

# **AKTiveMedia: Cross-media Document Annotation and Enrichment**

Ajay Chakravarthy, Fabio Ciravegna,  
and Vitaveska Lanfranchi

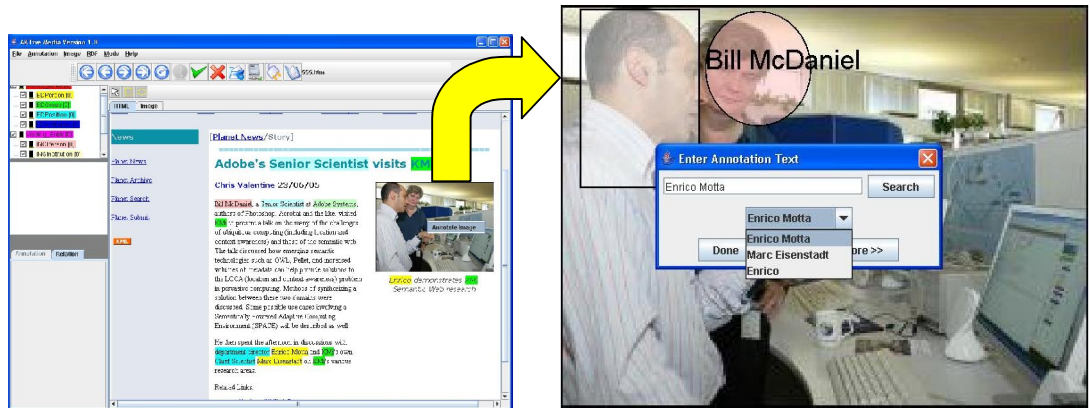
Department of Computer Science, University of Sheffield, Regent Court,  
211 Portobello Street, S1 4DP, Sheffield, United Kingdom  
{ajay, fabio, vita}@dcs.shef.ac.uk  
<http://nlp.shef.ac.uk/wig/>

## **1 AKTiveMedia**

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 textual and graphical information available in documents for effective and efficient use. For example, several activities focus on text annotation as a way to enrich a textual document, making it machine-readable and also accessible to people [1, 2, 4, 5]; other projects focus more on annotation of images e.g. [3]. However, we believe that the separation of text and images is artificial and there is a strong need for enabling true cross-media annotations that span the division of text and images. A constantly increasing number of information sources, like websites, 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.

AKTiveMedia is a user centric system for cross-media document enrichment; it uses Semantic Web and language technologies for acquiring, storing and reusing knowledge in a collaborative way, sharing it with other members of the community. AKTiveMedia's aim is to provide a seamless interface that guides users through the annotation process, reducing the complexity of their task. The system supports document enrichment through ontology-based and free-text annotations; these can be inserted manually by a user, inserted by reusing other users' annotations or suggested by the system. An Information Extraction system is integrated, learning from previous textual annotations and suggesting new ones, thus easing the annotation process. A context specific annotation mechanism that uses suggestions inferred from both the ontology and from the previously stored annotations has also been implemented: as an example consider a HTML document that has details about visits to the KMI institute. When the user is annotating the document, he will start annotating "Enrico Motta" as a visited person. If in the same document a picture of Enrico is present and the user is

annotating using the same ontology concept, the system will suggest as a description “Enrico Motta” (Figure 1).



**Figure 1 - Annotating HTML using AKTive Media**

If the suggestion is accepted by the user an identity is established between the instances in the text and the ones in the image (same URI). In case the image is annotated first, the system will search the text for descriptions compatible with those in the image. Matching is done using string distance metrics<sup>1</sup>. A sharing and reuse facility is available to view other users' annotations and comments, in order to share the knowledge inside the community.

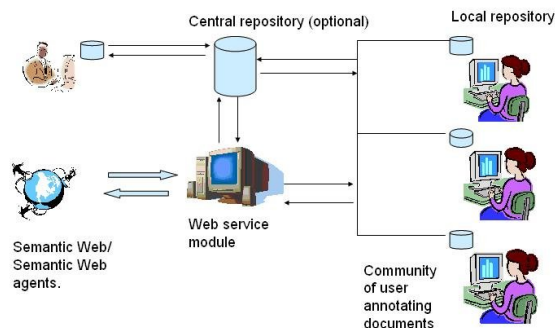
**Acknowledgments.** This work was funded by the AKT project, EPSRC (grant GR/N15764/01) and the X-Media project ([www.x-media-project.org](http://www.x-media-project.org)) (IST-FP6-026978.)

## 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),
2. 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.
3. 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.
4. 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
5. 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

## Demo Description

AKTiveMedia is a user centric system for multimedia document enrichment; it uses Semantic Web and language technologies for acquiring, storing and reusing knowledge. The aim is to provide a seamless interface that guides users through the annotation process, reducing the complexity of their task. The main modalities supported are: image annotation, text annotation and cross text/image annotation. All modalities share the same functionalities and very similar interfaces. They all offer ontology-based enrichment through a graphical interface: portions of text or images can be associated with concepts in the ontology with a straightforward point&click interface. Free-text annotations can also be added on top of the ontology-based ones, to insert more information. Every annotation is separately stored in the form of RDF file. The system is based on a configurable plug-in model in which the different components (e.g. ontology loader, annotation modalities, web services etc.) are independent sub-models that can be plugged in for creating a custom application. The architecture focuses on RDF as a way to store and query data and to communicate between components and web services and as a way to distribute the architecture. All the annotations are stored as RDF triples inside a local store and periodically updated into a central triple store using web services (Figure 1). This modular architecture has been implemented in order to allow different users to see other people's annotations and reuse them.



**Figure 1 - AKTiveMedia sharing model.**

The Demo will show the AKTiveMedia system running with some example html documents taken from KMI (Knowledge Media Institute) website. Users will be able to annotate document and images and see the Information Extraction system working, suggesting new annotations for the corpus. They will also be able to see other user's suggestions, to understand how the knowledge sharing process works. This will show how the cross-media annotation process is greatly eased and made faster by the integration of Semantic Web and Language Technologies.