

## Deliverable 15.6

### Final dissemination and exploitation report

Fabio Ciravegna  
University of Sheffield,  
fabio@dcs.shef.ac.uk

Christine Preisach  
University of Hildesheim  
Preisach@ismll.uni-  
hildesheim.de

Johannes Busse  
Ontoprise  
busse@ontoprise.de

Bernt Bremdal  
CognIT a.s,  
bernt.bremdal@cognit.no

Jan Erik Dahl  
CognIT a.s,  
jan.erik.dahl@cognit.no

Alberto Lavelli  
FBK-irst  
lavelli@fbk.eu

Steve Fullerton  
Solcara Limited  
steve.fullerton@solcara.com

Spiros Nikilopoulos  
CERTH-ITI  
nikolopo@iti.gr

Marina Giordanino  
CRF  
Marina.giordanino@crf.it

Sam Chapman  
K-Now Ltd  
sam@k-now.co.uk

Charlotte Wilson  
K-Now Ltd  
charlotte@k-now.co.uk

Luca Gilardoni  
Quinary Spa  
l.gilardoni@quinary.com

Clara Bagnasco  
Quinary Spa  
c.bagnasco@quinary.com

Jenny Benois-Pineau  
LABRI  
Jenny.benois@labri.fr

Jean-Philippe Domenger  
LABRI  
domenger@labri.fr

Martin Mozina  
ULFRI  
martin.mozina@fri.uni-lj.si

Sebastian Rudolph  
Karlsruhe Institute of  
Technology (KIT)  
sebastian.rudolph@aifb.uni-  
karlsruhe.de

---

Work-package: WP15

Type: Technical Report

Distribution: Public

Status: Final

Deliverable Coordinator: Fabio Ciravegna, University of Sheffield

Area Coordinator: Fabio Ciravegna, USFD

Project Coordinator: Fabio Ciravegna, University of Sheffield

EU Project Officer: Leonhard Maqua

---

## ABSTRACT

X-Media has been completed very successfully. Its dissemination and exploitation paths are a measure of its success. During the 4 years of project, we have obtained considerable results in terms of dissemination (e.g. number of papers, tutorials, workshops and invited talks) and in terms of exploitation (e.g. the creation of spin-out companies in Sheffield, Karlsruhe and Koblenz) and a number of new projects that spun out of X-Media). In this deliverable we list the final data regarding dissemination and exploitation during the project and outline the potential follow-ups of the project.

---

## TABLE OF CONTENTS

<b>INTRODUCTION.....</b>	<b>5</b>
1.1 EXPLOITATION PATHS.....	7
1.2 OPEN SOURCE SOFTWARE.....	7
1.3 SPIN-OUT COMPANIES.....	7
1.4 PATENTS AND AWARDS.....	8
1.5 FOLLOW-UP INVESTMENT.....	8
1.6 FOLLOW-UP PROJECTS.....	9
<b>2 EXPLOITATION BY TECHNOLOGY PROVIDERS.....</b>	<b>9</b>
2.1 QUINARY.....	9
2.1.1 <i>Target Market</i> .....	10
2.1.2 <i>Currently Exploited Results</i> .....	11
2.2 SOLCARA.....	11
2.2.1 <i>Target Market</i> .....	11
2.2.2 <i>Currently Exploited Results</i> .....	11
2.3 ONTOPRISE.....	12
2.3.1 <i>Currently Exploited Results</i> .....	12
2.3.2 <i>Target Market</i> .....	13
2.4 K-NOW.....	14
2.4.1 <i>Target Market</i> .....	14
2.4.2 <i>Currently Exploited Results</i> .....	15
2.5 COGNIT.....	19
2.5.1 <i>Target Market</i> .....	20
2.5.2 <i>Currently Exploited Results</i> .....	20
<b>3 EXPLOITATION BY INDUSTRIAL USERS.....</b>	<b>22</b>
3.1 ROLLS-ROYCE.....	22
3.1.1 <i>Meetings</i> .....	23
3.2 CENTRO RICERCH FIAT.....	24
<b>4 EXPLOITATION BY ACADEMIC PARTNERS.....</b>	<b>25</b>
4.1 UNIVERSITY OF SHEFFIELD.....	25
4.2 UNIVERSITY OF KOBLENZ.....	26
4.3 OPEN UNIVERSITY.....	27
4.4 UNIVERSITY OF HILDESHEIM.....	27
4.5 FONDAZIONE BRUNO KESSLER.....	27
4.6 KIT.....	27
4.7 UNIVERSITY OF LJUBLJANA.....	28
4.8 CERTH.....	28
4.9 LABRI.....	29
<b>5 CONCLUSIONS.....</b>	<b>29</b>
<b>6 APPENDIX A: DETAILS OF THE EXPLOITATION OVER 4 YEARS.....</b>	<b>30</b>
6.1.1 <i>Invited Talks</i> .....	30
6.1.2 <i>Tutorials</i> .....	32
6.1.3 <i>Conference and Workshop organisation</i> .....	33
6.1.4 <i>Book</i> .....	34
6.1.5 <i>Papers and Posters</i> .....	34
6.1.6 <i>Flyer</i> .....	42

6.1.7 *Sponsorships* ..... 42  
6.1.8 *Community Building and Clustering*..... 43

## Introduction

Critical to the success of X-Media was the effective communication and dissemination of results and the promotion of the usage of the project outputs to a wider audience, outside the Consortium. In this sense, the project delivered:

- High quality research output in terms of contributions to workshops, conferences and journals; at the end of the 4<sup>th</sup> year we have published
  - 17 journal papers,
  - 94 conference papers,
  - 1 book,
  - several dozens of demos.

A book to disseminate the final results is in preparation.

