Huang, X.: From web directories to ontologies: Natural language processing challenges. In: Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L.J.B., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., Cudré-Mauroux, P. (eds.) ASWC 2007 and ISWC 2007. LNCS, vol. 4825, pp. 623–636. Springer, Heidelberg (2007)
4. Siorpaes, K., Konstantinov, M., Popov, B.: Requirement analysis and architectural design of semantic content management platform. Technical report, Insemtives.eu (September 2009)
5. Andrews, P., Zaihrayeu, I., Pane, J., Autayeu, A., Nozhchev, M.: Report on the refinement of the proposed models, methods and semantic search. Technical report, Insemtives.eu (November 2010)