Annual Report 2010

The X-Media project addressed the issue of knowledge management in complex distributed environments. It carried out research into and developed large scale methodologies and techniques able to support sharing and reuse of knowledge distributed in different media (images, documents and data) across repositories. X-Media was a 4 year EU IST 6th Framework project which ran from March 2006 to February 2010 (IST- FP6 - 26978). 16 academic and industrial partners collaborated in the project. X-Media was coordinated by the University of Sheffield.

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, considerable infrastructure is needed to support large scale Knowledge Management across media.

X-Media provided the foundation for a new generation of services and tools: (1) it developed a new breed of single medium acquisition technologies more accurate and efficient than what was previously available; (2) it added on top of that innovative multi- and cross-media technologies able to capture and reuse knowledge across media; and (3) it has made available new Semantic Web-based ways of sharing and reusing knowledge in large organisations.

More concretely, X-Media has provided:

Effective and efficient new paradigms for knowledge retrieval, sharing and reuse, which enable knowledge workers to define and parameterise views on available knowledge.

Robust and scalable knowledge acquisition and data analytics tools that operate across media boundaries (text, images and data).

Novel and cutting-edge knowledge fusion methods to support knowledge workers in making decisions when confronted with possibly contradicting knowledge derived from different sources.

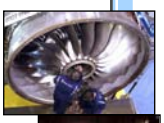
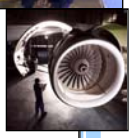
Techniques able to represent and manage (i) uncertainty, (ii) trust and provenance, as well as (iii) dynamic aspects of knowledge.

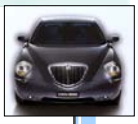
A methodology and a technical infrastructure able to cope with the complexity of large scale multimedia repositories.

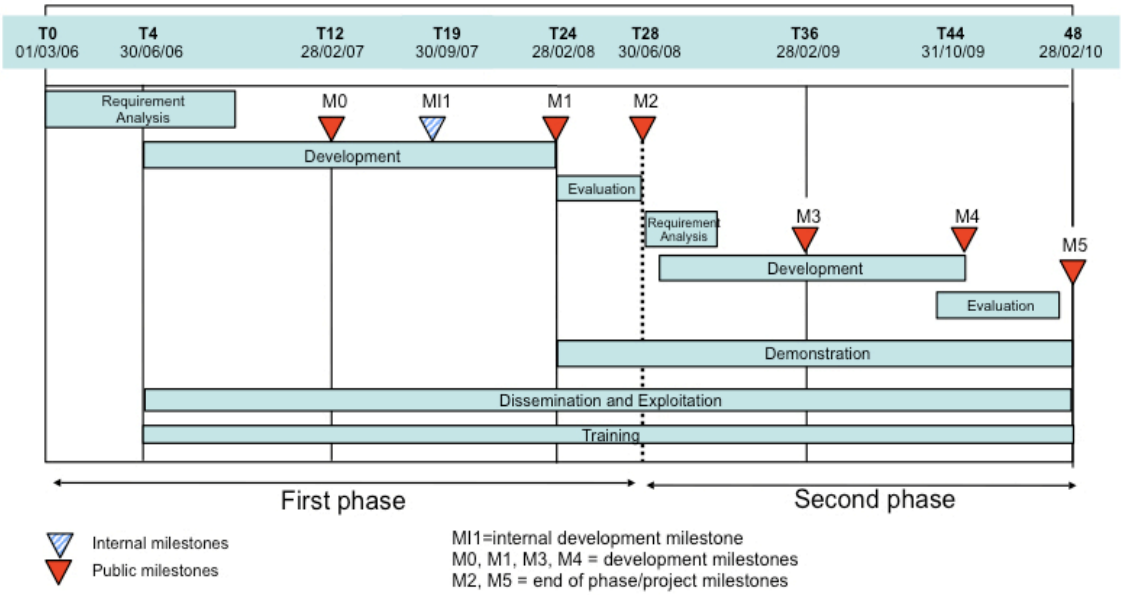
A generic and flexible architecture allowing end users to easily customise it and integrate it seamlessly with knowledge management practices.