Moreover we have organised:

- 4 Summer Schools;
- 11 international events;
- 1 Industrial workshop (at ESTC09)

We have also sponsored 10 international conferences and events.

Details are in the appendix.

- Innovative knowledge management technologies and methodologies in multi- and cross-media environments; the technologies concern the acquisition and capture of knowledge (developed as part of WP3, WP5, WP6, WP7 and WP8), and the knowledge lenses developed as part of WP4.

The target audience for dissemination and exploitation was mainly:

- The Knowledge Management scientific and industrial communities; to this end we have addressed e.g. the Conference on Business Information System, and journals such as the Journal of Intelligent Manufacturing.
- The Semantic Web scientific and industrial communities; we have published at all the major conferences in the field: the European Semantic Web Conference, International Semantic Web Conference, Asian Semantic Web Conference, the World Wide Web Conference, IJCAI, ESTC, etc. as well on major journals like the Journal of Web Semantics.
- The manufacturing industry as main target of the applications and testbeds developed in the project; we have written papers and given presentations, for example the paper published on the Journal of Intelligent Manufacturing.
- Venture capitalists and investors in the field of advanced technologies as main targets for exploitation of the most innovative parts of the technology; we had meetings with major venture capitalist where we have presented the technology developed;
- Standardization bodies related to the fields of Semantic Web, especially for the parts related to the extension of existing standards to support project related activities.
- Other EU projects; natural collaboration via several partners has occurred with the NEON, ACEMEDIA, WeKnowIt and SmartProducts projects. Moreover, X-Media has had a leading role in the CHORUS+ Coordination action, which aims

at creating the conditions of mutual information and cross fertilisation between the FP7 projects, More specifically, CHORUS+ objectives are to support the integration and strengthening of the European Research Area by stimulating interaction and co-ordination on an EU level, in the area of Audiovisual search engines and extend cooperation to Asian countries. CERTH has been a leading partner in CHORUS+. Moreover, our contribution to the standardization bodies of Semantic Web and multimedia analysis has also taken place through our participation in the related scientific competitions like TRECVID2009 and VideoOlympics 2009

Content exploited and disseminated covers an ample spectrum:

- Project main goals and objectives and research agendas, as well as updates on development; we have sponsored several international conferences and events (full list below), we have distributed leaflets to all the major conferences (including ISWC/ESWC 2007, 8 and 9) and given invited talks (see section on invited keynote speeches; important for EU dissemination was the talk given at the open day “Call 3 in Motion”, organized by INFOSO.E2 Content and Knowledge, European Commission, Luxembourg, December 2007);
- Material for training; we have given several tutorials at all the major conferences (ESWC, ISWC, ASWC, etc.) and Summer Schools (SSSW06/07/08/09, SSMS08/09, etc.);
- Technology investment opportunities, novelty issues, etc. (e.g. the open day at ESTC 2009);
- Single technologies: one patent on Hybrid Search has been submitted and one on Terminology Recognition is under submission.

Dissemination activities were centred mainly on raising awareness and knowledge, solution and technology transfer. Different activities were implemented for different target groups (e.g. industry and academia), as they differ in terms of literacy, cultural background and nature of objectives. Therefore the consortium implemented a range of dissemination activities and vehicles adapted to the characteristics of diverse, targeted audiences, such as:

- Seminars, Conferences and Journals: papers and presentations;
- Participation in major industrial events with presentations (several conferences including European Semantic Technology Conference in Vienna in 2007, 2008 and 2009), tutorials (Norwegian Semantic Days) and with dissemination material (several events). Booth presentations were a vehicle of dissemination and exploitation as well (ESTC 2009);
- Workshops were organised to invite the international community to discuss issues of interest for the project; Example is the SemSearch workshop, now at its third edition;
- Public Relations events: we have organised an open day/dissemination event targeted to industrial users at the European Semantic Technology Conference 2009 in December in Vienna;
- More than 20 potential Users’ Focus Groups were organised at Rolls-Royce and Fiat to disseminate project results and foster exploitation;
- Events organised by the European Commission (e.g. the already mentioned “Call 3 in Motion” event in Luxembourg).

The Web was a main dissemination vehicle; we have created a Web site including an internal intranet for the project participants (K@ developed and hosted by Quinary) and an external web for visitors (x-media-project.org hosted by Sheffield).

promote exploitation. Among these we can list:

- joint development and support of the X-Media architecture (lead by Quinary), with commitment from the consortium to make the infrastructure available for exploitation and dissemination;
- joint development of a public use case (covered in WP19, Bike Brakes) to help disseminate the X-Media technology and approach.

### 1.1 Exploitation Paths

Commercial impact of X-Media is first of all in technical domains such as the manufacturing industry, the automotive and the aerospace domain. The support of engineering processes (maintenance, design, manufacturing, competitor analysis) is a major category for X-Media-based solutions.

Exploitation paths are described below for the industrial partners. They include:

- Direct usage of project results by end user's organization (Fiat Group and Rolls-Royce) in new projects and activities; for example, Rolls-Royce is currently investing funding as follow-up of some of the X-Media technologies in order to reach the Technology Readiness Level 6 which is a legal requirement for use in an aerospace company.
- Spin-out of derived technology – whole or parts - from the industrial and academic partners. The industrial partners detail their plans in the sections below.

### 1.2 Open Source Software

The distribution of a considerable part of the software as open source, helped disseminate project results towards interested users and potential uptaker of the technology. At this point in time **more than 25 software packages have been released as open source**. Some of them have reached a very high level of uptake. For example the SimMetric Library is currently the most downloaded distance metric library on the Web, with over 30,000 downloads and an average of 700 downloads a month for the last year from prime site (SourceForge.net) and over 18,000 more from six other known mirror sites<sup>1</sup>.

