Annotation and Provenance Tracking in Semantic Web Photo Libraries

Christian Halaschek-Wiener, Jennifer Golbeck, Andrew Schain, Michael Grove,
Bijan Parsia, and Jim Hendler

University of Maryland, MIND Lab, 8400 Baltimore Ave., College Park, MD 20742, USA
{halasche, golbeck, hendler}@cs.umd.edu, andrew.schain@nasa.gov
mhgrove@hotmail.com, bparsia@isr.umd.edu

Abstract. As the volume of digital images available on the Web continues to
increase, there is a clear need for more advanced techniques for their effective
retrieval and management. In this paper, we present a domain independent
framework for both annotating and managing images on the Semantic Web. We
introduce a tool that facilitates creating and publishing OWL annotations of im-
age content to the Semantic Web. This is loosely coupled with a Semantic Web
portal with provenance tracking. We illustrate the effectiveness of this system
with an implementation of the approach and describe a hypothetical use case
that resulted in a proof-of-concept designed in collaboration with NASA.

1 Introduction

As the scale and infrastructure of the Internet have dramatically increased over the
past years, we have seen the incorporation of various digital media types onto the
Web, including images, video, and audio. As production of digital media content
continues to grow in the commercial and home use markets, and as Internet access
and wider bandwidth become even more pervasive, we can anticipate a continued
increase of these complex (non-textual multimedia) data types being made available
on the Web. Due to the format of such media, standard indexing techniques com-
monly used on text-based Web content, such as keyword-based approaches [2], are of
little use. Given the volume of unstructured digital media, it is clear additional ap-
proaches and techniques must be developed to allow for their effective management
and accurate retrieval.

Over the past few years, various approaches have been proposed to effectively re-
trieve and manage digital image content on the Web. Traditionally, these have in-
cluded techniques such as building keyword indices based on image content [7, 9],
embedding keyword-based labels into images [7], analyzing text immediately sur-
rounding images on Web pages [4], etc. More recently, there has been a research
focus to develop techniques to annotate the content of images on the Semantic Web,
using languages such as RDFS and OWL [1,3,5,6,11].

Recent efforts have largely focused on mapping low-level features of images to on-
tological concepts [1,3,11] and have involved the development of tools that are
closely tied to domain specific ontologies for annotation purposes [6,8]. Additionally,

past approaches have largely left unaddressed image metadata management and advanced interaction (browsing and search capabilities) that is enabled by employing Semantic Web technologies. While substantial progress has been made, we see the need for further work in defining a more generic approach for annotating and managing digital images on the Web.

In this work, we present an approach that provides generic, domain independent flexibility for publishing annotations of digital image content to the Semantic Web, as well as a mechanism for managing such annotations and tracking their provenance through a highly customizable, ontology-backed Semantic Web portal. Through the loose coupling of the annotation and management components of our approach, a seamless environment is provided in which users can annotate, share, and manage their digital images on the Semantic Web.

2 Motivation and Approach Overview

To understand the generic requirements that have driven the approach presented here, a representative use case based on a collaboration with NASA. While this motivation is presented in the context of NASA, we feel the model is sufficiently generic, thus capturing the general issues associated with managing metadata of digital images.

As an enterprise, NASA has hundreds of thousands of images, stored in different formats and locations, at different levels of availability and resolution, and with associated descriptive information at various levels of detail and formality. NASA also generates thousands of images on an ongoing basis that are collected and cataloged, often in accordance with needs of the image creator's specific disciplines and domain (preliminary investigators, mission specialists, public affairs, etc.). It is clear that a mechanism is needed to catalog all the different types of image content across different domains. Information is required about both the image itself (creation date, dpi, source, etc.) and also about the content of the picture (contains a satellite, astronaut, etc). The associated metadata must be maintainable and extensible so associated relationships between images and data can evolve cumulatively within a discipline or branching into other disciplines. The service must be available to a global consumer population but should be flexible enough to enforce restriction based on content type, ownership, authorization, or time.

A promising strategy for such image management requirements is an annotation environment that enables both providers and users to annotate information about images or regions in images using concepts in ontologies (OWL and/or RDFS). Thus, subject matter experts and consumers (regardless of their location) will be able to assert metadata elements about images and publish their annotations to the Semantic Web. There, such digital image annotations can be harvested and merged, resulting in advanced browsing, searching, and management.

