Semantic Annotation of Image Collections

Laura Hollink[1] Guus Schreiber[1] Jan Wielemaker[2]
Bob Wielinga[2]

[1]Free University Amsterdam, Computer Science
E-mail: {laurah,schreiber}@cs.vu.nl

[2]University of Amsterdam, Social Science Informatics
E-mail: {jan,wielinga}@swi.psy.uva.nl

ABSTRACT

In this paper we discuss a tool for semantic annotation and search in a collection of art images. Multiple existing ontologies are used to support this process, including the Art and Architecture Thesaurus, WordNet, ULAN and Iconclass. We discuss knowledge-engineering aspect such as the annotation structure and links between the ontologies. The annotation and search process is illustrated with an application scenario.

1. INTRODUCTION AND APPROACH

In this paper we show how ontologies can be used to support annotation and search in image collections. Many of such collections currently exist and users are increasingly faced with problems of finding a suitable (set of) image(s) for a particular purpose. Each collection usually has its own (semi-)structured indexing scheme that typically supports a keyword-type search. However, finding the right image is often still problematic.

Figure 1 shows the general architecture we used within this study. For this study we used four ontologies (AAT, WordNet, ULAN, Iconclass) which were represented in RDF Schema [1]. The resulting RDF Schema files are read into the tool with help of the SWI-Prolog RDF parser [19, 10]. The tool subsequently generates a user interface for annotation and search based on the RDF Schema specification. The tool supports loading images and image collections, creating annotations, storing annotations in a RDF file, and two types of image search facilities.

The ontologies, the annotation template and their interrelations are discussed in Section 2. The annotation and query process is discussed in, Section 3 in the form of an application scenario. Section 5 discusses related work. Finally, Section 5 provides a general discussion on the approach taken.

This work is a sequel to earlier work on semantic annotation and search of a collection of photographs of apes [13]. In the earlier study the emphasis was mainly on the subject-matter of the image. For art images both the image subject and the art-historic features, such as artist and style, are important. This requires the use of additional ontologies (AAT, ULAN) and poses research questions with respect to the links between ontologies (see Section 2.4).

2. ONTOLOGIES, ANNOTATION TEMPLATE AND THEIR INTERRELATIONS

2.1 Ontologies

For this study we used four thesauri, which are relevant for the art-image domain:

1. The Art and Architecture Thesaurus (AAT) [11] is a large thesaurus containing some 125,000 terms relevant for the art domain. The terms are organized in a single hierarchy.
2. WordNet [8] is a general lexical database in which nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept. WordNet concepts (i.e. "synsets") are typically used to describe the content of the image. In this study we used WordNet version 1.5, limited to hyponym relations.
3. Iconclass [16, 15] is an iconographic classification system, providing a hierarchally organized set of concepts for describing the content of visual resources. We used a subset of Iconclass.
4. The Union list of Artist Names (ULAN [2] contains information about around 220,000 artists. The information includes name variants and some limited biographical information (dates, locations, artist type). A subset of 30,000 artists, representing painters, is incorporated in the tool.

AAT, WordNet, Iconclass and ULAN were all translated into the RDF Schema notation. For example, WordNet was represented in the following fashion:

- WordNet concepts ("synsets" which have a numerical identifier) were represented as RDFS classes;
- word forms of concepts were represented as RDFS labels of the corresponding class;
- hyponym relations were represented as RDFS subclass relations;
- glossary entries of concepts were represented as RDFS comments.

In another paper [20] we discuss how we can use WordNet 1.6 as represented by Melnik and Decker[1]. In a prior publication [21] one can find a discussion on issues arising when

