Generating and Querying Semantic Web Environments for Photo Libraries

Adam Axelrod, Jennifer Golbeck, Ben Shneiderman
Department of Computer Science
Human Computer Interaction Laboratory
University of Maryland, College Park
axelroda@umd.edu, {golbeck, ben}@cs.umd.edu

Abstract

Online photo libraries require a method to efficiently search a collection of photographs, and retrieve photos with similar attributes. Our motivation was to incorporate an existing collection of over 250 photographs of over 200 faculty members and events spanning 7 decades into a library called CS PhotoHistory that is available in hypertext *and* on the Semantic Web. In this paper, we identify challenges related to making this repository available on the Semantic Web, including issues of automation, modeling, and expressivity. Using CS PhotoHistory as a case study, we describe the process of creating an ontology and a querying interface for interacting with a digital photo library on the Semantic Web.

1 Introduction

Most web pages today are built with HTML, successfully allowing browsing and exploration via hyperlinks. This is already useful, but the opportunity for improvement by adding semantic labels is dramatic. As Sean Palmer states, "[t]he problem with the majority of data on the Web that is in this form at the moment is that it is difficult to use on a large scale, because there is no global system for publishing data in such a way as it can be easily processed by anyone" (*The Semantic Web: An Introduction*). This opportunity has lead to the growing popularity of the Semantic Web, which as enabled data to have meaning associated with it that can be machine-processed to provide improved search, superior discovery, and more effective user experiences. As a first step, Semantic Web researchers have developed more robust query languages such as the Resource Description Frameworks (RDF), and the OWL Web Ontology Language. With the growing interest in the Semantic Web, many new approaches and formats have been accepted in order to standardize data-querying methods and enable new capabilities.

One such capability is being able to annotate images and allow the combination of images and text to be queried together. This paper presents the structure of Semantic Web technologies, describes current tools to annotate images which enhance querying, and culminates with the Semantic Web approach used in creating a photographic history of the Department of Computer Science (CS PhotoHistory) at the University of Maryland.

2 Semantic Web Background

The World Wide Web today can be thought of as a collection of distributed, interlinked documents, encoded using (primarily) HTML. Any person can create their own HTML document, put it online, and point to other pages on the Web. Since the content of these pages is written in natural language, computer programs to parse the pages are limited in their efficacy. The Semantic Web makes it possible for machine-readable annotations to be added, linked to each other, and used for organizing and accessing Web content. Thus, the Semantic Web offers new capabilities, made possible by the addition of documents that encode the contents ("knowledge") of a web page, photo, or database, in a publicly accessible, machine readable form. Driving the Semantic Web is the organization of content into specialized vocabularies, called ontologies, which can be used by Web tools to provide new capabilities. Ontologies are encoded in standard languages such as RDF (Resource Description Framework) and OWL (Web Ontology Language), and they contain a framework for describing classes and properties. Those ontologies are then used as a framework for expressing data about instances of the classes. Computers can process ("understand") this information, and load it into a model used in applications.

3 Background and Related Work

Organizing and annotating photographs can be an arduous and tedious task, but without such effort searching and understanding result sets is difficult. Strategies for automating and simplifying annotation are being developed, which could enable users to explore a large collection of photographs in a meaningful and efficient way [Kustanowitz and Shneiderman, 2005]. Commercial tools such as Adobe PhotoShop Album or ACDSee provide limited annotation of whole images, but research tools have begun to enable location based annotation [Shneiderman & Kang 2000]. These processes becomes even more complicated when searching for photographs over the Web, especially when trying to match

photographs to specific criteria (as opposed to simple keywords).

With the presence of the Semantic Web, annotations for photographs have become a leading issue. The development of ontologies, query languages, and tools to interact with photographs has lead to many interesting advances of the Semantic Web.