We generalize these (NASA specific) high level requirements into the following application independent requirements: support for *adhoc* ontology-based annotation of images on the Web, enabling support for annotation with respect to *any* domain; the ability to make assertions about images and the contents of specific regions in images; the ability to automatically publish annotations to the Semantic Web, where they can be shared, indexed, and maintained; provide a metadata management facility

for interacting with and maintaining image metadata that is accessible to a global community – the Semantic Web; the ability to accumulate metadata about a specific image over a period of time from different sources.

Given these requirements, we present a loosely coupled approach that provides generic, domain independent flexibility for creating and publishing annotations of digital image content to the Semantic Web, as well as a mechanism for managing such annotations through a highly customizable, ontology-backed Semantic Web portal.

3   Implementation Details

The first component of the approach presented in this work is a digital image annotation environment. In this section we present PhotoStuff, a semantic annotation environment.

3.1   Digital Image Annotation – PhotoStuff

PhotoStuff is a platform independent, open source, image annotation tool that allows users to annotate an image and its regions with respect to concepts from any number of ontologies specified in RDFS or OWL. PhotoStuff provides functionality to import images, their embedded metadata, ontologies, and instance-bases. In the tool, users can perform markup, and export the resulting annotations. The tool provides users the ability to load multiple OWL and/or RDFS ontologies, allowing annotation of image content with respect to any concept, defined in any number of ontologies. The ability to annotate images with respect to any ontology is extremely important because the content of images can span multiple domains; frequently, a single ontology cannot capture the complexity of the content.

In PhotoStuff, an ontology-based approach has also been adopted in order to make statements regarding the high level concepts depicted in images. An ontology is used to provide the expressiveness required to assert what is depicted within an image, as well information about the image itself. In this work, an image-region ontology[1] has been specified, using OWL, which defines a set of concepts and their relations for images, videos, regions, and depictions.

To demonstrate the use of PhotoStuff, Figure 1 shows a screenshot of the tool in which as user is marking up information about an astronaut taking a space walk. The ontologies are visualized in both a class tree and list, depicted in the far left pane of the tool. In this example, the FOAF (Friend of a Friend) ontology has been loaded, as well as a Shuttle Crew ontology that is expanded in the window. This allows the user to choose concepts from both ontologies to mark up the photograph and its sub-regions.

In this approach, users can annotate the entire image, or selected regions. Users highlight regions around portions of images loaded in PhotoStuff. Figure 1 illustrates this with a region drawn around the astronaut. Classes can be dragged onto the image or into any region creating a new instance of the selected class. When a class is used, a form is dynamically generated from the properties of the selected class. With region

---

[1] Image-Region Ontology: http://www.mindswap.org/2005/owl/digital-media
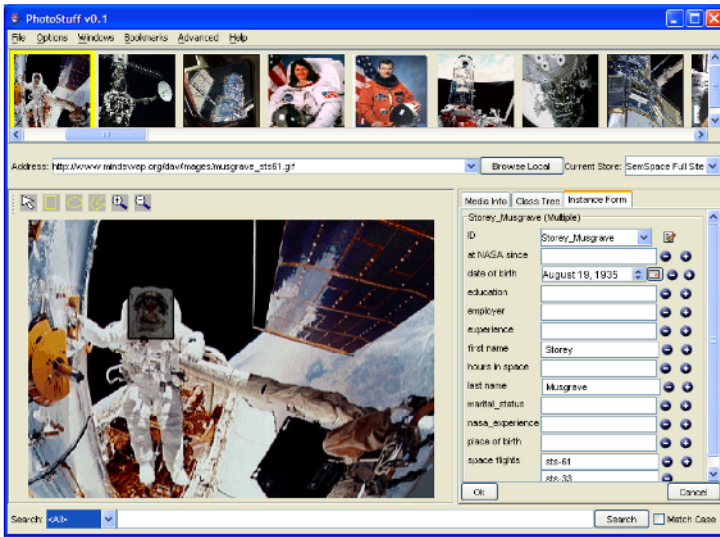
Fig. 1. PhotoStuff Screenshot

support, metadata can be more closely tied to the depiction it describes. Instead of simply stating that a photograph depicts several people, the metadata will contain coordinates for the regions of the photo that contain the depictions. The region is also semantically linked to the image, maintaining the connection between the image and the instance.