[1]See http://www.semanticweb.org/library/#wordnet

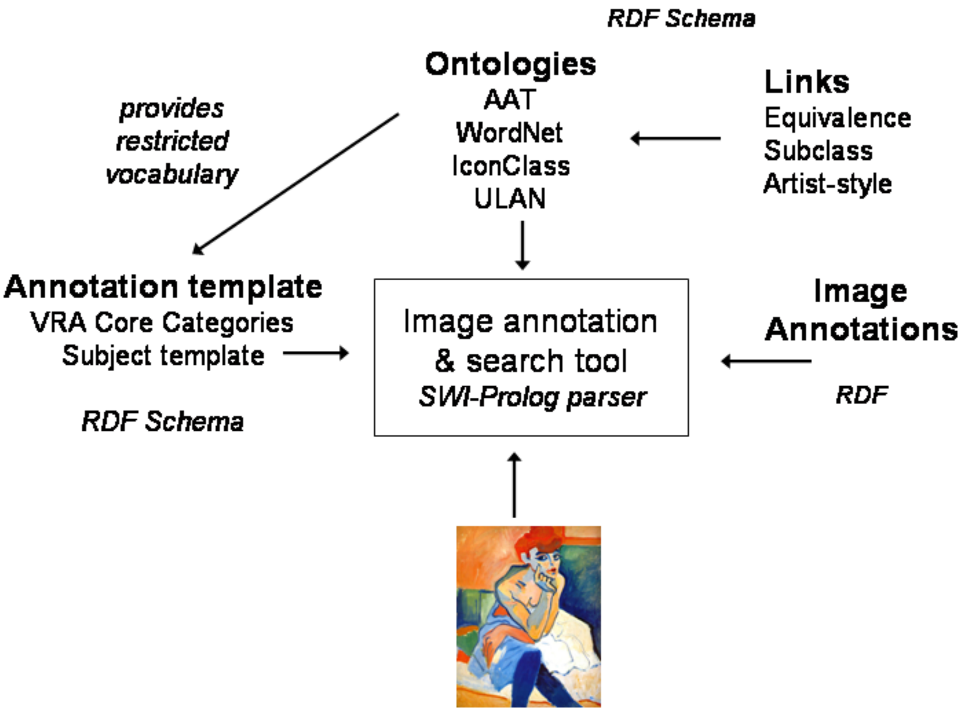

Figure 1: Overview of the approach in this study. The RDF Schema specifications of the ontologies, of the ontology links and of the annotation template are parsed by the SWI-Prolog RDF parser into the tool. The tool generates an annotation and search interface from these specifications. This interface is used to annotate and query images. The annotations are stored in an RDF file

Source	Triples
WordNet 1.5 (limited to hyponym relations)	280.558
Art and Architecture Thesaurus	179.410
Iconclass (partial)	15.452
ULAN (limited to painters)	100.607
Total	576.027

Table 1: Number of RDF triples in the four ontologies

representing AAT in RDF Schema.

Table 1 shows the number of RDF triples in the tool for each of the thesauri. The infrastructure of our current tool can handle this set of 576,000 triples efficiently, but it is expected to break down when the triple base becomes significantly larger. Based on our experiences in this work we have recently constructed a revised infrastructure that should be able to handle up to 40,000,000 triples [20].

2.2 Annotation template

For annotation and search purposes the tool provides the user with a description template derived from the VRA 3.0 Core Categories [17]. The VRA template is defined as a specialization of the Dublin Core set of metadata elements, tailored to the needs of art images. The VRA Core Categories follow the "dumb-down" principle, i.e., a tool can interpret the VRA data elements as Dublin Core data elements.[2]

[2]An unofficial OWL specification of the VRA elements, including links to Dublin Core, can be found at http://www.cs.vu.nl/~guus/public/vra.owl

The 17 VRA data elements were for visualization purposes grouped into three sets:

Production-related descriptors: title, creator, date, style/period, technique, culture and and relation. .

Physical descriptors: materials.medium, materials.support, measurements, type and record type.

Administrative descriptors: location, collection ID, source and rights.

Two VRA data elements are not included in the template: description and subject. Both are used to describe the content of the image. As we were interested in providing a more structured content description we used an adapted version of the "sentence structure" proposed by Tam [14] as a means of structuring image-subject descriptions. The subject of the image is described with a collection of statements of the form "*agent action object recipient*". Each statement should at least have an agent (e.g. a portrait) or an object (e.g. a still life). The terms used in the sentences are selected from terms in the various thesauri. Multiple sentences may be used to describe a single painting.

For example, the painting by Chagall in Figure 2, in which Chagall kisses his wive and gets flowers from her, can be described with the following two statements (source of the term parenthesized):

Figure 2: Painting of Chagall

```
Agent:     "Chagall, Marc"   (ULAN)
Action:    "kiss"            (WordNet)
Recipient: "wives"           (AAT)

Agent:     "woman"           (WordNet)
Action:    "give"            (WordNet)
Object:    "flower"          (WordNet)
Recipient: "Chagall, Marc"   (ULAN)
```

The scheme was developed for a previous experiment [13]. It avoids the problems of parsing natural language descriptions, while maintaining some of the naturalness[3] and richness. Note that the use of such concepts to describe the image allows one to do semantic matching during search. For example, one can find this picture when searching for a picture using a synonym or hypernym (e.g., "touch" instead of "kiss"). The application scenario in Section 3 gives an example of the use of this template.

In addition, one can describes the "setting", i.e., characteristics of the scene as a whole. We use three slots to describe the setting: event, place and time. These three slots are also filled with terms from the thesauri. For example, the painting by Chagall can be described with the event `birthday celebration` (concept from Iconclass) and the location `artist's workplace` (concept from WordNet).

The tool also provides a free text field, where information can be stored that doesn't fit into one of the slots, or is not present in any of the ontologies.

2.3 Linking the annotation template to the ontologies

Where possible, a slot in the annotation template is bound to one or more relevant subtrees of the ontologies. For example, the VRA slot style/period is bound to two subtrees in AAT containing the appropriate style and period concepts. The following VRA data elements are currently linked to parts of AAT: technique, style/period, type, record type, material.support, material.medium and culture. One VRA data element is linked to ULAN, namely creator.

The slots of the subject-matter description are also linked to subtrees of the ontologies. WordNet provides many general concepts for subject-matter description. AAT also provides some concepts useful for this purpose. There is some overlap here between AAT and WordNet. In the next subsection we come back to this issue.

