EFFICIENT INTEGRATION OF SEMANTIC TECHNOLOGIES FOR PROFESSIONAL IMAGE ANNOTATION AND SEARCH

Andreas Walter
FZI Forschungszentrum Informatik, Haid-und-Neu-Straße 10-14, 76131 Karlsruhe, Germany, awalter@fzi.de

Gábor Nagypál
disy Informationssysteme GmbH, Erbprinzenstr. 4-12, Eingang B, 76133 Karlsruhe, Germany, nagypal@disy.net

ABSTRACT

Professional image annotation and search concerns two user groups. The first group consists of image agencies who aim to sell their images. The second group is formed by image buyers who search and use these images for the illustration of printed materials or for advertising. Professional image annotation and search is based on text based techniques for many years. Semantic technologies help improve both, the quality of the image annotation for image editors and the possibilities for the navigation through an image archive for image searcher. In this paper, we present the overall result of our research. We first collected the requirements of the two user groups for an effective usage and integration of semantic technologies in image archives. Based on these requirements, we created a visual methodology and application which fulfils these needs.

KEYWORDS

Semantic Image Annotation, Ontology Maturing

1. INTRODUCTION

Image agencies and image buyers form the two user groups in the area of professional image annotation and search. An image agency runs an image search engine, nowadays mostly web-based, to offer image contents to potential buyers. Image recognition techniques are not yet mature and scalable enough to provide image search. Therefore, image editors have to annotate their images before image buyers can search for them. Using text annotations based on keywords, aka tags, is the traditional and well accepted way for the annotation of images in image agencies.

The creation of image annotations is a very time consuming task. We observed image editors to determine how they use tags for the creation of image annotations. Mostly, they create very large annotations using tags. The annotations consist of synonymous meanings of words, of the same words in singular and plural versions, of the same words in different languages and also of hierarchies of words, such as “Germany, Country and Europe”. The motivation of image annotators to perform this time consuming work is to ensure, that their images are shown in relevant search results of image searchers.

Text based image search is also a very time consuming task. Image searchers usually execute multiple searches to find their desired images because they have to try different synonymous meanings (also in different languages) of their search words to ensure that they found all of the images they are interested in.

Semantic technologies have proven that they have significant advantages compared to text based technologies [8]. It is possible to create a structured representation of background knowledge using ontologies. Each element of an ontology can offer support for multilingualism and synonymous meanings. Once such an element is created, it can be used to create semantic annotations. Consequently, image annotators are interested in using semantic technologies to save annotation time. Instead of writing the same

text-based tags for many images again and again, they may reuse the available ontology elements. Image searchers are interested in using semantic technologies, as well, because they improve the results of search requests. A semantic search request based on an ontology element can automatically search for different synonymous meaning in different languages [8]. Thus, image searchers save time because less search requests are needed. In addition, homonyms (such as computer mouse and mouse as an animal) have different elements in the ontology and therefore they are not mixed. Finally, relations between ontology elements may be defined, which provides the possibility to search for semantically related images based on a given search request (e.g. search for Germany after a search for Angela Merkel, the German chancellor).

In our current project, it has turned out very quickly, that the integration of existing semantic technologies were not usable for a real-world image search applications. The methodologies and workflows for the creation of semantic image annotations did not fit the identified scenarios of image annotators and searchers: image annotators could not extend ontologies when they needed new ontology elements, already available ontologies were not understandable and available tools were too complicated for them. This motivated our research for a practical usable workflow and methodology, which enables an efficient integration of semantic technologies in professional image annotation and search systems.

We analyze the usual scenarios of image annotators and searchers using a text based image search engine, and report on the required extensions for an efficient integration of semantic technologies in Section 2. In Section 3, we present our solution for a suitable integration of semantic technologies as expected by image annotators and searchers, the so called ImageNotion application. Section 4 concludes the paper.