Additionally, the approach presented here leverages current efforts in multimedia format standardizations that provide support to embed image metadata in actual image files. For example, the JPEG file format provides support for embedding a standard set of markers in the file header, defining metadata elements including file size, width/height, pixel density, etc. The Exchangeable Image File Format (EXIF), also provides metadata in the form of camera specific information (camera make, model, orientation, etc.)[2]. Our approach takes advantage of this existing metadata by extracting and encoding it into RDF/XML so it is accessible on the Semantic Web.

As mentioned earlier, PhotoStuff, and the approach in general, maintains a loose coupling with a Semantic Web portal. As briefly discussed before, there are three ways in which PhotoStuff interacts with the portal: retrieving all instances that have been submitted to the portal, submitting generated RDF/XML, and uploading local images so they can be referenced by a URI (thus allowing them to be referenced using RDF/XML). The following section outlines the metadata management and browsing functionality provided though the loose coupling of the annotation environment with the Semantic Web portal.

3.2 Image Metadata Management

Upon the completion of image annotation, the approach provides the capability for publishing resulting markup to the Semantic Web. This is accomplished through the

---

[2] EXIF Homepage: http://www.exif.org/

Fig. 2. Instance Depictions and Co-Region Browsing

coupling of the annotation environment with an ontology-backed, Semantic Web portal. Our existing work on a Web portal based on Semantic Web technologies (OWL) has been extended to provided communication with PhotoStuff. It is noted here that the Web portal's functionality extends what is presented here and is an ongoing project within the MINDSWAP[3] research group. Details are provided here through one of its configurations in the context of a proof-of-concept, SemSpace[4], developed as an experiment with NASA. A variety of other domain configurations have been developed at MINDSWAP; all configurations provide the same functionality, only differing by the ontologies and instances maintained by the system.

The portal technology is flexible enough to be used in a variety of domains, as it is not limited in the number of ontologies that it can manage; thus for the purpose of this work, any ontology can be used to annotate an image. The portal is designed to use information from the various ontologies to guide the display of and interaction with metadata and the site in general. The main interface for browsing images is driven by the underlying class of each instance, thus providing a high level view of all the metadata of images that have been annotated using PhotoStuff.

The portal provides the ability to browse data associated with instances, images, image regions, and to search metadata in the collection (in Figure 2). The portal component also provides various management capabilities. Metadata submissions can be audited, edited, or removed. Provenance information (submitter name, email, etc.) from all submissions is maintained and editable. For each statement, the provenance information is also provided when the user hovers their mouse over any annotation.

---

[3] MINDSWAP Research Group: http://www.mindswap.org/
[4] SemSpace Homepage: http://semspace.mindswap.org/

3.2   Provenance Management

As mentioned previously, PhotoStuff and the portal technology exploit and provide support for management of provenance data. When users submit annotations to a portal, their user name and a comment is required for input. When the portal receives this data, a timestamp is additionally recorded. This provenance data is stored, along with the annotations in a RDF triple store, RDFLib[5].

Each time a new submission is received, a new submission resource is created and stored in the triple store. Additionally, the submission object has various attributes related to it, including the submitter, timestamp, etc.
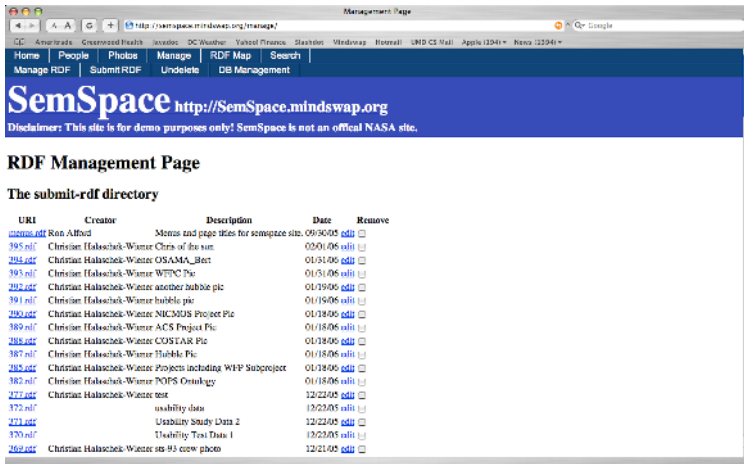

Fig. 3. RDF Submission Management Interface with Provenance Data

Fig. 4. An excerpt of a page in the site showing Mouse Pop-ups of provenance data when site browsing

---

[5] RDFLib Project Homepage: http://rdflib.net/

The portal technology provides a submission metadata management page, in which submission can be edited and/or removed. In order to assist system administrators in finding submissions, the provenance data is presented in the interface (see Figure 3).