The work done was driven by the requirements of real-world case studies in the manufacturing industry, provided by Rolls-Royce (jet engines) and Fiat (cars). Testbeds concerned product lifecycle monitoring (Rolls-Royce) and competitor analysis (Fiat). Usability was a major concern together with ease of customisation for new applications.

The ability to deal with knowledge arising from multimedia fits well environments such as the above, in which: (1) sensors and cameras provide basic data to be interpreted; (2) textual documents complement, describe, and help interpret data and images; (3) ontologies describe the domains and the applications.





Areas of Activity

The technical activity was divided into four areas:

Area 1: knowledge sharing and reuse. We defined and implemented technologies and methodologies for easy and intelligent access to and reuse of formalised and non-formalised knowledge.

Area 2: automated knowledge acquisition and extraction from documents, images and raw data. We defined and implemented methodologies and technologies for the annotation of multi-media resources (e.g. text, images, raw data), as well as the correlation of knowledge across different media; a move to largely automated

document annotation techniques was studied and implemented, with the support of knowledge acquisition technologies.

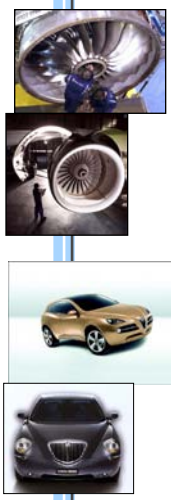
Area 3: infrastructure. We defined and implemented a knowledge acquisition, integration and sharing environment. Since X-Media was an application-oriented integrated project, integration was required at the implementation as well as at the conceptual level.

Area 4: application and testing. We designed and implemented showcases and prototype applications for the industrial testbeds. The project has been highly successful in transferring technology into the application domains, in particular via the testbeds defined by the two large industrial users in the consortium (Rolls-Royce and Fiat).

The Workplan

The work programme was iterative and organised into two distinct phases. This allowed tuning the activity where required, and ensured quality assessment of project results. Evaluation was performed both through scientific experiments and under operational conditions in the field. The two phases of the project were organised in a cycle of requirements analysis, design, implementation and evaluation.

The project has now completed its second and final phase. The first phase lasted 28 months, including evaluation. The second phase lasted an additional 20 months, during which time the output of the first phase evaluation was used to design and implement the final set of technologies, which were then successfully evaluated. The work carried high complexity because it involved integrating and developing a number of cutting-edge and emerging technologies in Knowledge Management, Information Extraction, Semantic Web, Machine Learning, Image Understanding and Interaction Design. Throughout, the consortium followed industrial best practice to control and minimise the risk, by, for example, incrementally evolving stable and tested versions of the software tools built. In some cases, tools were designed based on existing prototypes created during the first phase to support requirements elicitation. As our research progressed, these initial tools were revised and augmented to ensure that our effort was appropriately balanced between the “completed” and the “ongoing”.



Additionally, during the second phase aspects of the design that had not been fully developed were successfully implemented and integrated into the existing framework, allowing these to be evaluated as part of the integrated knowledge management system.

Summary of Activities

Following the requirements revision based on the evaluation at the end of the first phase, the last year of the project concentrated on the finalisation of the phase two prototypes and their evaluation, toward deployment in the industrial use cases.

Tool development took two forms: standalone or as part of an integrated framework, depending on the technologies in use and the target environments. Both cases allowed an assessment of the added value the X-Media project and knowledge framework could bring to knowledge management in general and specifically, to satisfy the requirements of the main industrial use cases.

The project was concluded at the end of February 2010 with a successful review on the evaluation results at the end of the second phase. This assessed the value of the research done, the improvements in the design and the technologies implemented, and finally, the contribution to the state of the art. X-Media's activities during the last year can be split into the following areas:

- Knowledge acquisition;
- Knowledge sharing;
- X-Media software architecture;
- Applications;
- Evaluation.

Knowledge Acquisition

All of the second phase knowledge acquisition systems, covering both single-medium (text, images and raw data) and cross-media areas, were finalised and evaluated. X-Media has provided:

- Multiple document formats processing;
- Technologies adaptable to several types of data and target semantic information to extract from document repositories;
- Integration of the technologies into a common X-Media architecture.
- In the first phase, the results of the evaluation were very encouraging, and were successfully translated into improved prototypes in the second phase of the project. The final evaluation of the knowledge acquisition systems was successfully carried out, both from an academic perspective and from the end users' viewpoint.