Schreiber et al. [2001] proposes a framework for setting up ontologies for photographs. To describe the photographs, they suggest a grouping of two types of ontologies: a photo annotation ontology, and a subject matter ontology. The photo annotation ontology is meant to be independent of the subject matter domain. Such information might include how, when, and why the photo was created, and the storage medium for the photo. The subject matter ontology is based off the domain of the photo's subject along with any vocabulary unique to that domain. The structure of the subject matter ontology makes it easily portable and universally usable for other subject matters as it is divided into four distinct parts: *agent* (the subject in the picture), *action* (what the object is doing), *object* (the object being acted on), and *setting* (location description).

Each of these components represents an attribute of the RDF triple used to create an annotation for the photo. Since a photo can be interpreted several different ways, a wildcard can be specified to keep the annotation more general. An annotation that replaces a specific object with a wildcard allows less-specific querying on that photograph. For example an annotation that reads "Jim riding his <wildcard> in the park" allows a user to query without specifying what Jim was riding, and still return the same image. Their work also includes a tool for searching and making queries to a database of photo annotations.

PhotoStuff [Golbeck, et al., 2002] is a tool that allows users to annotate specific areas of photographs. The goal behind this innovation was to allow markup of media as opposed to the traditional approach towards text-only markup. PhotoStuff allows the user to input their own ontology or set of ontologies without any specified format. In addition to adding information about the photo as a whole, specific regions of photographs can be designated and annotated, thus increasing the querying capability of the entire photo. The advantage of querying regions is the ability to return more specific results and support fine-grained queries. However, the process of running PhotoStuff, loading ontologies, and annotating regions is time consuming for our purposes. Because we are using a fixed ontology and making less expressive statements about a photo, a lightweight tool would be more appropriate.

The FotoNotes project considers the image as a collection of objects, each providing different meaning and interpretation. Instead of focusing directly on what attributes a photograph possesses, FotoNotes divides the image into different visible objects, and then assigns an annotation to each object. The different object may include people, stories, data, authorship, non-visual metadata, or other encompassing information about the image. As a whole, these annotations act to tell a story describing a picture. In addition, the separate objects that are devised from the larger photograph help to expand the archive of all objects. In reference to the Semantic Web, FotoNotes allows users to search based on specific objects or annotations rather than a broad range of attributes that may or may not be supported by a given object. The decomposition of objects from the larger source enhances the efficiency of searching semantics [Elin et al., 2004]. FotoNotes provides a stronger relationship between the photograph and annotation than simple attributes due to the smaller granularity of detail that can be expressed.

With the growing popularity of the Semantic Web, design and interaction have become key concerns in the way one uses Semantic Web pages.

One clear example of all these components is CS AKTive Space. Developed by the University of Southampton (UK), CS AKTive Space focuses on Computer Science research in the UK. The data gathering process is continuous using multiple techniques for harvesting and acquisition. Over ten million RDF triples have been derived from a variety of sources including published RDF, personal web pages, and other databases [schraefel et al., 2004].

The project was designed with the focus on the user and in allowing maximum flexibility to browse and search. To allow for exploration of the data, CS AKTive Space uses many formats for the user to query the data. Direct manipulation mechanisms allow the user to search by research area or geographical region. This initial selection reorients the page according to the choice selected resulting in a more focused interaction for the user. A graphical component is also included as another way for the user to interact with the data. Because of the large range of exploration of the semantic data, a visual tool aids in the efficiency of focusing a search. The organization of queries, combined with the capacity to handle multicolumn and geographical data, enhances the tools available to the user during navigation. The ability to create such robust paths allows users to explore extended relationships, thus creating structure to data on the Web.

The W3Photo project [Elin, 2004] is one example of a lightweight photo annotation experiment that takes advantage of work in interaction design. The goal is to create an annotated repository of photos related to the annual World Wide Web conference. Going active at the 2004 conference, users have been able to make quick annotations to photos online using a set of ontologies for representing people, locations, and conference events in an entirely web-based interface. Elements of AKTive Space are also present in this

project, including the AKTivePhoto search that allows users to find pictures by navigating through conference events.