### 1.3 Spin-out Companies

X-Media has been very successful in exploiting the generated results. One measure of success is the number of spin-out companies that have been generated. We have created two spin-out companies and another two are in the making. They are:

- **K-Now Ltd**, a spin-out company of the University of Sheffield who is commercialising X-Media technologies in the area of corporate search, knowledge capture by distributed communities and similarity matching (the later ones used mainly for spend analysis in large corporations). Current customers are – among others - Rolls-Royce, KPMG, Deloitte, Schroeders Bank, Coles, Adelle, Associated British World Foods, Granger, Silver Spoon, Allied Mills, Ryvita Jordon, GAME, Allied Bakery, Kew Gardens. K-now was spun out quite early in the program (2007-8) and has become a full

---

<sup>1</sup> A comparison with download figures of SimMetrics's bigger competitor software, SecondString, shows how Second String total 7,000 downloads since 2004 and how its rate of downloads and usage has dropped since the release of SimMetrics.

partner in X-Media since 2009. Their exploitation plans are detailed in the next section.

- **Dr. Johannes Busse Knowledge Engineering** (jbusse.de): it is a bureau for ontology engineering, i.e. for crafting formal terminologies. The company is specialized on pruning OWL full ontologies to less complex profiles (like OWL2 RL) such that the result can be executed fast and in large scale with rule engines like Ontoprise's OntoBroker. A typical customer is a knowledge intensive business which wants to (a) carve out, formalize and cast special knowledge into individualized ontologies; (b) document and refine own terminologies; (c) enable its staff to maintain and extend the corporate ontology on their own responsibility; (d) document the results of these processes in an comprehensible and accessible way. As for exploitation of X-Media results Johannes Busse has committed to maintain the tool mm2flo. (The tool was prototyped in WP4 and published under GPL licence in 2009 on the Ontoprise website.) An RDF/turtle version of the tool was redeveloped completely in 2010 and published under <http://semauth.sourceforge.net> (GPL licence). The tool semauth is used for empowering business end users to contribute in all phases of ontology engineering maintaining. While the RDF version of semauth is free, a more sophisticated F-Logic version is given for free to customers of Johannes Busse.

Another two companies are about to be created:

- **Kreuzverweis** (kreuzverweis.com), a spin-out of the University of Koblenz-Landau. They exploit X-Media technology developed in workpackage 4 on user interfaces and ranking of linked data. Kreuzverweis will develop software for media management and has been successful in receiving funding from the German EXIST programme. The foundation of Kreuzverweis GmbH is planned for 2011.
- An additional spin-out of the University of Koblenz-Landau will be launched in the beginning of 2011. The spin-out will provide the software "eNavi", which is navigation software for IT processes. The software re-uses technology developed in workpackage 1 on process provenance. A proposal to receive funding from the German EXIST programme will be submitted by end of June.

#### 1.4 Patents and awards

The project produced two patents:

- An international one was filed by the University of Sheffield, "Searching Methods and Systems" (GB2449501, EP2149097 and EP2149097) patenting the Hybrid Search methodology used in K-Search [http://v3.espacenet.com/publicationDetails/biblio?DB=EPODOC&adjacent=true&locale=en\\_gb&FT=D&date=20100203&CC=EP&NR=2149097A1&KC=A1](http://v3.espacenet.com/publicationDetails/biblio?DB=EPODOC&adjacent=true&locale=en_gb&FT=D&date=20100203&CC=EP&NR=2149097A1&KC=A1).
- An international one is being submitted by Rolls-Royce and University of Sheffield on Terminology Recognition for Aerospace.

The above X-Media technologies were shortlisted for the Rolls-Royce Director of Research's Creativity award in 2007 and 2009.

#### 1.5 Follow-up Investment

One of the two industrial users have committed significant funds to further support



and exploit some of the X-Media technologies. Rolls-Royce is currently funding K-Now Ltd and the University of Sheffield to the tune of £560k. The projects are detailed in the next section.

## 1.6 Follow-up Projects

The following projects are direct follow-ups of X-Media because they reuse technologies and even contain a subset of the X-Media consortium:

- [SmartProducts](#), an FP7 European Integrated Project on embedding proactive knowledge into Smart Products to help their design, manufacturing, servicing and use. Coordinator is SAP, users are Philips, Fiat and Eads. Three X-Media partners are involved: University of Sheffield, Centro Ricerche Fiat and Open University (2009-2012);
- [WeKnowIt](#): an FP7 European Integrated Project on Collective Intelligence. User partners are Yahoo!, Telefonica and Sheffield City Council Emergency Services; X-Media Partners involved: CERTH (coordinator) University of Sheffield, and Koblenz (2008-2011).
- [SEALS](#): Semantic Evaluation At Large Scale: an FP7 European Project on Evaluation of Semantic Web Technologies (2009-2012).
- IPAS XWB: a project to apply Knowledge Management Technologies to the design process of the jet engine XWB (the future Airbus A350's engine). The project involved Rolls-Royce and University of Sheffield and was fully funded by Rolls-Royce (2008-2009);
- [Samulet](#): a £90m research project aiming to revolutionize the lifecycle of aerospace engines to address environmental emissions. It is funded by TSB in the UK and Rolls-Royce, and involves University of Sheffield and Rolls-Royce (2009-11);
- [Siloet](#): a £90m research project on enhancing manufacturing methods funded by TSB in the UK and Rolls-Royce. It involves the University of Sheffield and Rolls Royce (2010-12);

## 2 Exploitation by Technology Providers

### 2.1 Quinary

Quinary is a system integration company, which mostly works with big or mid-sized corporations. Its main business is in supporting customers in the development of their IT infrastructure to meet goals, in such areas as Telecom, Media, Utilities and Public Administration. The main value Quinary is proud to pass on to their customers is in being able to stay at the forefront of the technology and in being able to understand how technology could be applied. Research activities in X-Media naturally fall in this scenario, as they offered us the chance to experiment and develop new technology on our own, to develop a network of relationships with advanced technology providers from the Consortium and to share experience in applying technology to concrete use cases. Quinary is not a 'product company' (such as, e.g. Ontoprise, who provide a range of services alongside a proprietary product line), hence exploitation from Quinary is expected mostly by reusing such experiences within customer projects. Nevertheless, on the one hand a niche vertical product exists, targeting the legal market, where by-products of X-Media are expected to find a place, and on the other hand Quinary strongly believes in partnering with pure technology providers and in exploiting open source solutions. To this extent, Quinary will actively support and

maintain the X-Media kernel developed within X-Media, both as a standalone package and in connection with technology components developed by Consortium partners, to exploit services based on the technology baseline.

