

# Image Browsing and Natural Language Paraphrases of Semantic Web Annotations

Christian Halaschek-Wiener, Jennifer Golbeck, Bijan Parsia,  
Vladimir Kolovski, and Jim Hendler

Maryland Information and Network Dynamics Lab, 8400 Baltimore Ave.  
College Park, MD, 20740 USA  
{halasche, golbeck, kolovski, hendler}@cs.umd.edu  
bparsia@isr.umd.edu

**Abstract.** Recently, there has been interest in marking up digital images with annotations describing the content of the images using Web based ontologies encoded in the W3C's Web Ontology Language, OWL. The annotations are subsequently exploited to improve the user experience of large collections of images, whether by enhance search or by a structured browsing experience. In the latter case, the complexity and unfamiliarity of logic-based ontology languages may do more to impede, than aid, the user. To alleviate this problem, we propose using automatic generation of natural language (NL) paraphrases of OWL statements to assist browsing image content. In this paper, we provide an overview of our NL generation approach and an empirical evaluation of the use of our paraphrases for image browsing.

## 1 Introduction

Recently there has been interest in using Semantic Web ontologies encoded in the Web Ontology Language, OWL, to formally represent the semantic content of digital images [1]. Given semantically rich image metadata, collections can be more accurately searched and browsed, with new knowledge derived from existing annotations.

The canonical OWL interchange syntax is based on the XML serialization of RDF (RDF/XML). Neither RDF nor XML were designed readability in mind (much less casual end user readability). There are several alternative surface syntaxes designed with more readability in mind and, due to the correspondence with first order logic, there are a number of traditional logic notations available. However, these all require a fairly deep understanding of OWL, logic, or both which, aside from being comparatively rare, seems unnecessary for the purpose of navigating through image collections. Recently, there has been work in providing natural language translations of OWL concept definitions, providing users with a format that is easier to both read and understand [3].

In this paper, we propose the use of automatically generated NL paraphrase of OWL classes and individuals to aid in browsing image. We feel that the paraphrases will make allow users to more effectively use and enjoying using the semantic annotations when browsing image collections, and to better understand the content of the

browsed images. Most semantic markup of images includes a large number of instance assertions, corresponding to features of concrete particular depicted in the images. Therefore, we extend the algorithm in [3] to provide natural language paraphrases of OWL assertions about individuals, as well as of class definitions. We have implemented this algorithm as a plug-in to an ontology-based, image annotation and browsing tool, PhotoStuff [1]. Lastly, we provide an empirical evaluation of the effectiveness of these NL paraphrases for the task of browsing and interacting with image collections.

### 3 OWL Individual NL Paraphrases

It is clear that images can be categorized based on the types of things they depict. In the SemSpace<sup>1</sup> portal, a nested list rendering of the class hierarchy is the initial interface to the collection. Thus, NL paraphrases of concepts might help users determine which categories are likely to contain images of interests. [3] presents an approach for automatically generating natural language paraphrases for OWL concept definitions that attempts to preserve the underlying meaning of the logic based definition. The approach is applicable to any ontology where the classes and properties are named using certain standard naming conventions. The algorithm is fairly straightforward considering the good quality results achieved. The most sophisticated NL processing tool the algorithm uses is a part-of-speech (POS) tagger, which is used to improve the fluency of the generated NL description. Due to space limitations, further details are omitted, however they can be found in [2,3].

Images however, especially photographs, generally depict concrete objects that are naturally represented in OWL as individuals with various types and properties. Thus, if paraphrases are to help the user understand the contents of particular images, the algorithm in [3] must be extended to handle NL paraphrases of OWL individuals as well.

#### 3.1 Approach Overview

The approach adopted here provides an NL rendering of OWL individuals based on the direct relations of that individual (e.g., type assertions, labels, defined relations, etc.). In order to generate an NL paraphrase for an OWL individual, first a NL parse tree is generated for that individual [3]. In contrast to the approach for OWL concepts [3], this tree will be at most one level deep, as only the direct relations are used. As the tree is created, only *rdf:type* assertions corresponding to the (possibly inferred) most specific classes of the individual are added as edges. (In certain ontologies with shallow, informative class graphs, it might be preferable to add all the types of the individuals, or to control the depth in a different way.) The relation edges are labeled by two distinct mechanisms: An *rdf:type* edge in the tree is given an “*is a*” label. For all additional relations, the POS tagger is run over the “local part” of the property’s

---

<sup>1</sup> <http://semspace.mindswap.org/>

URI (as in [3]) to provide more legible labels for the relations. Finally, the labels for the objects of the relations are added to the object nodes.

After this post-processing the NL rendering is generated. The current approach uses the following template for producing OWL individual NL renderings: first the label edges of the NL tree for the given individual are retrieved. This label is used to begin the NL sentence. Following this, a comma separated rendering of all *rdf:type* edges is generated. Then a bullet list of all the additional relations is produced. The template is shown below in Table 1. Additionally, the resulting NL rendering for the OWL individual *Storey Musgrave* is presented below in Table 1 (note that the original RDF/XML for Storey Musgrave is provided in [2]).