4 Goals of CS PhotoHistory

The CS PhotoHistory project started from the goal to create a chronological history of the Department of Computer Science at the University of Maryland by implementing a photo-based browser. The main driving force towards exploring the Semantic Web is its ability to connect multiple members of a similar community by compiling metadata from a variety of attributes. The ability to create more specific queries using the Semantic Web allows paths between individuals to be recognized that may not have been evident otherwise.

While there are many existing ontologies for faculty domains, we decided to create a new ontology to handle time line features (i.e. events that relate to specific faculty members), and to link photos to corresponding faculty. The CS PhotoHistory project attempts to focus on those attributes that other ontologies have not included. Developing this expanded ontology allows other groups (such as other universities or institutions) to adopt the CS PhotoHistory ontology, and further enhance the value of their own metadata. In the Semantic Web community, FOAF (Friend-Of-A-Friend) is the predominant ontology used to categorize people and their relationships. By linking the CS PhotoHistory ontology with the FOAF ontology, the faculty community at UMD becomes connected with the outside world through research projects, academic history, and other such similar biographical information. Connecting with the outside world allows greater collaboration and communication on a variety of projects. Other research institutions and universities that include their information in FOAF can become connected to the University of Maryland Computer Science community.

On a more local level, this project helps to define the Department of Computer Science at UMD, and allow users (students, other faculty, and outsiders) to explore the achievements of the faculty in the department. Determining which faculty members specialize in a particular research area, finding faculty who have participated on a particular project, and simply exploring the education background of the faculty are compelling reasons to incorporate such a tool in the Semantic Web environment.

The CS PhotoHistory project is also a foundation for examining when and how to annotate data in a focused way. The Semantic Web allows unbounded annotation of resources, but it can be important within a project to limit the expressivity available to users. We will address those issues and use this project to give them context.

5 Challenges of the Semantic Web

When moving to a Semantic Web environment, there are issues relating to expressivity, automation, and modeling. This section identifies some of the most significant challenges that we encountered in the CS PhotoHistory project.

5.1 Determining What to Annotate

When focusing on photographs and graphical features, deciding exactly what to annotate can be thought-provoking. Annotating anything and everything about a photo would produce the most robust database for querying. However, this operation can be quite expensive and potentially unnecessary if the domain of users will only focus on certain aspects of the photo. When starting a database of annotations for a group of images, there could be numerous features that potentially could be included, yet only a subset of annotations provide a sufficient base for user queries. Since one of the main goals of the Semantic Web is to encourage sharing of information and community-wide interaction, it may even be most appropriate to leave the extra annotations to be created by future users. For example, a user may recognize the setting of a particular photo that has not yet been annotated. This user should be allowed to add the annotation to the database in order to enhance this photograph's value. Deciding when to use a particular annotation could depend on how closely the annotation can directly link the images of a photo, or how well the annotation adds to a path from one independent entity to another (such as finding a path by which a picture containing yourself is linked to a picture containing a famous public figure).

One of the most efficient ways to determine which characteristics should be annotated is to conduct usability tests on several typical users to determine what is intriguing about a particular photograph or set of photographs. This provides a solid foundation to determine what features should be required in a querying tool, along with features that unite the set of photographs.

5.2 Extracting Semantics

In developing a photo-based approach to the Semantic Web, it became clear that there are several opportunities to explore the relationships of data and photographs. However, with these opportunities come challenges and decisions about the structure of semantics. The overriding theme is how to best extract semantics from the Web. The modeling process generally requires a human in the loop to produce high-quality results, and thus cannot be fully automated[Berendt et al., 2002]. Clearly, there is always a need for human involvement in determining and organizing the semantics. Yet, the decision of where to draw the line between computer automation and human interference is still a great challenge to Semantic Web

developers. Deciding how strong of a relationship an object is to a subject can hinder the effectiveness of a user query. Once the first step of determining the many sources of data (web pages, databases, existing text, etc…) is complete, the mapping onto an ontology requires a way for the machine to recognize the semantics and efficiently extract them. This process must involve keeping a general approach so that all the data can be incorporated, yet still be able to allow the strongest attributes to be included.

