Requirements for Multimedia Document Enrichment

Ajay Chakravarthy
University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield, S1 4P, United Kingdom
+44-114-2221945

Vitaveska Lanfranchi
University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield, S1 4P, United Kingdom
+44-114-2221945

Fabio Ciravegna
University of Sheffield
Regent Court, 211 Portobello Street,
Sheffield, S1 4P, United Kingdom
+44-114-2221945

a.chakravarthy@dcs.shef.ac.uk   v.lanfranchi@dcs.shef.ac.uk   f.ciravegna@dcs.shef.ac.uk

ABSTRACT
Nowadays a large and growing percentage of information is stored in various multimedia formats. In order for multimedia information to be efficiently utilised by users, it is very important to add suitable metadata. In this paper we will present AKTiveMedia, a tool for enriching multimedia documents with semantic information.

Categories and Subject Descriptors
H.5 [**Information Systems**] Information Interfaces And Presentation

General Terms
Human Factors

Keywords
Semantic annotation interfaces, multimedia enrichment.

1. INTRODUCTION
Nowadays a large and growing amount of information is stored in various multimedia formats, such as images, video, audio. Much research has been undertaken into the efficient and effective storage, access, usage and retrieval of textual information. Semantic annotation and enrichment has been proposed as a way to make information available in documents for effective and efficient use. For example, several activities focused on text annotation as a way to enrich a textual document, making it machine-readable and also accessible to people [1, 2, 3, 6, 8]; other projects focus more on annotation of images [4, 5].

However, we believe that the separation of text and images is artificial and there is a real need for enabling true multimedia annotations that go across the division of text and images. As a matter of fact, a great deal of information, like websites, intranets or company reports often contain both text and images that are interrelated: usually the text in the document contains references to the image or describes it. It is therefore crucial to develop strategies and interfaces for cross-media knowledge creation and sharing that will make these references explicit, increasing the value of the document itself.

In the rest of this paper we will outline some requirements for multimedia document enrichment and then we will present AKTiveMedia, a tool for enriching multimedia documents that meets those requirements.

2. REQUIREMENTS FOR MULTIMEDIA DOCUMENT ENRICHMENT
Different requirements can be identified for multimedia document enrichment in the various phases of the knowledge production, that is when the document is annotated (i.e. knowledge acquiring) and when the annotations are used to share the produced knowledge (i.e. knowledge sharing and reuse).

First all is important to outline which are the dimensions of the content that can be identified and enriched with annotations: (1) Metadata, that can be associated to the multimedia document, like creation date, time, author, and so on; (2) Content: the annotation should semantically describe the content of the image or document so to make it available for retrieval. For example, marking up elements in a picture (e.g. the person in the picture is Sam). (3) Contextual information, that is information about an instance in a precise time and occasion (e.g. Sam is awaiting a taxi in Southampton); (4) General knowledge, about the instances annotated, typically immutable knowledge, for example Sam was born on 26-07-1974; (5) Relations between the identified instance and other instances in the same document or in other document. Relations may be cross-media, e.g. that may interrelate what is presented in the text with part of an image, e.g. "Fabio" mentioned in the text is_employer_of Sam that is the person in the picture; (5) Comments about the instances or the information or the relations (e.g. "This is an artistic picture, in reality Sam doesn't have pink hair").

Following [6] we identified two main types of annotations that can be used to enrich the content: (1) *Ontology-driven annotations* can be used to make available the document content or its context. They empower better retrieval and reasoning; [1] and connect to semantically enabled services that are compatible with the ontology. (2) *Free text annotations* (or comments) add contextual knowledge to the document not explicitly mentioned within the content (often referred to as *braindump* or freetext tagging, e.g. folksonomies).

The task of annotating a document using an ontology may be tiresome [1] and overwhelming: usually ontologies are complex and users may find difficult to remember all the available concepts and to use them properly. As previous literature proved [7], when dealing with vast quantities of information users may want to zoom and visualise only the sections they are interested in, or filter out what is not relevant for the current task. A main requirement for annotation of multimedia documents is therefore to make the ontology "disappear", i.e. creating a more user-friendly annotation methodology can hide the complexity of an ontology whilst still maintaining its expressivity. Moreover it is possible to distinguish two main styles when annotating images:

(1) Whole image annotation: the entire image is annotated; this can be done for the single image or for image collections; (2) Regional annotation: different regions of the image are annotated. The two different styles of image annotation are accomplished using both ontology-based annotations and free-text annotations. n the phase of knowledge sharing and reuse the annotations and comments inserted by are used by the system to provide a user-friendly interface. A first is to use the already acquired knowledge to provide suggestions to the user while annotating: this will make the annotation task more user-friendly [1, 2, 8].