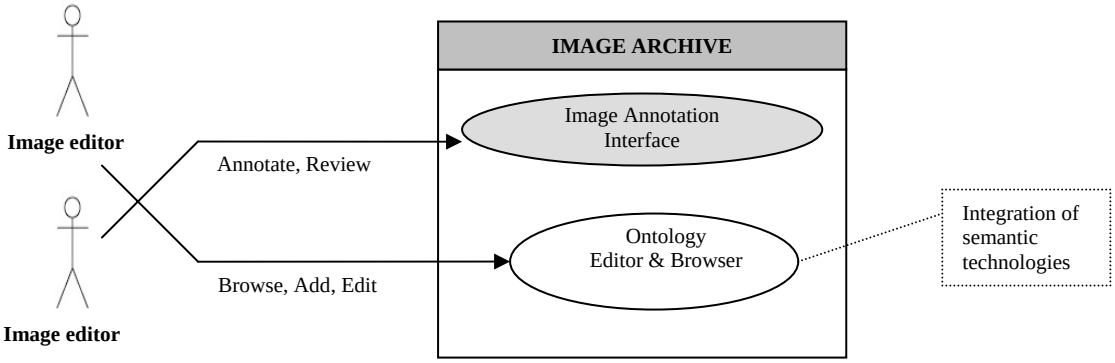
2. EXTENSION OF SCENARIOS IN AN IMAGE SEARCH ENGINE FOR THE INTEGRATION OF SEMANTIC TECHNOLOGIES

We have observed the interaction of image editors and searchers with classical text-based systems to identify the relevant scenarios and requirements for the integration of semantic technologies into an image search engine.

2.1 Annotation of images

The *image annotation interface* in Figure 1 is used in a text-based image search engine to create textual annotations. In an image agency, a team of image editors collaboratively creates annotations for the existing image contents. Their task is the identification of required tags which may be easily fulfilled by just typing in the new tags to the annotation. Other image editors may review annotations. They may add more tags or can remove existing ones. Normally, they agree in a common style of creating annotations, e.g. what languages to use.

Figure 1: Integration of semantic technologies for the annotation of images



For the creation of semantic annotations, the usage of ontologies is required. Ontologies may have different grades of formality, e.g. with a low formality for grouping words and their synonyms, containment of *is-a* relations as used in many thesauri (e.g. used in WordNet [8]) or in addition unnamed relations as used e.g. in the SKOS vocabulary [8]. If image annotators or searchers require it, it is also possible to use ontologies with a high formality, e.g. OWL [5]. Figure 1 shows the integration of semantic technologies in an image search engine for the annotation of images with the new component O*ntology Editor and Browser*.

To successfully create semantic annotations image editors require that they can browse the ontology to reuse existing ontology elements. Also, they expect the possibility to extend ontologies with new elements, when it is required for new image contents. Thus, the resulting workflow must allow a work integrated, collaborative methodology for the engineering and maintenance of ontologies and for the creation of semantic annotations. In addition, the image editors require one integrated application instead of a number of different applications to create semantic annotations.

Existing methodologies and tools for the creation of ontologies and semantic technologies could not support the expected workflow of image annotators [9]. The methodologies assume a separated creation of ontologies by ontology engineers. Thus, a work integrated extension of ontologies is not possible with these methodologies. The resulting ontologies are usually too complicated for image editors without much ontology experience or do not cover all ontology elements required for annotating the images. In addition, the creation of semantic annotations requires using different tools. These tools were too complicated for the image editors; in addition they did not allow a collaborative creation of semantic annotations.

Consequently, the creation of a new methodology was necessary to cover the expected workflow [8]. We will present this methodology in the next section.

2.2 Image Search

Image searchers use the component *Image Search Interface* in a text based image search engine to formulate their required image contents (see Figure 2). Therefore, they type a number of tags and start the search process. The image search engine then returns matching images and displays them on a result page.

Figure 2: Integration of semantic technologies for the search for images

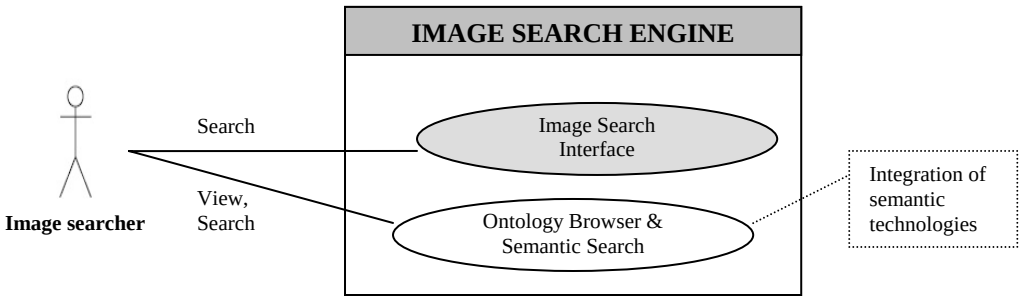


Figure 2 shows the integration of semantic technologies in an image search engine. A new component *Ontology Browser & Semantic search* was added. Image searchers can use this component to view existing ontology elements and to use them for the formulation of semantic search requests.