Last but not least, we need here to mention the possibilities opened by the fruitful collaboration with the Consortium partners. Cooperation is expected with both technology providers and research institutions. In particular, Quinary shares with Solcara an active interest within the legal market, with complementary skills and technology – other than with a diversified geographical customer base, which may easily lead to technical and commercial fruitful cooperation. The same applies to CognIT within the Oil and Gas industry, while semantic products from both Ontoprise and K-Now may open sales or referral opportunity in Italy. As for research institutions we are eager to cooperate further to take technology to the market and extend research results.

### **2.1.1 Target Market**

As mentioned, and as also detailed in [D16.1], one primary target market for exploitation is the legal sector. Quinary is already active in the field with a proprietary Knowledge Management solution for law firms and legal organisations (LKMS/Mnemosyne, see <http://mnemosyne.quinary.it> and [D16.1] for more details). Mnemosyne/LKMS already embeds a semantic layer, and exploits IE techniques to extract meaningful information from legal documents to fill the semantic layer, used in turn to provide a knowledge rich support to search and sharing. Part of X-Media technology has already found its way into the system as further discussed in the next section, and demonstration and dissemination activities are in place (see also [D16.3]); however, whilst features related to semantic representation on the one hand and to information extraction on the other are directly applicable in the field, needs for cross media analysis are not so relevant, at least in everyday use for legal applications. This is also true in other fields where we were already able to partly take advantage of X-Media technologies; one example, also discussed later is in providing richer information structures for searching and providing access services to complex E-Commerce sites.

Handling of images, and cross media, instead becomes more critical in the second target market we plan to exploit, which includes utilities and oil & gas industry. We already have active relationships with major Italian players in the field, namely the Italian handler of Electricity infrastructure Terna (and related services and R&D entities, such as Cesi and Erse – again see [D16.3]) and Eni, the Italian Oil Major. In either case, engineering and maintenance activities share several commonalities with the issues attacked by X-Media technology and coped with by the X-Media use cases developed at Rolls Royce and CRF. We expect to actively boost activities with these kind of organisations (the two specifically mentioned are among the main Italian players and we already have links in place, but others will be an obvious target) in these areas, exploiting X-Media technology to provide a leap from currently used support system, which mostly lack any capabilities for integrating data, text and image analysis. Both the issue resolution use case from Rolls Royce, as well as the EV one and the Noise Reduction one from CRF may be used as examples to draw upon, and the underlying technologies are immediate candidates for supporting new applications.

The third target market we are currently considering is the Health sector. In this area there is an evident need to be able to relate data, text and images when handling patient data (e.g. consider the management of clinical records), and we see compelling

evidence that X-Media technology may help significantly. Our assessment of our possibility to concretely enter this market here is still very preliminary; however we have already started discussing on the one hand with medical systems providers (Siemens) and on the other, with health organisations (which in Italy, generally means the public sector – and in specific the local governments – Regione Lombardia) about potential collaborations for exploitation.

### **2.1.2 Currently Exploited Results**

Most Quinary activities in X-Media were centred around the kernel and underlying infrastructure, including handling of compound documents and in sharing tools (scratchpad).

Parts of the X-Media kernel have already been used in the latest versions of our Mnemosyne/LKMS legal knowledge sharing environment. Work on the scratchpad and underlying ideas has partly been transferred in a preliminary version of a wiki component extended with semantic attachment embedded within Mnemosyne/LKMS since 2009, with further extensions currently being made.

Part of the infrastructure is also being used within a large project with an Italian provider of normative technical documents, where Quinary is developing an e-commerce service to sell such documents (with various modalities – one offs or subscribed); in this framework part of the X-Media kernel infrastructure has been adapted to support indexing and extraction services from the document base to provide advanced search and interlinking capabilities.

## **2.2 Solcara**

Solcara will continue to evaluate semantic search, information extraction from text and linked data integration technologies to enhance Knowledge Management and Enterprise Search in key markets. They will look to productise suitable technologies either alone or in partnership with other consortium members or clients and look to extend such technologies through further Research & Development grants and/or Knowledge Transfer Programs.

### **2.2.1 Target Market**

Law firms, legal information integrators and content providers, pharmaceutical research, the media and communications sector (including news aggregators) as used by police, local and national government bodies, the health sector and finally in police and security applications (see [D16.3] for details of dissemination to those sectors).

### **2.2.2 Currently Exploited Results**

As a first step, Solcara will announce in Quarter 2 of 2010 the launch of a new commercial product – PRISM (Precise Retrieval of Information using Semantic Methods) – built on the semantic enablement of the existing SolSearch federated search tool as used as an integration technology in the Rolls-Royce Experimental Vibration Use Case. The semantic federated search interface is currently available for demonstrations to existing and potential clients and will be shown in the Final Review Meeting, having previously been shown at ESTC2009.

The key features are:

- PRISM is integrated within a popular RDF repository, OpenRDF Sesame, which effectively adds federated search functionality to its query processing without having to extend the SPARQL language. PRISM will only invoke a federated search when called for by the SPARQL query, any unrecognised elements of the query are passed through to Sesame for processing as normal.