Iconclass is particularly useful for describing scenes as a whole (cf. the `birthday celebration` example earlier). ULAN contains specific persons, which are typically used to annotate images in which artists themselves are depicted (e.g., a self portrait). We are currently considering to include also some geographical terminology base, such as the Thesaurus of Geographical Names (TGN)[4], to be able to describe specific locations in a semantically meaningful way.

2.4 Links between ontologies

The four ontologies contain many terms that are in some way related. For example, WordNet contains the concept `wife`, which is in fact equal to the AAT concept `wives` (AAT uses the plural form as the preferred one). One could consider to design a new ontology by merging them. However, to make the Semantic Web work, we will need to *reuse* existing ontologies rather than redoing them. Thus, we decided to use the ontologies "as-is" and create separate corpora of ontology links. We added three types of ontology links. Equivalence relations and subclass relations are often mentioned in the literature as useful link primitives (e.g. [9]). In addition, we added links specific for the art-image domain.

2.4.1 Equivalence links

We added equivalence relations between terms appearing in multiple ontologies that refer to the same concept. For example, the `artistic movements` branch in WordNet is linked to the equivalent `styles and periods` subtree is AAT. Similarly, the WordNet concept `wife` is linked to the AAT concept `wives`.

As RDF Schema does not provide an equivalence relation[5], we had to introduce our own special-purpose property for this. In forthcoming versions of the tool this relation will be replaced by the OWL language construct **owl:equivalentClass** [18].

2.4.2 Subclass links

When differences in the structure of the ontology are large (a common feature), equivalence relations are sometimes only possible at the lowest, most specific branches of the hierarchies. We use the RDFS subclass relation to create links at a higher level in the hierarchies. Consider the example in Figure 3 which show two subtrees of respectively AAT and WordNet. One can see that the term `artist` in WordNet, does not refer to the same concept as `artist` in the AAT, since some subconcepts of `artist` in WordNet, such as `musician`, are not subconcepts of `artists` in AAT, which contains only `people in the visual arts`. To link WordNet to AAT we need to create a subclass link: AAT `artist` is a subclass of WordNet `artist`.

[3]The naturalness is limited, see. the term "wives" in the first statement. This is because AAT uses the plural form for concepts.

[4]http://www.getty.edu/research/tools/vocabulary/tgn/

[5]The revised version of RDF Schema allows cycles of subclass relations. This means that one can now represent equivalence of A and B by stating the A is a subclass of B and that B is a subclass of A.

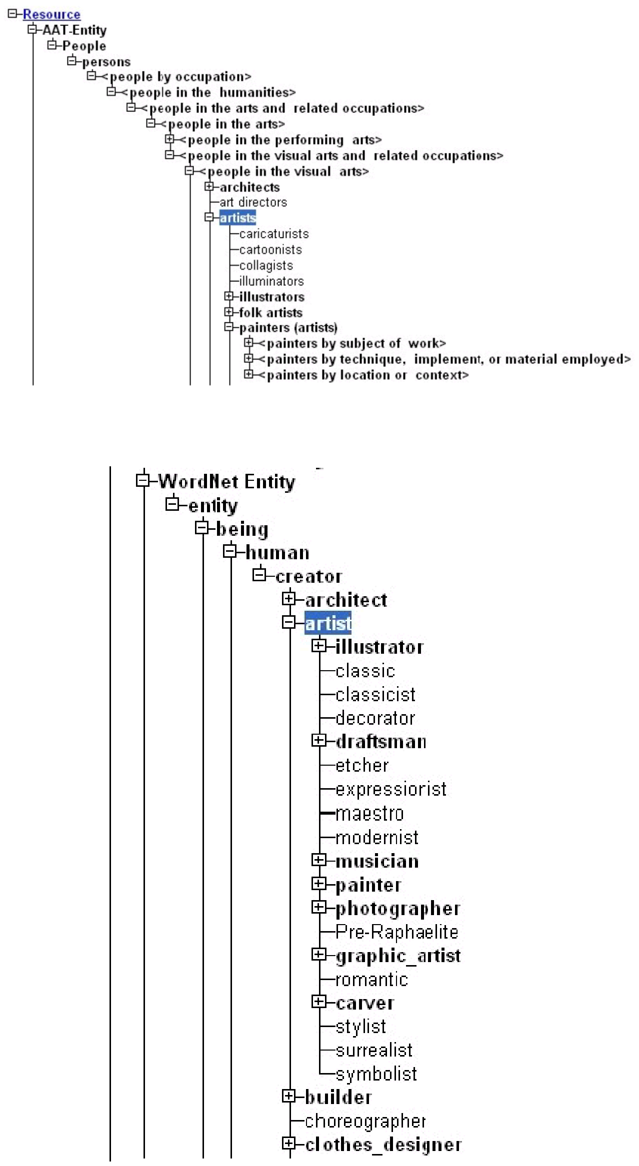

Figure 3: Subtrees of AAT (above) and WordNet (below) in which the concept `artist` appears. The figures are snapshots of the RDF browser of our tool

2.4.3 Domain-specific links

In addition to equivalence and subclass links, we also use *domain specific* relations. For example, by linking painting techniques to materials, we were able to derive the value of the technique slot from the values of the material.support and material.medium slots. Similarly, a link between artists in ULAN and painting styles in AAT, made it possible to suggest to the user the value of the style/period slot once the creator was known. In this way, Picasso is linked to cubism, Matisse is linked to Fauve, Van Gogh to impressionism, and so on. This relation is many-to-many: a artist may belong to multiple styles.