Knowledge Sharing

All of the second phase knowledge sharing systems dealing with dynamic and uncertain knowledge, fusion of knowledge and semantic approaches were finalised and evaluated. X-Media has provided:

- A metaknowledge management infrastructure;
- A text annotation tool prototype;
- Systems that are able to perform knowledge fusion;
- Systems that enable searching and browsing knowledge.
- Systems that encourage and enable reuse and sharing of knowledge beyond the initial point of creation and use. A number of methods have been made available to support end users, from text-based to interactive visualisation tools.

As for the case of the acquisition systems, the results of the first phase evaluation were very encouraging, and fed into improved prototypes that were successfully evaluated with target end users.

X-Media Kernel Infrastructure

The final software implementation of the X-Media kernel infrastructure was made available in good time in our third period enabling our other tool developers to build upon it. This stable and final version marked the culmination of more than 3 years of exhaustive consulting with all areas of the project and industrial partners, determining the right conceptual framework and software architecture for X-Media needs.

Throughout the project duration, the X-Media kernel was a crucial enabler of software integration amongst partners, and it can be regarded as one of the key contributors to a successful evaluation of the technologies. The kernel design enabled seamless communication and information exchange between the independently developed tools, enabling the construction of a user interface as a single access point to multiple tools within an integrated framework for each use case.

Applications

X-Media continued to invest very significant effort in the application of its innovative technology in the industrial partners' use cases.

For product lifecycle monitoring at Rolls-Royce:

- **Issue Resolution**: a use case that supported the identification and resolution of emerging issues in jet engines; its goals were knowledge acquisition and sharing for both the resolution of the open issue and for future reference.
- **Experimental Vibration**: management of knowledge about engine vibration in order to ensure safety of engines and avoid high cycle fatigue.
- For competitor analysis at Fiat:
- **Driver Comfort**: noise curves analysis and evaluation aimed at supporting performance engineers in setting up the best solution for aerodynamic comfort.
- **Competitor Scenario Forecast**: aimed to support the activity of publishing the calendar forecast of competitors' new cars.

Evaluation

A number of evaluation activities were successfully carried out on the industrial partners' sites, with the participation of real users performing simulations of real-world tasks. This was preceded by system tests with academic data sets, to provide independent verification of the genericity and reusability of the technologies developed. The evaluations served as a means of disseminating the results of the project to target end users, increasing the likelihood of adoption in the production environments, and ensuring the benefits reach well beyond the project's life.

Project Appraisal

The final annual review of X-Media was held in Koblenz, Germany, on the 25th and 26th of March, 2010. The review was very successful. For the fourth consecutive year, the reviewers noted that the project goals were very relevant to the IST programme objectives and that the project had proceeded according to plan and had met its objectives.

The final review unanimously rated the scientific and technical results of the project as excellent, due to its contribution to the state of the art, the exceptional exploitation activity during the project and its continued potential for further exploitation. The co-ordination, co-operation and involvement of all project partners were also lauded.

User Involvement, Promotion and Awareness

Spin-out Companies

In addition to *K-Now Technologies*, a spin-out from the University of Sheffield during the project, *jbusse.de Knowledge Engineering* (jbusse.de) was created at the end of the project by one of the principal investigators, to craft formal terminologies using OWL ontologies. jbusse.de ontologies are able to work with, among others, the Ontoprise software suites.

Two spin-out companies are to be created by KIT (formally the University of Karlsruhe) in 2011, to exploit the research done in X-Media: *Kreuzverweis* (kreuzverweis.com) exploits the user interfaces developed for the ranking of linked data, and *eNavi* to exploit the research done on process provenance.

Results Dissemination

As noted also during the final project review, a significant amount of dissemination and exploitation was carried out during the project. This is set to continue beyond life of X-Media. The main dissemination and exploitation targets were the Knowledge Management and Semantic Web communities, relevant standardisation bodies and the manufacturing industry.