- PRISM uses RDF descriptions of federated search connectors that map their input and output properties to RDF predicates. The search connectors that are described in RDF can be targeted to both Internet and Intranet resources. This enables users to concentrate on what to search for rather than where the data resides.
- The RDF descriptions can be mapped to particular data types returned by a search connector's underlying resource.
- PRISM enables non-RDF published sources to participate in SPARQL queries.
- PRISM negates the need for resources, such as internal application databases to be aggregated into the RDF repository.
- PRISM allows for data that is not under the control of the client organisation, such as public or subscription resources to be included in SPARQL queries with local data.
- PRISM will dynamically create federated searches based upon the input and output predicates in each triple group of a query, enabling multiple sources to be searched simultaneously and their results to be included in the SPARQL result set.

Where triple groups reference variables from other triple groups in the same query, PRISM will generate multiple, iterative federated searches as required by the query processing to complete the result set. This powerful feature enables the output from one or more federated sources to automatically be used as input to other sources.

PRISM's use of a mature federated search engine, SolSearch, means that sources being queried can be protected from being overloaded with too many requests using application configuration options. The use of SolSearch's connector interface means that internal and external resources can be interrogated in flexible, efficient ways taking advantage of APIs, XML output if present or emulating a browser user to navigate a search query form and results pages if required.

PRISM features are suitable for Linked Data Exploitation and in that context Solcara are looking at the newly published UK Government information sources at <http://data.gov.uk> in order to define potential applications.

It also has the potential for incorporating IE and Knowledge Lens tools from WP5 and WP4 to extract, link and visualise entities and concepts within returned result sets (proof of concepts being designed).

Solcara will also look to integrate with and compliment technologies being exploited by K-Now and the University of Sheffield in the UK, and CognIT in Norway.

Finally Solcara will continue to work with Rolls-Royce, if applicable, to take the Use Case systems (or subsets of) created during the project to the required Technology Readiness Levels for potential production environments.

## 2.3 Ontoprise

Additionally to providing ontology engineering services and other knowledge-oriented tasks, Ontoprise focused on developing two material products (i.e. Software) in X-Media: Crafting an RDF store with sophisticated inferencing capabilities, and prototyping an innovative text annotation and ontology outlining tool.

### 2.3.1 Currently Exploited Results

#### **RDF store.**

Pure RDF stores share the shortcoming that (beside trivial RDFS) no inferencing is supported. X-Media developed an RDF store which allows for sophisticated rule

based inferencing. Ontoprise released the store as a product in summer 2008; it was successfully deployed to customers i.e. from automotive and telecommunication industry. As a result of the evaluation of phase 1 it became clear that in the long term a semantic representation is required which smoothes (or allows for workarounds for) the differences between RDFS, OWL and F-Logic modelling patterns.

Starting with 2009 Ontoprise piped the experiences from the RDF store development into the definition of F-Logic2, which is called now Object Logic. While ObjectLogic bases on F-logic from a syntactical point of view, its internal data model is significantly redesigned in order to meet the RDF data model. Object Logic now is well aligned with RDF and the graph model of the OWL2 RL profile. As a consequence OB6 will transform RDF, OWL2 RL and RDF(S) ontologies to each other without extra effort.

As a result the OntoBroker RDF store which was developed in X-Media phase 1 is now included seamlessly into OntoBroker 5.2. Early research results of X-Media have influenced the definition of the new language significantly, thus has become integral part of the Ontoprise technology stack.

### **Annotation tool mm2flo**

The functional prototype "freemind to F-logic" (mm2flo) allows end users – i.e. domain experts – to author semantic models in a tree-based editable graphical visualization. The tool is already used in projects of the FhG IITB Karlsruhe and various projects in Karlsruhe University.

Being a light weight tool with a flat learning curve, mm2flo can be used easily in knowledge elicitation processes (i.e. ontology workshops) from users who are not trained ontology engineers.

While in phase one of X-Media the tool translated a freemind map into RDF, OWL and F-logic separately, the current version focuses on writing F-Logic (resp. ObjectLogic) in order to let OntoBroker 6 do the transformation from or to OWL and other languages.

### **2.3.2 Target Market**

The market addressed by OntoBroker RDF store is characterized by the strategic decision of companies to foster RDFS (as opposed to OWL or F-logic) for semantic modelling. Ontoprise RDF store addresses customers which have decided to work with the RDF(S) technology stack, but don't want to abandon the option to perform sophisticated inferencing.

Starting with OntoBroker 5.3 the Ontoprise RDF store is not a stand alone product any more. Instead it is realized as a lightweight view on full OntoBroker inferencing engine. Thus the Ontoprise RDF store allows for scaling inferencing complexity seamlessly from RDFS over OWL2 to fully rule based large scale reasoning without having to change technologies. With respect to the Ontoprise technology stack the OP RDF store serves as a low-threshold entry into nontrivial rule based inferencing.

The market addressed by mm2flo is characterized by the end users of the tool. While formalizing ontologies with expert tools like Protégé or OntoStudio used to be a task for well-trained professionals only, a lightweight approach like mm2flo also allows for calling domain experts earlier and simpler into ontology engineering tasks. This is an important step to facilitate semantic approaches within complex corporate knowledge management projects.



## 2.4 K-Now

K-Now is a University of Sheffield spin-out company (direct result of X-Media activities) that is applying X-Media technologies in the area of knowledge sharing and reuse, with a focus on corporate search and similarity analysis.

K-Now has developed and refined semantic-based tools that are based on X-Media technologies to support the knowledge management needs of complex organisations. K-Now solutions enable dynamic communities freedom in organising their knowledge repositories, while empowering immediate and effortless sharing and reuse across the rest of an organisation. Moreover, as different parts of the organisation access organisational knowledge using different point of views, K-Now technologies enable access to shared data, information and knowledge using different point of views, making reuse simpler and more immediate.