Other derivations are possible, but are not yet supported by the tool. ULAN contains information about the country of origin of the artists. This means that the VRA slot culture could in principle be derived from the slot creator. The type of a painting can sometimes be derived from descriptions of the content. If the only description of a painting is an agent, the painting is probably a portrait. If the agent is equal to the creator, we are looking at a self-portrait. The suggested values act as default values and can be overridden by the annotator.

2.5 Using the links

Equivalence and subclass relations increase the recall of the tool. They make it possible to retrieve images annotated with concepts from one ontology while searching with concepts from another ontology. Domain-specific relations are especially useful for annotation. Values in the annotation that are suggested by the tool reduce the time and effort spend by the human annotator. Domain-specific relations can also be used to improve search. For example, if a user is searching for Fauvist paintings, the tool can retrieve paintings by Matisse, Derain and De Vlaminck, all Fauvist painters. Domain-specific relations like artist-style. have to be interpreted by the annotation and search algorithms in a domain-specific fashion. A more general mechanisms for handling this would require a rule language.

3. AN APPLICATION SCENARIO

3.1 Annotating art-historic features

Figure 4 shows a screenshot of the annotation interface. In this scenario the user is annotating an image representing the painting by Chagall of Figure 2. The figure shows the tab for production-related VRA data elements. The four elements with a "binoculars" icon are linked to subtrees in the ontologies, i.e., AAT and ULAN. For example, if we would click on the "binoculars" for style/period the window shown in Figure 5 would pop up, showing the place in the hierarchy of the concept `Surrealist`. We see that it is a concept from AAT. The top-level concepts of the AAT subtrees from which we can select a value for style/period are shown with an underlined bold font (i.e., <`styles and periods by general era`> and <`styles and periods by region`>).

3.2 Using existing annotations

The collection of art paintings that was used for this study, was accompanied by short semistructured textual annotations. For example, this is the text accompanying the chagall painting:

```
Chagall, Marc
Birthday
1915
Oil on cardboard
31 3/4 x 39 1/4 in.
The museum of Modern Art, New York
```

We included in the tool a parsing facility, implemented as a special-purpose set of definite-clause grammar rules, This facility is able to create a partial annotation from these texts. For the image in Figure 4 the following VRA slot values could be derived directly from the text: title, creator, date, materials.support, materials.medium, measurements, location and ID . For the style/period slot a value is suggested based on the slot value for creator. The same is done for technique, the value for which can be derived from the two material

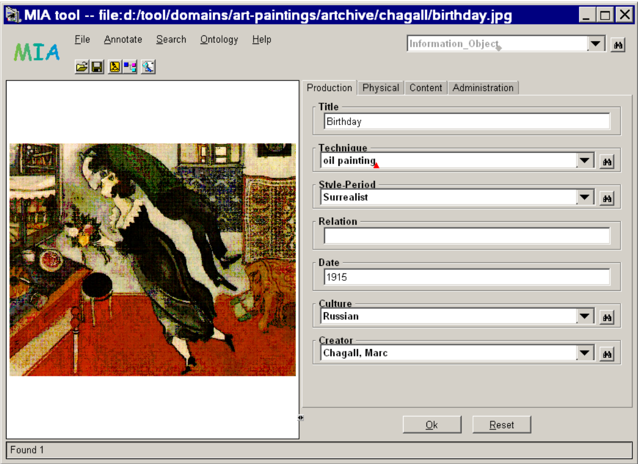

Figure 4: Screenshot of the annotation interface The figure shows one tab with VRA data elements for describing the image, here the production-related descriptors. The slots associated with a "binoculars" button are linked to one or more subparts of the underlying ontologies, which provide the concepts for this part of the annotation

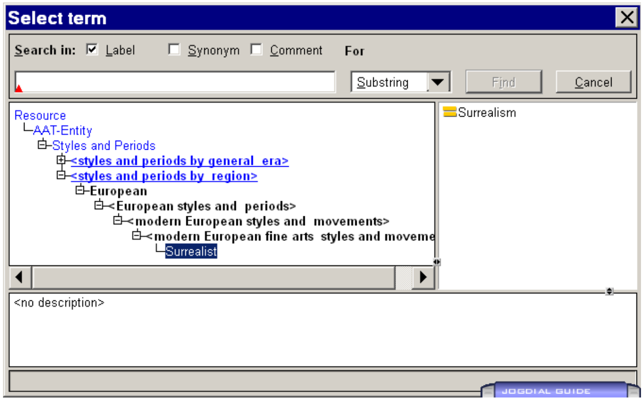

Figure 5: Browser window for values of style/period. The concept `Surrealist` has been selected as a value for this slot. The top-level concepts of the AAT subtrees from which we can select a value for style/period are shown with an underlined bold font (i.e., `<styles and periods by general era>` and `<styles and periods by region>`)

slots. In Figure 4 all values except for culture are derived automatically from the existing annotation.