3. AKTIVE MEDIA

AKTive Media is a user centric system for multimedia document enrichment, developed at the University of Sheffield, designed to meet the above mentioned requirements; it uses Semantic Web and language technologies for acquiring, storing and reusing knowledge.

The aim is to provide a seamless interface that guides the user through the process, reducing the complexity of the task. Languages technologies and a web service architecture are adopted to provide a context specific annotation mechanism that uses suggestions inferred from both the ontology and from the previously stored annotations to help the user: the ontology is pre-filtered to present only the top-level concepts (the most generic ones); when the users identify an instance of a top-level entity the system suggests the possible relevant sub-entities, e.g. when annotating a region of an image as a "part", the system suggests all the possible parts present in the ontology for that engine and the user can select the right one. The same happens for relations, again inferred from the ontology and suggested to the user on the base of the concept selected: for example, when the part has been chosen, the user can select a "has_fault" relation and drag and drop the text in the document that describes the fault (see Figure 1).
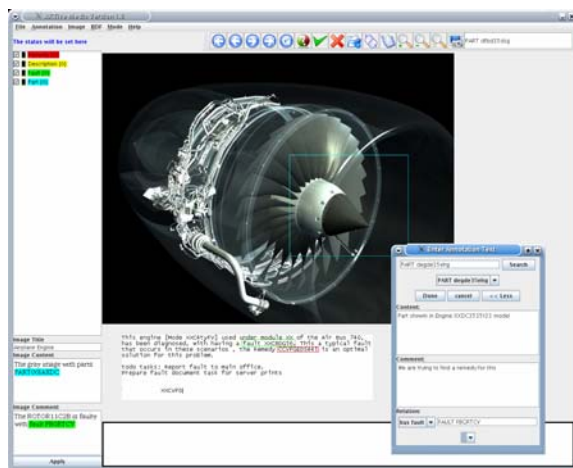

Figure 1 - AKTiveMedia interface

When an instance has been identified it is possible to add free-text comments, to highlights findings or opinions, or to state some generic or contextual information about the instance, e.g. findings on why the failure happened.

The produced knowledge is then used as a way to establish connections with and to navigate the information space: when the user annotates a part of an image as a "sand-damage" on a "turbine" the system uses those annotations to retrieve other related images and documents. New relationships can then be established with the found knowledge, e.g. the damage can be related to other previous cases, and through free-text comments the relationship may be made explicit (e.g. this type of failure happens constantly on this blade in hot conditions, and this is proved by document x). AKTiveMedia architecture uses a two steps persistence model to save the annotations: 1) the annotations are fist saved in a local repository, 2) then they are imported by a web service into a RDF central repository.

5. REFERENCES
[1] Ciravegna F., Dingli A., Petrelli D. and Wilks Y. User-System Cooperation in Document Annotation based on Information Extraction. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), 1-4 Oct 2002 - Spain, Lecture Notes in Artificial Intelligence 2473, Springer Verlag

[2] Handschuh S., Staab S., and Ciravegna F.. S-CREAM- Semi-automatic CREAtion of Metadata. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), 1-4 Oct 2002 - Spain, Lecture Notes in Artificial Intelligence 2473, Springer Verlag

[3] Kahan J., Koivunen M., Prud'Hommeaux E., and Swick R.. Annotea: An Open RDF Infrastructure for Shared Web Annotations. In Proc. of the WWW10 International Conference, Hong Kong, 2001.

[4] Kang, H. and Shneiderman, B. Visualization Methods for Personal Photo Collections: Browsing and Searching in the PhotoFinder Proc. IEEE International Conference on Multimedia and Expo (ICME2000), New York City, New York.

[5] Kuchinsky A., Pering C., Creech M.L., Freeze D., Serra B., Gwizdka J., FotoFile: A Consumer Multimedia Organization and Retrieval System, Proceedings of ACM CHI99 Conference on Human Factors in Computing Systems, 496-503, 1999.

[6] Lanfranchi V., Ciravegna F., Petrelli D. Semantic Web-based Document: Editing and Browsing in AktiveDoc, Proceedings of the 2nd European Semantic Web Conference , Heraklion, Greece, May 29-June 1, 2005

[7] Shneiderman B. The eyes have it: A task by data type taxonomy for information visualizations, In Proceedings IEEE Visual Languages , pages 336-343, Boulder, CO, Sept 1996

[8] Vargas-Vera M., Motta E., Domingue J., Lanzoni M., Stutt A., Ciravegna F.: MnM: Ontology driven semi-automatic or automatic support for semantic markup. In Proceedings of the 13th Int Conference on Knowledge Engineering and Knowledge Management, EKAW02. Springer Verlag, 2002