K-Now solutions cover the whole knowledge life-cycle, from modelling to acquisition and capture, to sharing and reuse and all technologies are semantic based, exploiting the power of the Semantic Web, Web 3.0, Information Extraction from Text as well as traditional approaches.

### 2.4.1 Target Market

K-Now technologies have been applied to different verticals, amongst which:

#### **Service Based Industry: Customer Retention and Feedback Management**

Within service based industries customer retention is often key to success. A single bad customer experience can lose a customer unless properly handled. In large organisations the problem of feedback becomes more acute as feedback must be handled at the point of relevance and be considered globally to an organisation.

In large organisations where multiple services are offered each collating a variety of essential feedback the management of which becomes a complex task. Managing specialised feedback for each service is fundamental for smooth operation of each specialised service also globally it remains important to be flexible gaining a global insight across an organisation, to help address issues and prioritise resources.

K-Now provides a dynamic way to collect feedback from customers whilst still linking relevant knowledge across services, thus allowing instantaneous searching, analysis and direct handling of customer feedback, locally or globally. K-Now gives much needed business Intelligence to the management of feedback empowering greatly improved customer retention and interaction.

#### **Aerospace: Knowledge Support**

Aerospace manufacturing, service and design is a complex task where designers utilise previous designs, countless reports from worldwide service events, manufacturing reports and component analysis in order to address design needs. Designers must access relevant documents (or the knowledge in them) to create or optimise component parts in very complex interconnected devices. K-Now aids a large multinational aerospace manufacturing firm by capturing knowledge either at generation or in legacy documents regardless of worldwide sources, formats or varying terminology. K-Now technology enables searching, reuse and analysis of knowledge across the whole organisation.

#### **Procurement and Retail: Spend Analysis**

With today's economic climate it is increasingly important to scrutinise company expenditure. Within any large organisation differing business sectors, locations,

subsidiaries or acquired businesses frequently purchase similar goods but without wider knowledge of potential savings which could be made across the organisation. In theory economies of scale in larger organisations implies that savings should be made in bulk buying of materials or services with strategic purchasing and longer-term contracts. In practice however these potential savings are rarely employed as large organisations struggle to align procurement data across increasingly segmented organisations. The task of Spend Analysis aims to bring together disparate corporate knowledge resources regarding procurement in such a manner to provide means to enable economic savings.

Spend analysis techniques adopt different technologies: from Machine Learning approaches requiring prior annotation, domain dependent rule based systems, to simple interface alignment support. The most widely adopted approaches however use a classification approach by collating procurement information in an unique datastore and categorising each item of spend to a specific category in a predefined schema. This categorisation sometimes utilises a pre-defined organisational hierarchy or ontology but more commonly employs an industry agreed taxonomy for categorisation to enable wider interoperability of spend data.

#### **2.4.2 Currently Exploited Results**

In more detail, K-Now has developed tools to manage knowledge in large organisations such exploiting five types of X-Media technologies:

- Form-based ontology-based knowledge acquisition and capture (K-Forms);
- IE-based knowledge acquisition for legacy data (K-Extraction);
- User-centred search of ontology-based annotated documents (K-Search)
- RDF store based on Sesame (K-Store);
- Similarity-based matching and integration (K-Similarity).

#### **K-Forms**

K-Forms enables knowledge acquisition and sharing at the point of its creation via easy delivery of Web-based forms.

Forms are a natural way for capturing knowledge in most communities. Unfortunately very often forms are designed using MS-Word or similar tools often copied and sent to the individual users to fill. This process creates isolated data silos, without limited linkage to pre-existing knowledge.

To make matters worse prior captured documents are unstructured and therefore the contained knowledge is not easily shareable. Moreover it is very difficult to impose a quality check on the filled content of such unstructured documents.

K-Forms can be designed using a simple user interface and released over the Intra/Inter-net with one click. Forms can be filled directly by connecting to a server via a networked device. When a form is filled, it is immediately available over the Intra/Inter-net, its content can then be reused via K-Search.

At design time it is possible to selectively reuse parts of existing forms. When parts are reused, a formal connection is created automatically between the shared parts. Shared parts of forms create a network of interconnected semantic knowledge collating valuable knowledge across an organisation.

#### **K-Search**

K-Search is a powerful enterprise search solution for seamless access to structured and unstructured knowledge across an organisation.

It uniquely enables flexible searching within the same query:

1. Semantic concepts (i.e. unique values in conceptual fields)
2. Keywords in the context of captured knowledge (values in specified conceptual fields containing a specified set of keywords)
3. Keywords (search over all the resources)

K-search enables searching across repositories that share knowledge, i.e. it enables searching multiple (and partially) connected repositories.

K-Search enables:

1. searching for documents over an intranet
2. summarising information via graphs
3. exporting information to spreadsheets

### **K-Store**

K-Store is a persistent store of organisational knowledge. Uniquely K-Store uses the power and versatility of a triple store combined with standard keyword index in a novel single repository. This storage repository is designed to store knowledge suitable for using K-Forms, K-Search or K-Extraction but can be used across organisations to store combined conceptual and keyword knowledge.

### **K-Integrate**

K-Integrate is a means to collate organisational knowledge from across a variety of sources. K-Integrate works by automatically examining knowledge resources and aiding integration tasks between structured data. K-Integrate is based on two sub technologies: K-Classify and K-Similarity. These are used to segment and analyse knowledge to maintain a consistency of knowledge across an entire organisation.