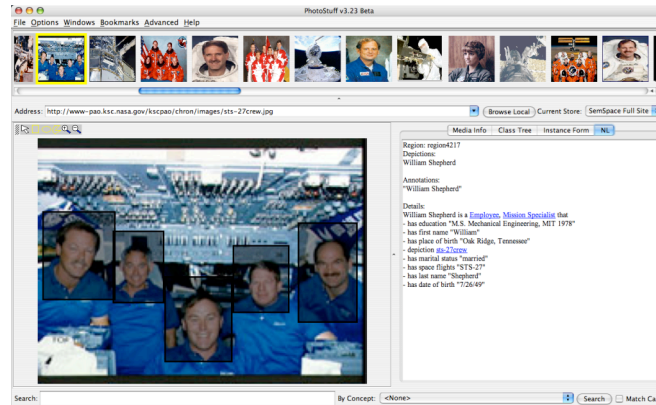
**Table 1.** NL rendering template. Italicized words are invariant of the individual being rendered. Additionally, automatically generated natural language rendering of Storey Musgrave

Individual <i>is a rdf:type that</i>	NL Rendering
<ul style="list-style-type: none"> <li>- relation 1</li> <li>- relation 2</li> <li>- ...</li> <li>- relation n</li> </ul>	<p>Storey Musgrave is a Payload Commander, Employee that</p> <ul style="list-style-type: none"> <li>- has first name "Storey"</li> <li>- has last name "Musgrave"</li> <li>- depiction musgrave_sts61, STS033, musgrave , sts-61crew, STS061_storeyL, STS61_storey, musgrave WFPC</li> <li>- has space flights "sts-61", "sts-33"</li> <li>- has date of birth "08/19/1935"</li> </ul>

## 4 NL Captions for Browsing Image Collections

In this section we provide details of using automatically generated NL paraphrases for browsing image collections in our image annotation tool, PhotoStuff [1]. The main motivation for using NL paraphrases for browsing image collections and their annotations is that it will provide a more enjoyable user experience. Essentially, the user will be provided with the details about image contents in a human readable and understandable format, which allows them to focus on the images and what they depict rather than the cryptic details of the representation language. One additional benefit of publishing the natural language rendering of the images is that it can allow exiting search engine technology to index such content.

An experimental implementation of our NL generation algorithm has been provided as a plug-in for PhotoStuff, an ontology-based, image annotation tool [1]. Within PhotoStuff, the NL captions are used in two main ways. First, when an image is selected and loaded into the image canvas, NL renderings for the instances depicted within regions are provided as pop-ups when the regions are moused over. This provides the user with a quick, human readable display of the individual depicted within that region.



**Fig. 1.** Region NL Caption Pop-Up

Additionally, the NL for individuals depicted within regions can be viewed by right-clicking the regions, and selecting “View NL”. This puts the natural language paraphrases in a NL info pane, as shown above in Figure 1. The figure illustrates that in the NL info pane, the generated natural language includes hyperlinks for all additional individuals and OWL class that occur in the rendering of the current individual or class. This allows the user to browse the image collection based on the existing annotations in a Web-like manner. Further, the hyperlinks can be right clicked, providing the option to filter the thumbnail strip to only show images that depict that instance or instances of that class.

## 5 Evaluation

To evaluate the benefits of the NL format, we conducted a pilot user study. Our hypothesis was that users would prefer the NL format for viewing data when the task was to gain an understanding of the meaning of a concept. Our pilot study included seven subjects. Subjects ranged in age from 20 to 37 and all were students working toward bachelors, masters, or Ph.D. degrees. We tested the subjects’ preferences for method of viewing classes separately from viewing instances. For both classes and instances, we choose three examples: one very simple, one of medium complexity, and one very complicated example.

When viewing classes, we used the Wine<sup>2</sup> ontology. Our simple class is AlsatianWine, which only had one restriction. Anjou was the medium complexity class, with four simple restrictions and an intersection combined with a restriction. The most complex class was Beaujolais, with six simple restrictions, including cardinality, and an intersection combined with a restriction. Each class was presented to the subjects in SWOOP [4]. For the study, four additional formats for viewing each class was provided in SWOOP, including concise format, Turtle, Abstract Syntax, and RDF/XML. Users were asked to view the concept with the goal of understanding it’s

<sup>2</sup> Wine Ontology: <http://www.w3.org/2001/sw/WebOnt/guide-src/wine.rdf>

meaning, rather than the modeling features. They were allowed to take as much time as they needed and, when finished, they were asked to rank the formats according to their preference, with 1 being best and 5 being worst. If the subjects felt two formats were equally good, they were allowed to give them the same ranking. For all three classes, the order of the rankings was the same. From best to worst, the formats were ranked, and shown below in Table 5. Additionally, the average ranking and standard deviations for the various formats is presented below in Table 6.