In addition, image searchers expect navigation mechanisms when using semantic search. Therefore, they like to start with a text-based image search. Then, based on the text-based query, the image search engine should look up matching semantic elements and propose them for the refinement of a text based query. Also, they would like to navigate to images related to a given search requests. This is possible by using the relations between semantic elements to generate a proposal for related semantic elements to refine the search. Another requirement of image searchers is the possibility to navigate through image parts. Therefore, image annotators must first create semantic annotations for image parts. Then, image searchers can use these annotations by clicking on image parts to start a new semantic search.

So far, there were no image search engines which could fulfill all the discussed requirements for the integration of semantic technologies [8]. Therefore, we have included relevant features in our ImageNotion application.

2.3. Photo sharing platforms

Based on the collaborative features of Web 2.0 applications, photo sharing platforms, e.g. Flickr [4] or Riya [8], have emerged. These platforms enable everyone (mostly private persons with non-profit interests) to upload, annotate and search for images. In these platforms, a user can represent both user groups of a professional image search engine at the same time. As an image annotator, a user can annotate her own images or images of other users; as image searcher she can search for images. Since these platforms are based on textual image annotations, photo sharing platforms have similar requirements as professional image search engines for the integration of semantic technologies, but a much bigger number of participants. Because of smaller user groups and higher experiences for using image search engines, we focused on professional image annotators and searchers for the evaluation of suitable methodologies and applications.

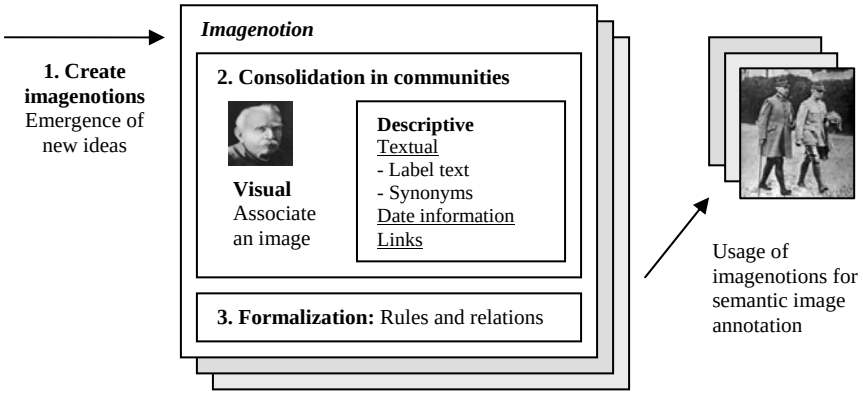
3. THE IMAGENOTION APPLICATION

Based on the identified user scenarios and requirements, we have created a suitable methodology and application to fulfill these requirements, called ImageNotion.

3.1 The ImageNotion Methodology for a work integrated ontology maturing

The basis of our ontology formalism is a concept we call imagenotion [8]. An imagenotion (formed from the words image and notion) graphically represents a semantic notion through an image. Furthermore, similarly to many existing ontology formalisms, it is possible to associate descriptive information with an imagenotion. A part of the descriptive information is textual information: a main textual label and alternative labels in different languages. Additionally, date information allows searching for images based on an exact date or a time interval. Further, it is possible to add links to related web pages for an imagenotion. Links help users and also automated processes such as text mining algorithms to gather background information from web pages, and thus to support the maturing of existing imagenotion. In addition to descriptive information, relations between imagenotions are also supported. To achieve maximal understandability, ImageNotion makes no distinction between concepts and instances. Using imagenotions, users do not need to understand this somewhat artificial separation of ontology elements.

Figure 3: The ImageNotion methodology



The aim of the ImageNotion methodology (see Figure 3) is to guide the process of creating an ontology of imagenotions. The main steps of this methodology are based on the ontology maturing process model [8]. Step 1 is the creation of new imagenotions, step 2 is the consolidation of imagenotions in communities and step 3 is the formalization of imagenotions with rules and relations. Imagenotions from each maturing grade may be used for semantic image annotation.

3.2 Extended scenario for the annotation of images using ImageNotion

Using the ImageNotion methodology allows for the integration of semantic technologies in the scenario for annotation images as desired by image editors. We have implemented the methodology in the web based ImageNotion application. ImageNotion allows for a collaborative, work integrated creation and maturing of ontologies and semantic annotation of images.