K-Classify allows automatic classification of text to organise and automatically sort your knowledge. K-Classify can be applied to textual information regardless of its source allowing automatic organisation and sorting of your information. K-Classify is:

- fast to fit in with existing business processes;
- automatically updatable;

K-Similarity is a means to automatically determine the similarity of information between tested resources. K-Similarity uses a variety of research techniques taken from DNA analysis, fraud detection, database integration, statistics and puts them into a single easy to use framework. This framework contains a means to automatically, given examples, identify the best method to identify similar information.

These techniques are used for a range of tasks from spelling correction, terminology recognition to plagiarism detection automatically determining similar information.

These tools have then been exploited in other research projects or for commercial customers, detailed below.

### **Talkback**

Talkback360 is a University of Sheffield project aiming to empower the use of collated knowledge by facilitating the analysis and exploration of the data and automatically linking relevant knowledge across an enterprise.

This project brings together a range of departments and separate business units to collectively share capability and where sensible operational knowledge between partners:

1. U-Sport



2. Department of Accommodation and Campus Services (ACS)
3. Department of Estates
4. Department of Finance
5. Department of Human Resources
6. K-Now

The role of K-Now in the project is to tailor K-Search and K-Forms to the individual partners needs, in order to provide means to quickly and easily create feedback forms, search across them and integrate knowledge collected by the different department into a unique, searchable knowledge repository. The partners are provided with advanced semantic search mechanisms and interfaces to easily retrieve and share knowledge.

### **Grassportal**

Online digital information about the world's grasses is currently distributed across various suppliers of information or collated into dated physical repositories.

The University of Sheffield department of Animal and Plant Sciences, in collaboration with the Royal Botanic Gardens, Kew, is leading a JISC-funded project producing a web portal, GrassPortal<sup>2</sup>, offering public access to evolutionary and ecological data on the world's grass species, greatly expanding the services currently offered by Kew.

GrassPortal will allow any of the 11,000 species of grasses to be defined by its geographical range, climate preferences, and evolutionary relationships to other species. This new online resource is at the forefront of a digital revolution in biodiversity research, providing a one-stop-shop for resources on each species that are currently dispersed around the globe.

Grasses are present on every continent on the planet, and play major roles in the global economy and ecosystem. They account for most of our staple food crops and livestock feeds, including rice, barley, wheat, maize, millet and sugarcane. Grasses are the main winners from deforestation, and form grasslands significantly influence our climate by altering the cycling of carbon and water between the land and air.

Understanding the evolutionary and ecological relationships between species is fundamental for biology, and there is tremendous current interest in rolling out the same scientific approach used by GrassPortal to cover all of the world's half million plant species during the coming decade.

GrassPortal will offer scientists tools for easily accessing and analyzing data on an unprecedented scale, and will be particularly useful for researchers in the fields of climate change, conservation biology, and evolutionary biology. Educational case studies and sample datasets in these areas will also be packaged as research-led teaching resources for university and school students.

GrassPortal is a collaborative project involving plant and computer scientists from the University of Sheffield, K-Now and the Royal Botanic Gardens, Kew. Its international partners are the University of Lausanne, University of Washington, Brown University and the University of California, Santa Barbara.

K-Now's role in the project is focused on collating and organising knowledge using similarity technologies and in providing all technology and interfaces for sharing and reusing the collated knowledge.

### **Samulet**

SAMULET (Strategic Affordable Manufacturing in the UK through Leading

---

<sup>2</sup> For more information: <http://www.grassportal.org/>

Environmental Technologies) is a Rolls-Royce-led collaborative programme in a consortium alongside other high profile manufacturers, small and medium sized enterprises (SMEs) and several of the UK's top universities to accelerate the development of manufacturing and product technologies. It will focus on productivity and environmental improvements, including efficient advanced manufacturing processes and lower engine fuel consumption. The programme will be closely linked with the advanced manufacturing research centres.

The consortium partners and sub-contractors are: BAE Systems, GKN, Tacit Connexions, Granta, K-Now and BERU F1. The UK Technology Strategy Board is investing £28.5m in the programme and the EPSRC, £11.5m. Further support is under discussion with regional bodies. Total cost of the project including industry investment is around £90m.

SAMULET will focus on productivity and environmental improvements including reductions in raw material usage, efficient advanced manufacturing processes and lower engine fuel consumption. The programme aims will be achieved by developing new technologies and delivering a number of knowledge transfer initiatives, it will be closely linked with the advanced manufacturing research centres (in Sheffield, Glasgow, and Ansty near Coventry) so strengthening the position of UK aerospace manufacturing and its supply chain.

K-Now role in the project focuses on providing knowledge capture and sharing methodologies, developed in tools such as K-Forms and K-Search and in upgrading and customising such tools to reflect the needs of the domain under investigation. The project aims to deploy K-Now's technologies for active use in the aerospace industry to help address these issues.

### **Spend360**

Typically spend analysis can identify 5-10% in potential savings in large organisations but the large scale, inconsistency and variability of spend data makes it very hard and time consuming for a company to actually prepare data for the analysis: moreover it is of fundamental importance that the analysis is done frequently and quickly to maximise potential savings. A delay in the analysis may mean critical decisions cannot be taken. Many organisations typically rely on data-warehousing solutions, or periodic, incomplete, proprietary transactional reports from multiple Enterprise Resource Planning (ERP) systems, therefore introducing delays in the collection of, and inconsistencies in, the data. Critical decisions are then made, based upon erroneous information and savings may well not be maximised.

Looking at procurement market analysis, it has been discovered that more than 80% of organisations do not have comprehensive, accurate, repeatable spend analysis capability. (source Aberdeen Group) and consistently over the last three years, on average, over 51% of organisations have reported that the volume and complexity of data sources, the inability to extract data, and its poor quality is inhibiting their ability to do spend analysis (source Aberdeen Group). For this reason specialised IT approaches have been developed for collecting, categorising and analysing spend data. Usually spend analysis systems are rule-based, with data extraction rules, transformation rules and classification rules. Such rules-based integration solutions are costly, inefficient and time consuming when dealing with large scale, variable and inconsistent data.