Project Websites

The project web site, x-media-project.org, which will continue to be hosted by the University of Sheffield, serve as a key portal for information on and related to the project. It was regularly updated with project news and links to publications and software. A flyer that describes the project goals was regularly distributed at conferences and other events, as well as directly to interested potential customers and technology adopters. The flyer is available for download from the project web site.

Other web sites dedicated to specific tools and technologies developed during the project are hosted by respective consortium partners. These include the GIRL - General Image Representation Library - site (girl.labri.fr) and the RelEns - Rela-

tional Ensemble Classification - site (www.ismll.uni-hildesheim.de/software/relens/index_en.html).

Research Publications

The large number of and wide variation in publication targets reflects the research output of more than 30 researchers in fields from Knowledge Management to Semantic Web to Natural Language Processing to Image and Media Analysis to Information Retrieval to Human-Computer Interaction to Information Visualisation.

111 research papers had been published in journals and at major conferences by X-Media researchers by the end of the project. Publication venues included the WWW and Semantic Web conferences: ISWC, ESWC, ESTC and ASWC; the Knowledge Management conferences: K-CAP and EKAW. Other related conferences included IJCAI and ACM CHI and Multimedia. A number of working demos and posters were also presented at top international conferences, tutorials and workshops, in addition to participation in scientific competitions. Target journals included the Journal of Web Semantics and selected IEEE and ACM Transactions.

Further publications were submitted toward the end of the project, so that the final publication count should be even higher. One book, on *the Foundations of Semantic Web Technologies*, was authored by researchers during the project. Finally, a collection of articles, grouped by research area, is to be published in book form, to disseminate the research as a whole done in X-Media.

Conference, Tutorial & Workshop Organisation

Several X-Media principal investigators have been involved in the organisation of top research events, greatly increasing the visibility of the project. Prof. Enrico Motta, Dr. Yiannis Kompatsiaris and Prof. Fabio Ciravegna all served as general chairs of major research conferences. Prof. Steffen Staab and Prof. Philipp Cimiano served as a vice chairs of major conferences. Prof. Steffen Staab is also a Chief Editor of the Journal of Web Semantics. Consortium members worked collaboratively to organise tutorials and workshops at major conferences. This is in addition to the industrial workshop "*Our Vision - X-Media: Knowledge Sharing and Reuse across Media*", used to disseminate the results of the project as a whole at the 3rd European Semantic Technology Conference in Vienna in December 2009. The workshop, targeted at an industrial audience, emphasised the value of the practical application of the research done in X-Media, using presentations about key technologies developed by research partners and testimonials by the industrial partners.

Both PIs and senior researchers in the consortium were invited keynote speakers at major conferences, symposia and tutorials, and also served as tutors at key summer schools, including the annual Summer Schools on Ontological Engineering and the Semantic Web (SSSW) and Multimedia Semantics (SSMS).

The research partners also played a key role in community building, through participation in consortia such as STI (the Semantic Technology Institute) and SmaRT (Semantic MultimediA Research and Technology), and by working with decision-making bodies and EU government agencies.

The industrial partners played a key role in dissemination to the non-academic community, holding a number of internal and external presentations and tutorials, to demonstrate the output of the research activity to target sectors in industry and to state and governmental bodies. They also participated in industry-focus conferences and workshops, where they showcased the output of the project. Venture capitalists and investors in leading edge technology have been among the project's target audience.

Sponsorship

To promote awareness for the project in the research community X-Media sponsored a number of international events during its final year: SSSW 2008/9, SSMS2009, ISWC 2008/9, ESWC 2009, EKAW2008, CBMI2009, ESTC2009 and the Semantic Search Workshop at WWW2009.

Software, Standardisation & Patents

The release of over 25 software packages to the Open Source community and work done toward extending current Semantic Web and multimedia analysis standards have also widened the scope of dissemination. Two patents were submitted during the course of the project, on Hybrid Search and Terminology Recognition for the aerospace domain.

Collaboration with other Projects

A number of X-Media partners have gone on to participate in other integrated projects, including WeKnowIt, SEALS, SmartProducts, Active, MyMedia and DIGIBOOK. Software developed in X-Media is also being re-used in some of these projects. Other collaboration included NEON and ACEMEDIA, and X-Media played a leading role in the CHORUS+ Coordination action, created to foster information exchange and collaboration between EU IST FP7 projects.