3.3 Annotating image content

Figure 6 shows the annotation of the content of the painting called "Portrait of Derain" by Maurice de Vlaminck. The template on the right-hand side implements the subject template as described in Section 2.2. The content has been tersely described with the following terms:

```
Agent:    "Derain, Andre"                (ULAN)
Action:   "smoke"                        (WordNet)
Object:   "pipes(smoking equipment)"     (AAT)
```

As with the art-historic features, the slots are linked to one or more subparts of the underlying ontologies, which

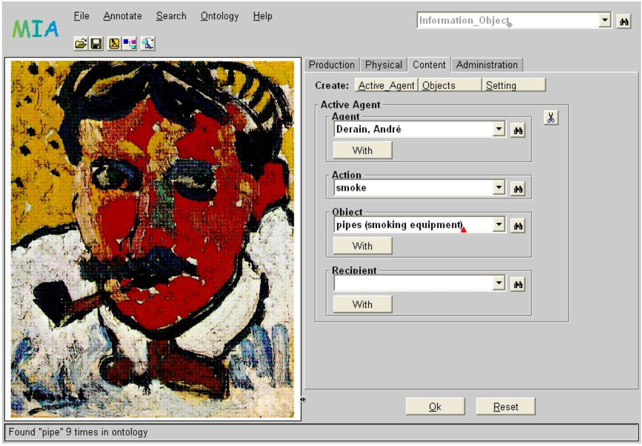

Figure 6: Description of the content of the painting "Portrait of Derain"

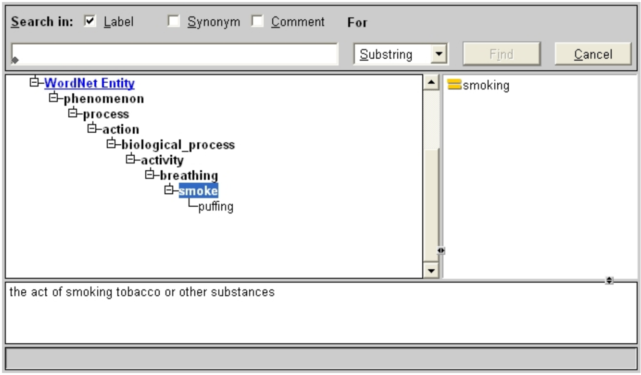

Figure 7: Browser window for the concept `smoke`

provide the concepts for this part of the annotation. For example, if we would click on the binocular icon for action the window shown in Figure 7 would pop up, showing the place in the hierarchy of the concept `smoke`. We see that it is a concept from WordNet.

The user interface provides some support for finding the right concept. For example, the user can type in a few characters of a term and then invoke a completion mechanism (by typing a space). This will provide a popup list of concepts matching the input string. In the browser window, more advanced concept search options can be selected, including substrings and use of synonyms. One synonym of `smoke` is provided, namely `smoking`. The ontology makes it easier for people to select the correct concept. For example, seeing the specialization `puffing` of the concept `smoke`, the user might decide to use this term.

For annotation purposes the ontologies serve two purposes. Firstly, the user is immediately provided with the right context for finding an adequate index term. This ensures quicker and more precise indexing. Also, the hierarchical presentation of concepts helps to disambiguate terms. When the user types in the term "pipe" as the object in a content-description template, the tool will indicate that this an ambiguous term. In the user interface the term itself gets a green color to indicate this and the status bar near the bottom shows the number of hits in the ontologies. If

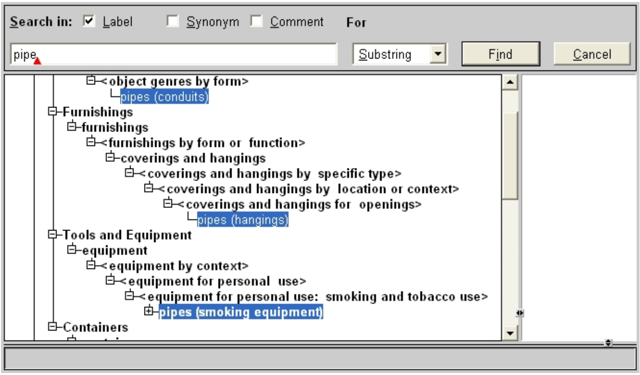

Figure 8: Browser window for the `pipe` concepts

one clicks on the binoculars button, the tool will provide the user with a choice of concepts from the ontologies that are associated with this term. Figure 8 shows three of the concepts associated with `pipe`, namely `conduits`, `hangings` and `smoking equipment`. From the placement of the terms in the respective hierarchies, it is usually immediately clear to the indexer which meaning of the term is the intended one. Term disambiguation is a frequent occurrence in this type of application.

