

LARGE SCALE CROSS-MEDIA KNOWLEDGE SHARING AND REUSE

J. Magalhães, A. Chakravarthy, A.-S. Dadzie, L. Xia, J. Iria, M. Greenwood,
D. Petrelli and F. Ciravegna

Departments of Computer Science
The University of Sheffield
Regent Court, 211 Portobello Street, S1 4DP Sheffield, UK

ABSTRACT

This paper describes an application that demonstrates several large scale methodologies and techniques to support sharing and reuse of knowledge distributed across different media (images, documents and data).

1. INTRODUCTION

The growing size of an organisation's archives has serious implications on the way knowledge management can be implemented. Dimensions of complexity include the presence of multiple media across distributed archives, the difficulty in focusing and integrating knowledge from different sources, and the dynamic and uncertain nature of knowledge. Moreover a considerable infrastructure is needed to support large scale Knowledge Management across media.

X-Media application [1] provides a demonstration of new technologies and tools addressing the above concerns: (1) a new breed of single medium information extraction technologies; (2) multi- and cross-media information extraction technologies able to capture and reuse knowledge across media; (3) new Semantic Web based ways of sharing and reusing knowledge in large organisations.

2. X-MEDIA DEMO APPLICATION

The goal of this multimedia knowledge management application is to assist human users in their tasks by seamlessly integrating current work practices in a more efficient and effective way. The application integrates human knowledge and machine driven technologies to help humans acquiring, sharing and reusing knowledge.

The X-Media demo application integrates several technologies to support knowledge workers in an effective way, (i) hiding the complexity of the underlying search/retrieval process, (ii) resulting in a natural access to knowledge, (iii) allowing interoperability between heterogeneous information resources and (iv) including heterogeneity of data type (data, image, texts).

Next we will describe the application modules.

3. X-MEDIA ARCHITECTURE

All X-Media applications are based in a unique architecture designed as a flexible plug-in framework where functionalities have dedicated plug-ins for knowledge acquisition, knowledge fusion, and knowledge sharing and reuse. See [3] for a discussion of some of the requirements from Aerospace Engineering related applications.

Figure 1 depicts the internal architecture of the X-Media application demo. In this framework, applications use the plug-ins that they require and are responsible for orchestrating all individual plug-ins. An initial requirements analysis phase has refined the architecture design.

3.1. Knowledge Acquisition

Starting from the acquisition tools, bottom part of Figure 1, knowledge is extracted from single-media and across media. In X-Media, acquisition and extraction is performed by integrating techniques like information extraction from text, image understanding and data mining. Data, text mining, and image analysis make sense of unstructured sources of information such as numerical data, unstructured texts and raw images (e.g. results or parameters in experiments, textual documents, pictures, etc.) for discovering hidden knowledge and emerging patterns. Information is classified automatically against ontologies. Information from different sources, including the Web, and different media is integrated into a common repository (the company's knowledge base). Several methodologies, technologies and tools have been developed for (1) single medium IE technologies designed with integration within multimedia environments in mind; (2) cross-media methodologies where single medium technologies interact to provide cross-media services [2].

The strength of our approach comes from the flexible combination of media, according to the application needs.

3.2. Knowledge Sharing and Reuse

When knowledge is extracted, it must be fused, i.e. integrated across documents, media and archives. Uncertainty and dynamicity are two fundamental issues in representation and fusion:

- Knowledge is uncertain in any real environment, either because it is so in nature or because of the use of automated means of extraction. Uncertainty is represented as probabilistic knowledge.
- Knowledge in real environments is dynamic, i.e. it changes in time or in different contexts. Again this is an issue that we cope with by defining means of exploiting provenance, changes and knowledge about processes.

3.3. Knowledge Access

Finally, knowledge is made available for reuse and sharing by people or automated systems. X-Media provides dynamic and effective access to knowledge by means of flexible browsing and searching technologies, organised around a methodology for knowledge sharing and reuse.