Figure 4: Semantic annotations of images using ImageNotion

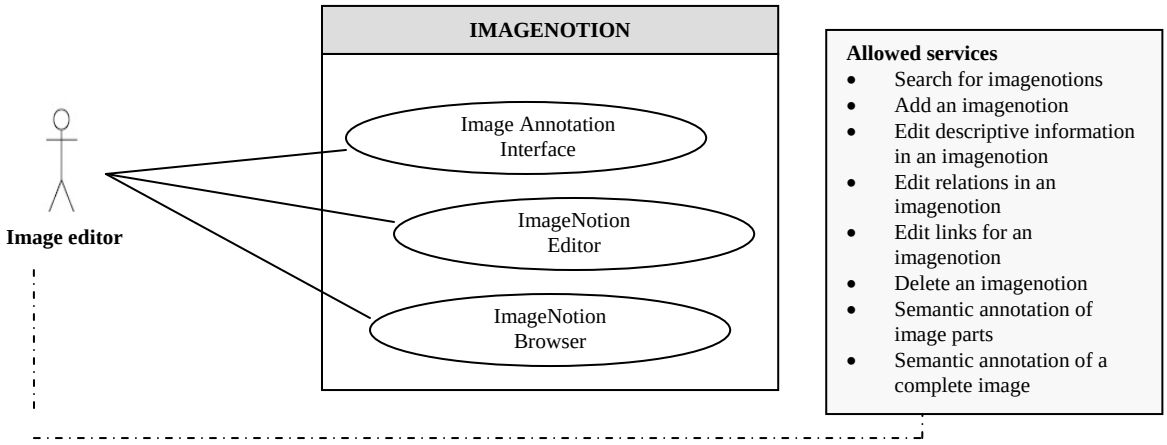
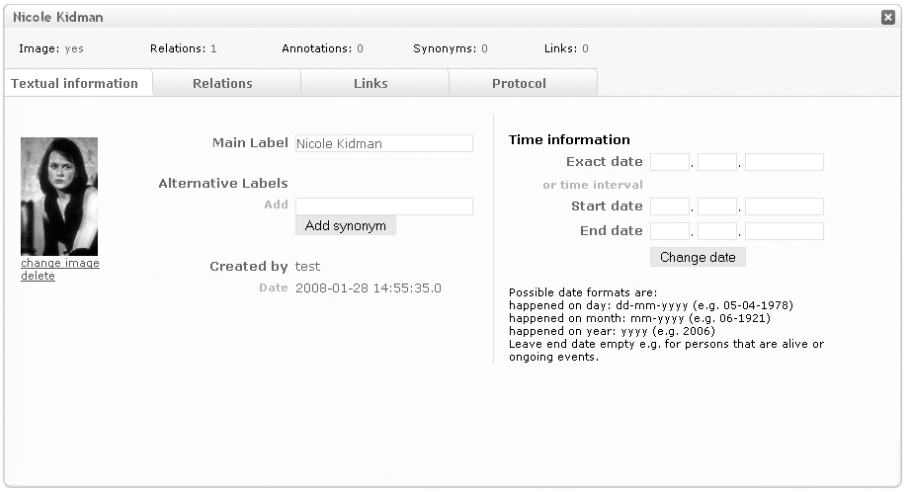


Figure 4 shows the scenario for the semantic annotation of images using the ImageNotion application. In a preparation step, a group of image editors discusses which ontology formalism should be used for the semantic annotation of images. The ImageNotion application currently allows for the creation of maximal three relations (similar to SKOS): is-a (broader), narrower and unnamed relations. In our evaluations, the groups of image annotators were interested in keeping things simple and considered this formalism as adequate for initial versions of their ontologies.

Figure 5: Example for editing an imagenotion



Based on the selected formalisms, the group of image editors can create an initial ontology. In our evaluations, the groups were interested in semantic interoperability with other systems and therefore were interested in using main parts of standardized ontologies as core ontology, e.g. CIDOC-CRM [3]. Also, it would be possible to start with a collaboratively discussed and agreed initial version of an own ontology or with an empty ontology.