The ontologies provide a wide range of concepts for the subject-matter descriptions. Although the choice of concepts depends on the indexer, and although the quality of an annotation is subjective, there are some general guidelines for good annotations. An annotation is most effective if the annotator chooses the concepts as specific as possible. Experiments [5] have shown that users describe images in terms of the agents and objects that are depicted. An annotator should therefore focus on agent and object descriptions.

3.4 Searching for an image

The tool provides two types of semantic search. With the first search option the user can search for concepts at a random place in the image annotation. Figure 9 shows an example of this. Suppose the user wants to search for images associated with the concept `Aphrodite`. Because the ontologies contain an equivalence relation between `Venus` (as a Roman deity, not the planet nor the tennis player) and `Aphrodite`, the search tool is able to retrieve images for which there is no syntactic match. For example, if we would look at the annotation of the first hit in the right-hand part of Figure 9, we would find "Venus" in the title ("Birth of Venus" by Botticelli) and in the subject-matter description (`Venus (a Roman deity) standing seashell`). The word "Venus" in the title can only be used for syntactic marches (we do not have an ontology for titles), but the concept in the subject description can be used for semantic matches, thus satisfying the "Aphrodite" query.

General concept search retrieves images which match the query in some part of the annotation. The second search option allows the user to exploit the annotation template for search proposes. An example of this is shown in Figure 10. Here, the user is searching for images in which the slot `culture` matches `Netherlandish`. This query retrieves all images with a semantic match for this slot. This includes images of `Dutch` and `Flemish` paintings, as these are subconcepts of `Netherlandish`.

4. RELATED WORK

The architecture shown in Figure 1 is in the same spirit as the one described by Lafon and Bos [7]. The main difference lies in the fact that we place more emphasis on the nature of the ontologies. Koivunen and Swick [6] discuss an architecture semantic annotation, but mainly from the perspective of the shared collaborations. CREAM [3] also provides an architecture for semantic annotation including both manual and semi-automatic techniques. The present work differs from the latter two approaches through its focus on images (which creates special problems, such as annotating the image content) and the practical work on integrating multiple existing ontologies. The work of Hyvönen and colleagues [4] combines ontology-based image retrieval view view-based and topic-based retrieval and is probably closest to the present work. So far, they have not reported many details on the ontologies being used.

5. DISCUSSION

This paper gives some indication on how a semantic web for images might work. Semantic annotation allows us to make use of concept search instead of keyword search. It paves also the way for more advanced search strategies. For example, users can specialize or generalise a query with the help of the concept hierarchy when too many or too few hits are found.

In a previous study on a collection of ape photographs [13] we did some qualitative analysis on the added value with respect to keyword search. The provisional conclusion was that for some queries (e.g., "ape") keyword search does reasonably well, but for other sightly different queries (e.g., "great ape") the results are suddenly poor. This is exactly where semantic annotation could help.

In another prior study [12] we reported on a small experiment concerning the usability of the annotation toll. Although our approach relies to some extent on manual annotation, it should be possible to generate partial semantic annotations from existing annotations (which vary from free text to structured database entries). The application scenario in Section 3 shows an example of this. However, the example is based on a special-purpose parser. Systematic use of NLP techniques should be considered here. Also, content-based image analysis techniques could be used to derive image features, such as the location and color of objects.

Our experiences with RDF Schema were generally positive. We made heavy use of the metamodelling facilities of RDF Schema (which allows one to treat classes as instances of other classes) for defining and manipulating the metamodels of the different thesauri. In our experience this feature is in particular needed in cases where one has to work with existing representations of large ontologies. This is a typical feature for a semantic web: one has to work with existing ontologies to get anywhere, even if one disagrees with some of the design principles of the ontology.

For our purposes RDF Schema has some limitations in expressivity. We especially needed a notion of property cardinality and of equivalence between resources (classes, instances, properties). For this reason we plan to move at some near point in the future to OWL, the Web Ontology