Provenance data maintained on the portal is also utilized to enhance the user experience when browsing data on the portal. When resources are browsed on the website, pop-ups are provided if provenance data is known about the particular data be viewed. This is shown in Figure 4.

4 Discussion and Future Directions

The approach discussed in this paper allows ad hoc, manual annotation of image content. This provides a cumulative technique where metadata can be incrementally added or repurposed for future users on a per-need basis. While manual annotation is essential for such ad hoc additions or edits, it can prove to be quite time consuming. This may be slightly alleviated through use of various image processing and automated vision techniques. First, region segmentation techniques may be used to suggest possible regions of interest. Additionally, image-processing techniques could potentially be used to recognize similar regions among photos, allowing the tool to suggest potential instances that may be depicted in the image. By exploiting similarity in images that are part of the same "album", automated recognition techniques can be used to take a first pass at labeling parts of the image. Once images are automatically labeled, users can then simply verify the resulting annotations [10].

Additionally, in this work it has been observed that generating effective, yet generic forms based on class definitions can be quite difficult (in this context, instance creation forms are generated when classes are dragged into image). We have adopted an approach in which the form is directly built from the underlying properties of the class. While this approach is a plausible first step, it can result in a very messy or congested form. We would like to explore allowing the user to create custom forms for classes. Additionally, we would like to investigate allowing ontology creators to embed HTML forms or XForms[6] into comments on class definitions.

5 Conclusions

In this work we have presented a generic, domain independent framework for annotating and managing digital image content using Semantic Web technologies. We loosely couple an annotation component with a Semantic Web portal that supports browsing, searching and managing digital image annotations and provenance information. Additionally, we have provided details of an open source implementation of this framework and an overview of a representative proof-of-concept. Potential future work includes automating portions of the annotations process, possibly by using image processing and computer vision techniques. We also plan to extend our work here to support annotation of additional digital media types, including video and audio.

---

[6] XForms 1.0: http://www.w3.org/TR/xforms/

References

1. Addis, M., Boniface, M., Goodall, S., Grimwood, P., Kim, S., Lewis, P., Martinez, K. and Stevenson, A. SCULPTEUR: Towards a New Paradigm for Multimedia Museum Information Handling. Second International Semantic Web Conference (2003) 582 -596

2. Brin, S., and Page, L. The Anatomy of a Large Scale Hypertextual Web Search Engine, *In the Proceedings of the 7th International World Wide Web Conference* (1998)

3. Dupplaw, D., Dasmahapatra, S., Hu, B., Lewis, P., and Shadbolt, N. Multimedia Distributed Knowledge Management in MIAKT. *ISWC 2004 Workshop on Knowledge Markup and Semantic Annotation*. Hiroshima, Japan, November 2004

4. Frankel, C., Swain, M., and Athitsos, V. Webseer: An Image Search Engine for the World Wide Web, Tech. Report TR-96-14, Computer Science Dept., Univ. of Chicago, July (1996)

5. Hollink, L., Schreiber, G., Wielemaker J., and Wielinga. B. Semantic Annotation of Image Collections. Knowledge Capture - Knowledge Markup & Semantic Annotation Workshop (2003)

6. Lafon, Y., and Bos, B. Describing and Retrieving Photos Using RDF and HTTP. W3C Note available at: http://www.w3.org/TR/photo-rdf/ (2002)

7. Rui, Y., Huang, T. S., and Chang, S. F. Image Retrieval: Current Techniques, Promising Directions, and Open Issues. Journal of Visual Communication and Image Representation, Volume 10 (1999), pp. 39-62

8. Schreiber, G., Dubbeldam, B., Wielemaker, J., and Wielinga, B. Ontology-Based Photo Annotation. IEEE Intelligent Systems, 16(3) (2001) 66-74.

9. Smith, J. R., and Chang, S. F. An Image and Video Search Engine for the World Wide Web. Proc. SPIE 2670 Storage and Retrieval for Still Image and Video Databases IV, SPIE, Bellingham, Wash., (1996) pp. 84-95.

10. Suh, B., and Bederson, B. Semi-Automatic Image Annotation. University of Maryland Computer Science Department Technical Report, HCIL-2004-15, CS-TR-46 (2004)

11. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, I., Staab, S., and Strintzis, M. G.: "Semantic Annotation of Images and Videos for Multimedia Analysis", 2nd European Semantic Web Conference, 2005.