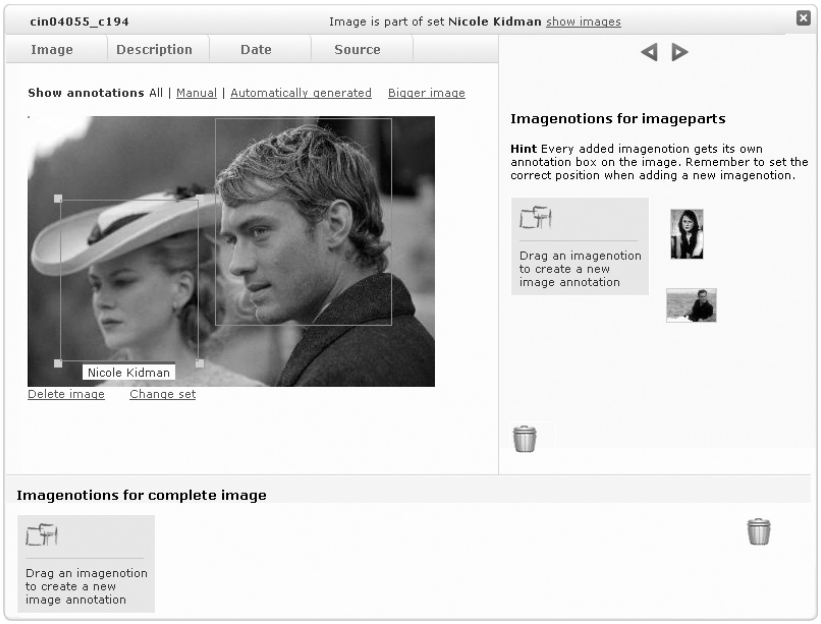
For the annotation of images, each image editor can use the services of the ImageNotion application as shown in Figure 4. Using the *ImageNotion Browser*, they can search for existing imagenotions they need for the semantic annotation of images. They can view the contents of each imagenotion. In case they have identified missing contents in an imagenotion, they can edit descriptive information, relations and links to web pages. In case of missing imagenotions, they can add them. They can also remove incorrect or unneeded imagenotions. All image editors can work collaboratively in the application and review the created imagenotions of other image editors. All steps together help to increase the quality of the imagenotions. Each imagenotion may be used for the semantic annotation of images. The more information is available for an imagenotion, the better becomes the quality of the semantic annotations for images. Thus, maturing the imagenotions increases the quality of semantic image annotations.

In our evaluations, image editors succeeded in performing the tasks for the creation of semantic annotations. Also, they were very happy about the visual ImageNotion methodology, its understandability and the easy usability of the ImageNotion application for the semantic annotation of images.

3.3 The ImageNotion Editor

Figure 5 shows an example for the creation of an imagenotion using the ImageNotion editor. The image editor wants to create semantic annotations for an image archive containing film stars. For images displaying “Nicole Kidman”, the image editor creates a new imagenotion.

Figure 6: Example for the creation of semantic image annotations

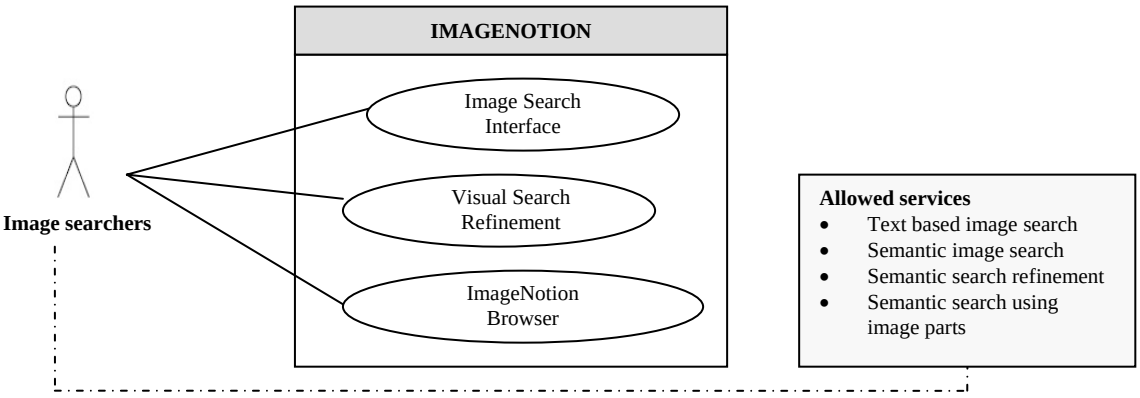


3.4 Semantic annotation of images with imagenotions

Existing imagenotions may be used for the semantic annotation of images and image parts. Figure 6 shows an example for the annotation of an image showing the film stars “Nicole Kidman” and “Jude Law”. For an easy usability, the semantic annotations may be added using drag & drop mechanisms.