Future Work

Following the conclusion of X-Media, the project website will continue to act as the main dissemination portal. The output of WP19, the public use case created to provide a small-scale demonstrator for the various technologies developed during the project, will be made available from the web site as a demonstration video. Open source software contributed to the integrated framework and whose use is seen in the video will be available for download from the project site.

The two main industrial partners, Rolls-Royce and Fiat, have identified key technologies developed during the project, which will be used as is or developed further, with a view to adoption in target communities, either as standalone tools or integrated into existing in-house software. In Rolls-Royce this is being done through the Samulet and Siloet UK TSB-funded projects. The web portal created during the project for the Fiat competitor forecast scenario will continue to be used by CRF; this serves as yet another point for dissemination and exploitation.

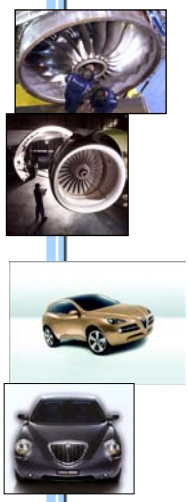
The other industrial partners will also continue to carry out dissemination and exploitation activities in their target communities, focusing predominantly on the technologies they contributed to the project.

Impact

From the start of the project in 2006, we observed, at an increasingly faster rate, knowledge being made available in all forms of media objects (data, text, audio, video, etc.) and in several contexts of use (commercial, organisational, educational, etc.). The trend in generating multimedia data to support knowledge-rich activities has intensified, and X-Media technology is in a very good position to help tackle the challenges arising from managing knowledge latent in large volumes of multimedia data.

X-Media aimed to play a fundamental role in supporting the competitiveness of European industries and in developing the European knowledge-based economy. X-Media therefore proposed to create technologies for beyond 2010, when the project was to come to an end. We have seen, as the output of the project, novel capability that more accurately and efficiently organises and manages knowledge. X-Media technology:

- Improves the amount and quality of collected and derived knowledge, by supporting advanced, semi-automated knowledge capture and focused knowledge retrieval;
- Enables better sharing and reuse of knowledge, within dedicated communities of practice and beyond these to the wider needs of a single and across other related organisations;
- Reduces the time knowledge workers spend managing knowledge, reducing cost and making valuable resources – time and human – available for important activities;
- Helps in standardising knowledge processes, which translates into an increase in productivity and a reduction in cost, enabling more effective and confident decision-making, shorter time to market and better design of new products and services.



The Consortium

The consortium was composed of research institutions and industrial partners of international standing. The research partners have specialty in Semantic Web Technologies, Knowledge Management, Information Extraction from Text, Image Understanding, Data Mining, Document Enrichment and Knowledge Sharing. The academic institutions were:

University of Sheffield (coordinator, UK), University of Koblenz (D), FBK-Irst (I), University of Ljubljana (Slovenia), University of Hildesheim (D), CERTH-ITI (G), LaBRI (F), Karlsruhe Institute of Technology (D) and the Open University (UK).

The technology providers in the project were:

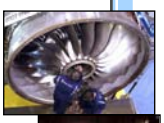
Quinary (I), Ontoprise (D), Solcara (UK), CognIT (N) and K-Now (UK).

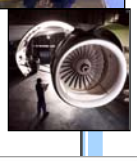
Their role was mainly to study, design and implement infrastructure to support the knowledge management methodology, and to support research and applications.

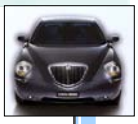
End users were:

Rolls-Royce (UK) and Centro Ricerche Fiat (I).

Their role was to provide requirements and testbeds for the knowledge management technology.







CONTACTS AND INFORMATION

Professor Fabio Ciravegna
Organisations, Information and Knowledge Laboratory
Department of Computer Science
University of Sheffield,
Regent Court,
211 Portobello Street,
S1 4DP Sheffield, United Kingdom
http://www.x-media-project.org
xmedia-coordinator@dcs.shef.ac.uk