5.3 Performance against Existing Search Technologies

With regards to performance, the Semantic Web aims to be faster and more accurate in querying then similar search engines. The product created by Schreiber et al. was analyzed against existing search engines (including Alta Vista) and they concluded that "search technology on the Web is either too specific (low recall, reasonable precision) or too general (high recall, low precision)," demonstrating that using more specific terms in a query causes recall performance to decrease [Schreiber et al., 2001]. This study showed that most search engines do not distinguish similar keywords (i.e. ape vs. gorilla), and also do not filter some irrelevant data such as personal home pages and out-of-context results. When searching for photographs based on the same search criteria, this study produced similar results. Searching for "an ape scratching his head" would result in very few photographs, and may not match the criteria being searched. As a result, this demonstrates the need for the Semantic Web, in addition to the challenge Semantic Web developers face with when trying to separate new tools from existing search engine technology.

One related issue to the effectiveness of the Semantic Web is the richness and amount of data that can be collected and queried. Many argue that the amount of data should be unbounded and open to all portals. However, the structure of a Semantic Web collection is determined by the format of the data, and how well the metadata can interact. In order for paths between two entities to be related, the ontology must support a related attribute in both. If the data does not fall within the ontology, then it is increasingly difficult to create links between data sources. As seen with CS PhotoHistory, the amount of data will dictate how successful it is in linking to outside domains, and in being able to extend a relatively new ontology.

6 Implementation

6.1 Creating the Data Models

In order to organize the photographs that were collected, we categorized the features needed to store a picture. Because the domain of photographs represents a wide range of activities and time periods, we allowed users to

add a photographer name and a caption to each photo. Since the subject matter of the photographs was restricted to the domain of faculty and staff members, we required the following attributes for faculty members: *First, Last name*, *Start, End Years* as members of the department, *Field Committee* (research area), *Short biographical information* (including academic history and career achievements)

These two domains served as the driving model towards a timeline of photographs (Figure 1) and events, along with a complete index of all faculty members since the department's inception. Each member listed in the faculty index has his or her own page devoted to displaying all photographs and biographical information.

Figure 1: This depicts the initial design of the CS PhotoHistory photo libraries as a timeline of events and photographs of the Department of Computer Science at the University of Maryland. The hyperlinked faculty names lead to a page detailing information about that faculty member including field committee, dates, awards, and other biographical information.