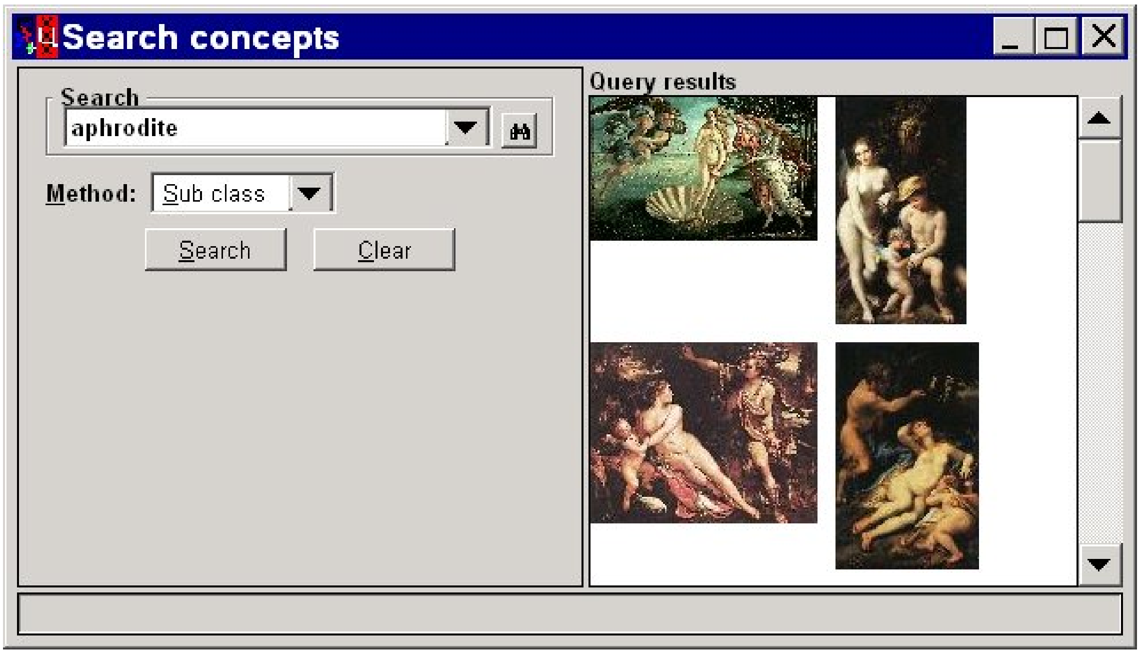

Figure 9: Example of concept search. The query "Aphrodite" will retrieve all images for which we can derive a semantic match with the concept `Aphrodite`. This includes all images annotated with the concept `Venus` (as a Roman deity). Only a small fragment of the search results is depicted

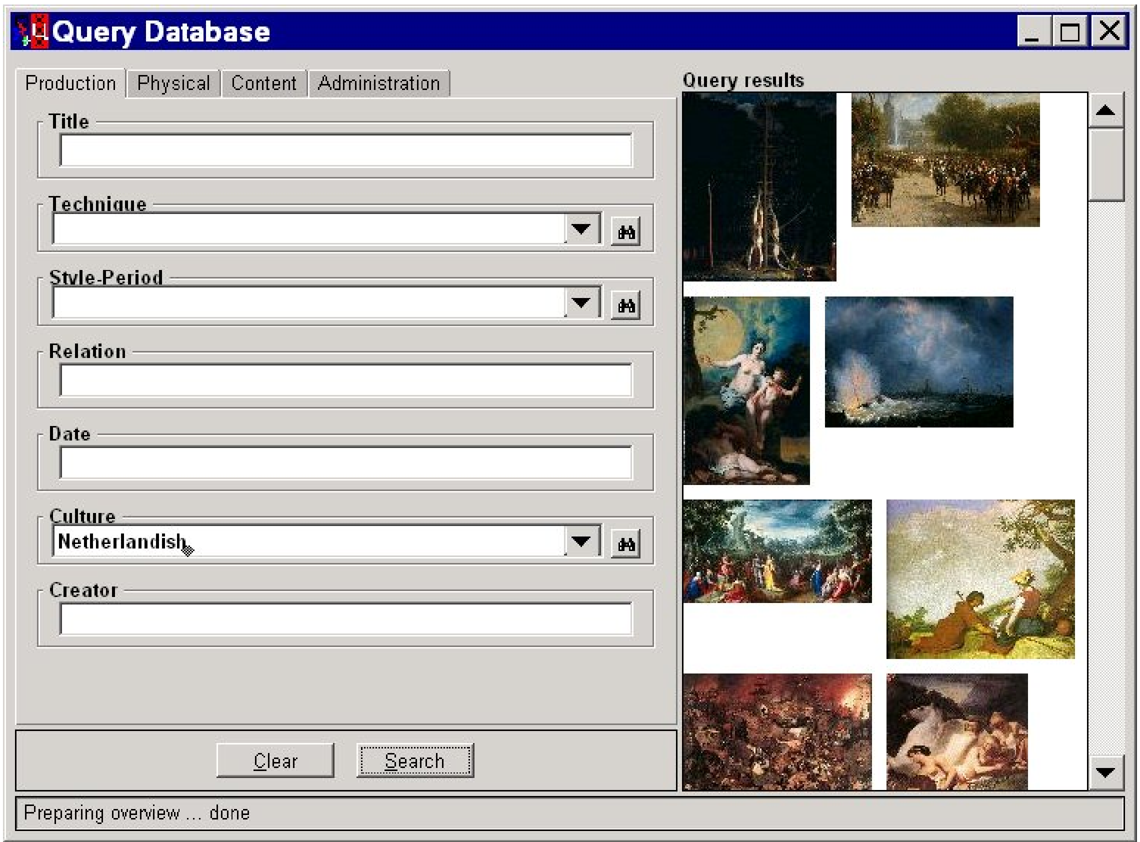

Figure 10: Search using the annotation template. The query "Netherlandish" for the slot culture retrieve all images with a semantic match for this slot. This includes images of `Dutch` and `Flemish` paintings, as these are subconcepts of `Netherlandish`

Language currently under development at W3C [18].