**Table 5.** Rankings from best to worst of view format of class descriptions

1. Natural Language
2. Concise Format
3. Abstract Syntax
4. Turtle
5. RDF/XML

**Table 6.** Average rankings and standard deviations for the five formats used to display class information

Class	Average Rank (Standard Deviation)				
	Concise Format	Abstract Syntax	Natural Language	RDF/XML	Turtle
AlsatianWine	2.00(1.41)	3.57(1.13)	1.57(0.79)	4.86(0.38)	3.86(0.69)
Anjou	1.86(0.69)	2.71(0.49)	1.14(0.38)	5.00(0.00)	4.00(0.00)
Beaujolais	2.14(0.69)	2.71(0.49)	1.00(0.00)	5.00(0.00)	3.86(0.38)

Using a Wilcoxon matched-pairs signed-ranks test, we found that the NL format significantly outperformed the Concise Format (ranked second on average) for both Anjou and Beaujolais with  $p < 0.05$ . There was not a significant benefit over the Concise format for the simplest class, AlsatianWine, but NL did significantly outperform the Abstract Syntax, which was the average third ranked format. This allows us to conclude that in the pilot study the NL format offers significant benefits to users when they are trying to understand the meaning of classes, particularly complex classes.

To evaluate the NL format with instances, subjects were given three instances from the collection of astronaut data<sup>3</sup>. The first instance, Bryan O'Conner, had only one property: the depiction that tied him to an image. The second instance, John Young, had a depiction and one additional property. The last instance, Storey Musgrave, was tied to three regions and four properties, including both Datatype and Object properties. Subjects viewed the data about each instance in a tabular format and in the NL format with the goal of understanding the information about each instance. Subjects took the time they needed without limits and then were asked to choose which format they preferred. For each of the three instances, the NL format was preferred 6:1 over the instance form. When subjects commented on their preference, it focused largely on the fact that the NL format was more concise, displaying only relevant information. We conclude this section by stating that the results of this pilot study show that users feel that the NL format provides them with an advantage when trying to understand both classes and instances, thereby suggesting that it may help in understanding and browsing image content.

<sup>3</sup> SemSpace portal instance data. Available at <http://semSpace.mindswap.org/rdf/dump>

## 6 Discussion and Conclusion

While our initial evaluations support the usefulness of the approach presented here, we note there are a few limitations. First, if individuals participate in a large number of relations, then the generated bullet list, or concatenated sentence, can be quite long and cumbersome to read through. One potential solution to ease this problem is to use a ranking metric when displaying the relations in the NL generated format. Additionally, we would like to investigate utilizing text summarization techniques for the output from the NL generation algorithm. In this work, we use a predefined template view of the NL rendering of instances. Under some circumstances, particular users may wish to define their own templates. Thus we would like to support using custom templates for particular ontologies or class descriptions. Lastly, we wish to further evaluate the extent of which these paraphrases work to improve the overall image browsing experience.

In conclusion, we have provided an extension of our previous work [3] to automatically generate natural language paraphrases for OWL individuals. Given this, we have proposed using this approach for browsing Semantic Web annotations of image collections. We have additionally presented an experimental implementation of the approach in an ontology based, image annotation tool, PhotoStuff. Finally, we have provided an empirical evaluation of the usage of the NL paraphrases for browsing image collections, finding that it was in fact useful for the task.

This work was supported in part by grants from Fujitsu, Lockheed Martin, NTT Corp., Kevric Corp., SAIC, the National Science Foundation, the National Geospatial-Intelligence Agency, DARPA, US Army Research Laboratory, and NIST. We would also like to thank Aditya Kalyanpur and Mike Grove for all of their contributions to this work.

## References

1. Halaschek-Wiener, C., Schain, A., Golbeck, J., Grove, M., Parsia, B., Hendler, J.: A Flexible Approach for Managing Digital Images on the Semantic Web, 5th International Workshop on Knowledge Markup and Semantic Annotation, Galway, Ireland, November 7, 2005.
2. Halaschek-Wiener, Golbeck, J., Grove, Kolovski, V., Parsia, B., Hendler, J.: Image Browsing and Natural Language Paraphrases of Semantic Web Annotations, UMIACS Technical Report 2006. Available at <http://www.mindswap.org/papers/2006/NLP-Images-TR.pdf>.
3. Hewlett, D., Kalyanpur, A., Kolovski, V., Halaschek-Wiener, C.: Effective Natural Language Paraphrasing of Ontologies on the Semantic Web. End User Semantic Web Interaction Workshop, International Semantic Web Conference (ISWC), November 2005, Galway, Ireland.
4. Kalyanpur, A., Parsia, B., Sirin, E., Cuenca-Grau, B., Hendler, J. Swoop - a web ontology editing browser. *Journal of Web Semantics*, 4(1), 2005.