In X-Media, this is enabled by the definition of:

- A new effective and efficient paradigm for knowledge retrieval, sharing and reuse which enable users to define and parameterize views on the available knowledge according to their needs. Users are supported by a prediction system capable of tracing the user profile and help focusing in the large search space.
- Knowledge fusion methods can help users in situation where knowledge is derived from different resources.
- Techniques able to present and manage (i) uncertainty, (ii) trust and provenance as well as (iii) dynamic aspects of knowledge, in order to help users balancing their judgment.

3. CONCLUSION

The X-Media demo application shows how to enable new and more powerful ways of organizing and managing multimedia knowledge, enabling better re(use) of proprietary knowledge and making organizations more effective and efficient. Moreover, it:

- improves the amount and quality of collected material, by supporting advanced semi automatic knowledge capture and focused knowledge retrieval;
- enables better sharing and reuse of knowledge;
- reduces the time knowledge-workers spend in managing knowledge, reducing costs and making time available for important activities;
- helps in standardizing knowledge processes, allowing an increase in productivity and a reduction in costs, enabling shorter time to market;
- and better design of new products and services.

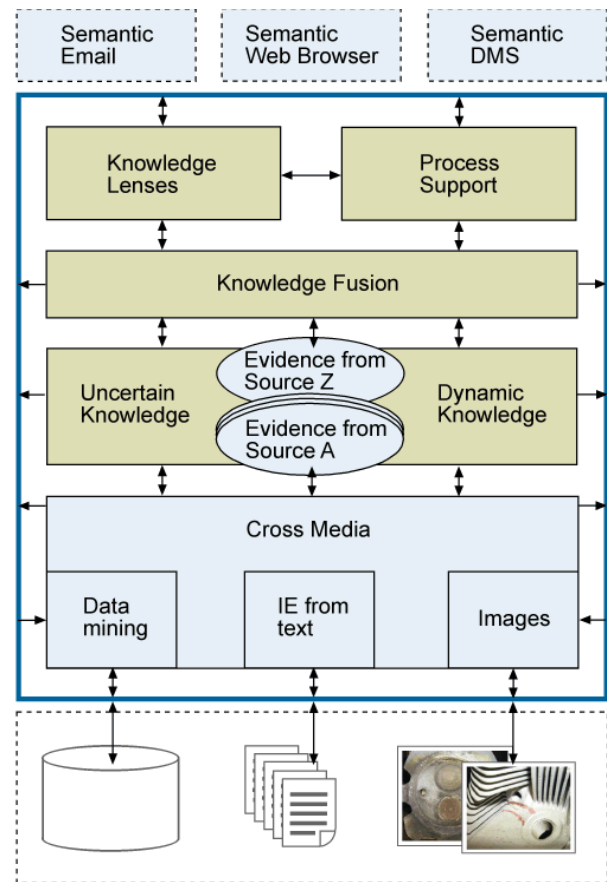


Figure 1: The X-Media Architecture

4. ACKNOWLEDGEMENTS

This work was partially funded by the X-Media project sponsored by the Commission of the European Communities as part of the Information Society Technologies programme under grant numbers IST-FP6-026978.

5. REFERENCES

- [1] "X-Media Project Web page: <http://www.x-media-project.com>," 2008.
- [2] A.-S. Dadzie, R. Bhagdev, A. Chakravarthy, S. Chapman, J. Iria, V. Lanfranchi, J. Magalhães, D. Petrelli, and F. Ciravegna, "Applying Semantic Web Technologies to Knowledge Sharing in Aerospace Engineering," *Journal of Industrial Manufacturing*.
- [3] J. Iria, V. Uren, A. Lavelli, S. Blohm, A.-s. Dadzie, T. Franz, J. Magalhães, S. Nikolopoulos, C. Preisach, and P. Slavazza, "Enhancing enterprise knowledge processes via cross-media extraction," in *Knowledge capture conference* Whistler, BC, Canada, 2007.