To make this information available on the semantic web, we constructed an ontology that represented the information and relationships contained in the original database. Because the database for the project was specifically designed to support certain types of interaction in the interface, we felt that our best option was to develop a new ontology with terms that we could later map to classes and properties in existing ontologies, such as FOAF. The ontology comprises two classes: Person and Image. We created Datatype Properties for each of the name attributes, year attributes, biography, and field committee with a domain of the Person class and photographer and date with a domain of the Image class. Finally, we created Object Properties to connect instances of the Person class to instances of the Image class. This facilitates browsing the data by co-depiction; users can move from a person, to a photo, to another person who is in the same photo.
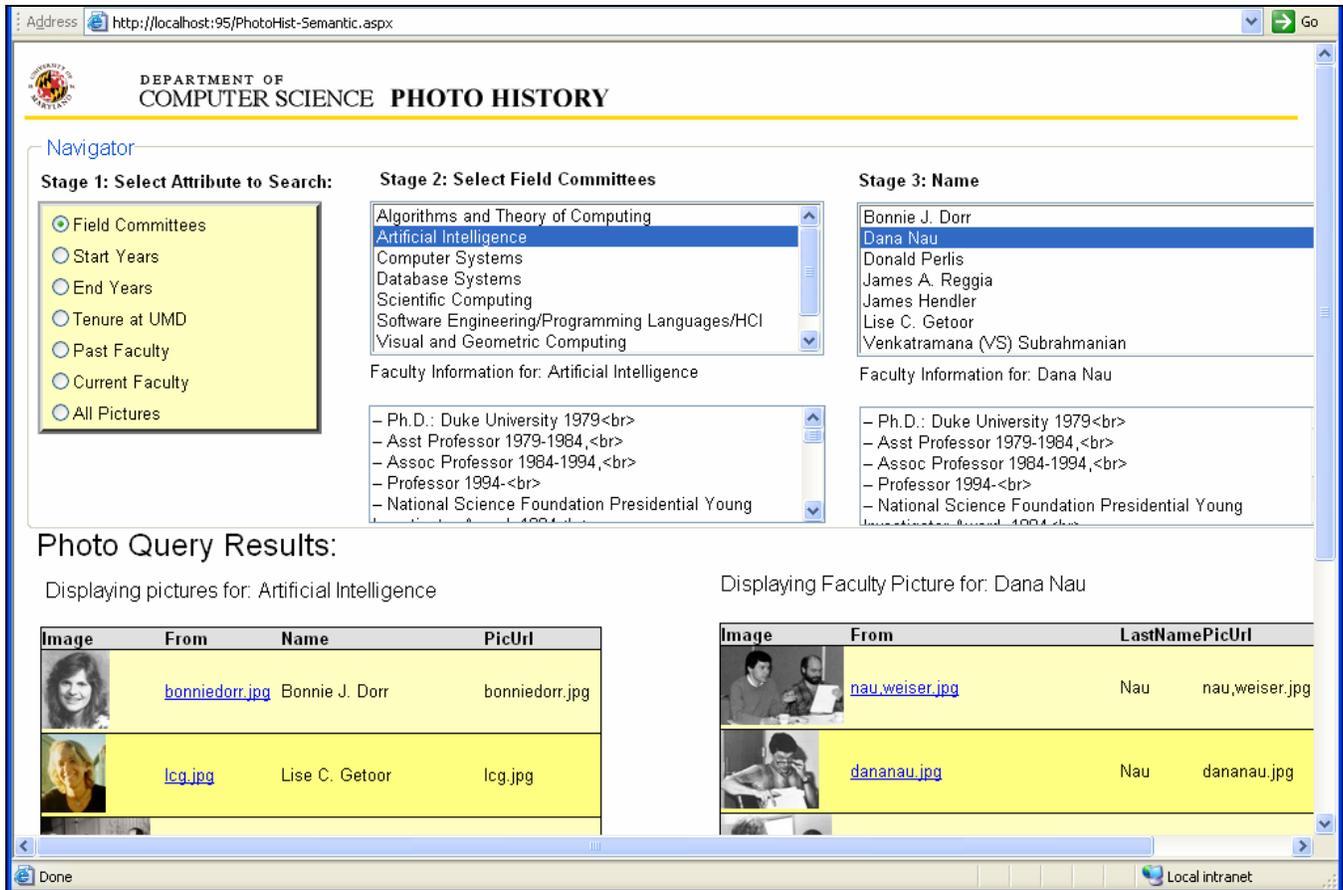
Figure 2: This figure shows the three stage process for querying the CS PhotoHistory libraries. The user chooses the initial attribute to query, then the browser fine-grains the search to all entities lying within that domain. Upon the second user selection, the browser retrieves a third-level list of results. Below the 'Navigator' are the actual photo results of the query listed alphabetically by last name.

6.2 Interface Design

Once users select the initial attribute, the query bar becomes more fine-grained and displays the results of the initial query. User choose one of the results and display the photographs related to that query. They also has the option to sharpen their search one more level with finer-grained attributes (Figure 2). In addition to displaying the photographs for the second query, the system expands the second query to retrieve all the faculty information related to the second query. For example, if users initially selects 'Field Committee', then select 'Database Systems', the third query field will display the faculty information of all faculty in 'Database Systems'. Users can then display the photographs related to those faculty members. Figure 2 shows the query interface for the CS PhotoHistory Data. It allows querying based on select attributes such as by Field Committee, Start Year, End Year, Tenure, Current Faculty, Past Faculty, and by Picture.

knowledge bases and methods for interacting with that data.