Because the web based application was created with state of the art techniques for web applications (e.g. AJAX), the application quickly responds and adds annotation boxes for image parts without reloading the whole page. The image editor may then set the correct position of an image box, e.g. for the face of “Nicole Kidman”.

Figure 7: Semantic search using ImageNotion



3.5 Extended scenario for image search in ImageNotion

Figure 7 shows the extended scenario for image searchers using the ImageNotion application. To allow for the expected navigation mechanisms of image searchers, they can still start with a text based image search. The new component *Visual Search Refinement* allows for the refinment of text based search requests to semantic search requests. Therefore, the system displays imagenotions, which contain the textual search requests. With drag & drop techniques, users then may visually refine their search request using their desired imagenotion. The system starts a new search and displays only images which contain the used imagenotion as semantic annotation for the complete image or for image parts. Moreover, this component checks combinations of semantic annotations with a given semantic search request. This allows the refinement of a semantic search with a number of imagenotions to create a more specific search request.

The *ImageNotion Browser* allows the browsing of all available imagenotions. The image searcher can start a new semantic search with each available imagenotion

3.6 Visual search refinement

Figure 8: Semantic search using visual search refinement

Figure 8 shows an example of the visual search refinement component. The image searcher has started with a text based search “Kidman”. The component has proposed the refinment to the semantic search for the imagenotion “Nicole Kidman”. In the next step, the component checked for combinations of semantic annotations. In this example, the component proposes the refinment to images containing the semantic annotations base on the imagenotions “Nicole Kidman” and “Jude Law”.

4. CONCLUSION AND FUTURE WORK

In this paper, we have first analyzed the existing scenarios of professional image editors and annotators using a text-based image search engine. Based on these scenarios, we described how semantic technologies can be integrated in these scenarios and what new requirements arise. Since none of the existing methodologies and applications has fulfilled all these requirements, we have created and proposed our ImageNotion methodology and application. It allows for the semantic annotation and semantic search of images in one integrated web based application. Work-integrated and collaborative ontology editing is also supported by the same application. For image searchers, the application allows for the semantic navigation through an image archive by refining textual search requests to semantic search requests and implementing the novel technique of visual search refinement.

Our future work is the extension of the methodology to support more complex ontologies, e.g. with more relations, or more complex formalisms, such as OWL. In addition, we are extending the manual process for the creation of semantic annotations with automated processes. We combine face detection and recognition algorithms, object identification algorithms and text mining algorithms to achieve automatically created semantic annotations with a high quality. This will help to reduce the human effort needed for the semantic annotation of images.

ACKNOWLEDGEMENT

For the final version only

REFERENCES

1. Braun, S et. al.,2007, Ontology Maturing: A Collaborative Web 2.0 Approach to Ontology Engineering, *Proc. of the Workshop on Social and Collaborative Construction of Structured Knowledge*, *WWW 2007*, Banff, Alberta, Canada
2. Brickley, D; Miles, A., 2005, SKOS Core Vocabulary Specification, W3C, *http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20051102*
3. Crofts, N. et. al., 2005, Definition of the cidoc conceptual reference model version 4.2. *CIDOC CRM Special Interest Group*.
4. Flickr: Welcome to Flickr – Photo Sharing. 2007, *http://www.flickr.com/*
5. Mc Guinness, D.L., van Harmelen, F., 2004, OWL Web Ontology Language. *W3C working draft, W3C*
6. M. Lew, et.al.: 2006, Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications, vol.2, issue 1*.
7. Nagypal, G., 2007, *Possibly imperfect ontologies for effective information retrieval*. In: PhD thesis, University of Karlsruhe.
8. Riya. Riya - Visual search. 2007, *http://www.riya.com/, 2007*
9. Walter, A., Nagypal, G.. 2007, Imagenotion - methodology, tool support and evaluation. GADA/DOA/CoopIS/ODBASE 2007 Confederated International Conferences DOA, CoopIS and ODBASE, Proceedings, LNCS. Springer, Vilamoure, Portugal
10. Walter, A., Nagypal, G., 2008, Evaluation of semantic technologies for the navigation through image archives, currently under review for ESWC 2008
11. WordNet. A lexical database for the English language, 2007, http://wordnet.princeton.edu/