Acknowledgments

This work was supported by the IOP Project "Interactive disclosure of Multimedia Information and Knowledge" and the ICES-KIS project "Multimedia Information Analysis", both funded by the Dutch Ministry of Economic Affairs. We gratefully acknowledge the contributions of Marcel Worring, Giang Nguyen and Maurice de Mare.

6. REFERENCES

[1] D. Brickley and R. V. Guha. Resource description framework (RDF) schema specification 1.0. Candidate recommendation, W3C Consortium, 27 March 2000. See: http://www.w3.org.

[2] The Getty Foundation. ULAN: Union List of Artist Names. http://www.getty.edu/research/tools/vocabulary/ulan/, 2000.

[3] S. Handschuh and S. Staab. Annotation of the shallow and the deep web. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*, volume 96 of *Frontiers in Artificial Intelligence and Applications*, pages 25–45. IOS Press, Amsterdam, 2003.

[4] E. Hyvönen, S. Kettula, V. Raatikka, S. Saarela, and K. Viljanen. Finnish museums on the semantic web. In *Proceedings of WWW2003, Budapest, poster papers*, 2003.

[5] C. Jörgensen. Indexing images: Testing an image description template. In *ASUS 1996 Annual Conference Proceedings*, 1996.

[6] M-R. Koivunen and R. R. Swick. Collaboration through annotation on the semantic web. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*, volume 96 of *Frontiers in Artificial Intelligence and Applications*, pages 46–60. IOS Press, Amsterdam, 2003.

[7] Y. Lafon and B. Bos. Describing and retrieving photographs using RDF and HTTP. Note, W3C Consortium, 28 September 2000. URL: http://www.w3.org/TR/2000/NOTE-photo-rdf-20000928.

[8] G. Miller. WordNet: A lexical database for english. *Comm. ACM*, 38(11), November 1995.

[9] I. Niles and A. Pease. Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In *Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas, Nevada*, June 23-26 2003.

[10] Bijan Parsia. Rdf applications with prolog. http://www.xml.com/pub/a/2001/07/25/prologrdf.html, 2001.

[11] T. Peterson. *Introduction to the Art and Architecture Thesaurus*. Oxford University Press, 1994. See also: http://www.getty.edu/research/tools/vocabulary/aat/.

[12] A. Th. Schreiber, I. I. Blok, D. Carlier, W. P. C. van Gent, J. Hokstam, and U. Roos. A mini-experiment in semantic annotation. In I. Horrocks and J. Hendler, editors, *The Semantic Web - ISWC 2002*, number 2342 in Lecture Notes in Computer Science, pages 404–408, Berlin, 2002. Springer-Verlag. ISSN 0302-9743.

[13] A. Th. Schreiber, B. Dubbeldam, J. Wielemaker, and B. J. Wielinga. Ontology-based photo annotation. *IEEE Intelligent Systems*, 16(3):66–74, May/June 2001.

[14] A. M. Tam and C. H. C. Leung. Structured natural-language description for semantic content retrieval. *Journal of the American Society for Information Science*, to appear.

[15] J. van den Berg. Subject retrieval in pictorial information systems. In *Electronic Filing, Registration, and Communication of Visual Historical Data. Abstracts for Round Table no 34 of the 18th International Congress of Historical Sciences. Copenhagen*, pages 21–28, 1995. http://www.iconclass.nl.

[16] H. van der Waal. ICONCLASS: An inconographic classification system. Technical report, Royal Dutch Academy of Sciences (KNAW), 1985.

[17] Visual Resources Association Standards Committee. VRA Core Categories, Version 3.0. Technical report, Visual Resources Association, July 2000. URL: http://www.vraweb.org/vracore3.htm.

[18] Web Ontology Working Group. OWL Web Ontology Language Overview. W3C Candidate Recommendation, World Wide Web Consortium, 18 August 2003. Latest version: http://www.w3.org/TR/owl-features/.

[19] J. Wielemaker. *SWI-Prolog RDF Parser*. SWI, University of Amsterdam, 2000. URL: http://www.swi-prolog.org/packages/rdf2pl.html.

[20] J. Wielemaker, A. Th. Schreiber, and B. J. Wielinga. Prolog-based infrastructure for rdf: performance and scalability. In *Proceedings ISWC'03*, 2003.

[21] B. J. Wielinga, A. Th. Schreiber, J. Wielemaker, and J. A. C. Sandberg. From thesaurus to ontology. In Y. Gil, M. Musen, and J. Shavlik, editors, *Proceedings 1st International Conference on Knowledge Capture, Victoria, Canada*, pages 194–201, New York, 21-23 October 2001. ACM Press.