In order to increase the extensibility of the CS PhotoHistory project, the next stage involves further categorizing the data and adding photographs. Adding more attributes to the faculty data expands the value of the data for querying, and gives more detailed meaning to the photographs. With more attributes, more complex queries can be drawn. These queries help to distinguish more relationships between the data and photographs, and

7 Conclusion and Future Work

The CS PhotoHistory project is a case study of creating photo libraries on the Semantic Web. We addressed issues of deciding what to annotate, how to extract semantics and build models, and creating an interface for interacting with the modeled data. Our hope is that this project will serve as an introductory example for other digital libraries being moved to the Semantic Web, and help guide decisions about creating ontologies to represent existing

allow many more users to define unique paths between objects.

One of the more interesting possibilities is to link the local domain of photographs to a more global domain, such as all other RDF data on the web relating to other universities. By having a framework to collect the photographs of similar domains, linking the data will allow greater expansion and visibility. With the development of the Semantic Web, this would be an intriguing way to show the power that the Semantic Web has in creating communication paths between unique groups.

With several different attributes represented in the data and photographs, it is important to also consider the types of queries that different users request. While this project has focused on including the most important data, it should be further examined to determine how users interpret the different photo results from different queries. This will allow a stronger focus on how to structure different levels of queries after each initial query. Understanding how users interpret photographs also aids adding efficiency in associating regions of photographs with particular attributes.

Acknowledgements

REFERENCES

Berendt, Bettina et al. "A Roadmap for Web Mining: From Web to Semantic Web." *First European Web Mining Forum, EWMF* September 2003.

Berendt, Bettina, Andreas Hotho, and Gerd Stumme. "Towards Semantic Web Mining." *First International Semantic Web Conference 2002*. Sardinia, Italy, pp 264-272.

Elin, Greg. "Is a Picture Worth a Thousand Clicks? Challenges of Adding Semantic Data to Images" *Proceedings of the First International Workshop on Interaction Design and the Semantic Web*, May 18, 2004. New York, New York.

Elin, Greg, Marc Rohlfing, and Michael Parenti. "FotoNotes." <http://fotonotes.net>. 2004.

Golbeck, Jennifer, Michael Grove, Bijan Parsia, Adtiya Kalyanpur, and James Hendler, "New Tools for the Semantic Web", *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2002)*, October 1-4, 2002, Siguenza, Spain.

Kustanowitz, J. and Shneiderman, B., "Motivating Annotation for Personal Digital Photo Libraries: Lowering Barriers While Raising Incentives", Univ. of Maryland Technical Report HCIL-2004-18, January 2005.

Lassilla, Ora, Ralph R. Swick, "Resource Description Framework (RDF) Model and Syntax Specification." <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>. February 1999.

Manola, Frank, Eric Miller, and Brian McBride. "RDF Primer." <http://www.w3.org/TR/rdf-primer/>. February 2004.

May, Wolfgang, José Júlio Alferes, and François Bry. "Towards Generic Query, Update, and Event Languages for the Semantic Web." *PPSWR* 2004: 19-33. 2004.

Robie, Jonathan et al. "The Syntactic Web: Syntax and Semantics on the Web." *Markup Languages: Theory & Practice 3.4* 2001: 411-440.

schraefel, m.c., Nigel R. Shadbolt, Nicholas Gibbins, Hugh Glasser, and Stephen Harris. "CS AKTive Space: Representing Computer Science in the Semantic Web." *WWW2004*: 384-392.May 17, 2004.

Schreiber, A. Th., Barbara Dubbeldam, Jan Wielemaker, and Bob Wielinga. "Ontology-Based Photo annotation." *IEEE* 1094-7167/01. 2001.

Shneiderman, B. and Kang, H., Direct annotation: A drag-and-drop strategy for labeling photos, *Proc. International Conference on Information Visualization 2000*, IEEE Press, Los Alamitos, CA (July 2000), 88-95.

Tammet, Tanel, and Vello Kadarpik. "Combining an Interface Engine with Databases: A Rule Server." *RuleML* 2003: 136-149. 2003.