As large-scale retail companies can generate in excess of a million procurement items in a single day this proves a huge challenge for most Machine Learning Classifiers. This is made more problematic as automatic classification approaches typically rely

upon trained data representative to the future data relevant to classify. With procurement data classification this reliance is less certain as the nature of procurement is very changeable over time with external influences altering spend patterns. For example alterations in economic/currency climates, consumer demand, advertising, new goods and changing environmental conditions like the weather and seasons all affect procurement trends. Another challenge to procurement classifiers relates to the variability and wide scope of input data. Frequently goods are purchased from around the world to minimise costs leaving variably formatted records related to who filed the data worldwide, the first language of this person, local acronyms and textual shorthand. For example the following procurement items are typical examples; these can clearly challenge most classification technologies.

- 800 SLICED THK WHT BG -> flexible plastic packaging (for 800g loaf of bread)
- WHMEAL SS12 94 BAG -> cereal grain flour mix (bag of)

Due to the variable nature of spend data there is a very strong need to enable classification using fuzzy classification on the textual input. Similarity approaches normalise raw data to enable more efficient classification to numerous spend categories.

A solution based on similarity approaches ensures:

- Increased precision - inconsistency in the raw data can be solved automatically with high precision levels automatically coping with a large degree of data variance and providing the ability to process a wider range of previously unseen examples.
- Timeliness – the time needed to collect, normalise and categorise the data is very minimal.
- Flexibility - as to how source data can be collected (from low-tech manual to high-tech automated) and classified removes traditional barriers to adoption and dramatically reduces time to benefit.

The output of this project is a commercial solution (Spend360) specialising in Spend Analysis<sup>3,4</sup> that has so far been applied to spend analysis project in different sectors, such as retail, logistics, financial for customers such as KPMG, Deloitte, Schroeders Bank, Coles, Adelle, Associated British World Foods, Granger, Silver Spoon, Allied Mills, Ryvita Jordon, GAME and Allied Bakery.

## 2.5 CognIt

CognIT intends to apply X-Media results in their effort to pursue commercial and government markets related to strategic and operational intelligence. The software and concepts developed will be incorporated into the Navigator suited developed and marketed by CognIT. The Summarizer technology provided for X-Media, is widely deployed throughout our Navigator platform. The Summarizer kernel is especially vital for extracting metadata such as names, collocations and summarization extracts from documents for quicker and better search results. Navigator is a combined search and extraction based system that manages federated sources and various forms of advanced search-in-search. It also features an analyst workbench for visual analytics. It builds and extends ontologies together with the analyst by suggesting extensions through semantic analysis of text. Currently the system supports formal and semi-formal text sources. Some support is given to management of images. The ambition

---

<sup>3</sup> <http://spend360.ewebbed.com/>

<sup>4</sup> [http://www.k-now.co.uk/index.php/solutions/spend\\_analysis/](http://www.k-now.co.uk/index.php/solutions/spend_analysis/)

will be to extend Navigator analyses into other media such as video and sound using concepts developed in X-Media.

### 2.5.1 Target Market

CognIT's primary market is insurance (anti-fraud and money laundering) and government bodies such as police, national tax authorities and military. In addition CognIT is targeting oil (global oil exploration and gas market intelligence) as well as media (journalist support).

X-Media results combined with CognIT's existing Navigator platform has been presented before some of the more notable energy companies in the North Sea including ConocoPhillips, Total and Statoil. As part of the X-Media exploitation plan these companies have pledged to take part in a future demo where X-Media/Navigator elements will be slightly customized to demonstrate the potential of a full fledged application aimed towards oil exploration. A demo application directed towards geological intelligence and exploration is now being built based on this. An illustration of the anticipated demo is found below.



The demo and the underlying technology will be presented for the companies mentioned above in February 2010.

### 2.5.2 Currently Exploited Results

X-Media technologies have been deployed for solving cases of antifraud. CognIT has targeted the major insurance companies in Norway, and offered support and technology for their fraud investigation. We have been working closely with Storebrand among others, one of the largest insurance companies in Norway. When there are suspicions against one of the company's clients for breaking the rules for receiving insurance premium, an investigation will normally be launched. One of the tasks will be collecting and analyzing information that can highlight a possible fraud attempt. CognIT has as consultants been running Navigator to collect and analyze information from the Internet. We have with a minimum of efforts repeatedly been able to find crucial information that their own long lasting investigation has not been able to reveal. Based on this information the analyst at the insurance company can take further steps as stopping the benefit or go further with the investigation.

One of CognIT's primary focus for many years has been the government sector. CognIT has a well established portfolio of customers within this business area. CognIT already offers a knowledge management system in this domain called Best Practice™. More than 40 municipalities in Norway are licensed to use this software that was developed partly as a result of CognIT's effort in the OntoKnowledge

project, an EU-funded project which was part of the 4<sup>th</sup> Framework Programme. In addition to local governments, customers within this business area also include state agencies, the military, health care organisations as well as several universities and colleges. For X-Media we have solicited a small group of government agencies that can be considered technology pioneers. They have a profound need to systematically gather intelligence in ways similar to that described for FIAT. As state agencies of this type they are not so much interested in competitors – their attention is predominantly occupied with law enforcement. This implies that they have to monitor a large amount of data distributed across many types of media and sources both within and beyond organisation. Although focus will naturally vary their needs can be well aligned with those defined for FIAT. They need to fuse content from many different sources in order to establish trends and identify new practices and behaviours among individuals or larger groups. This group consists currently of the National Lottery Board, The Norwegian Tax Agency, The District Police of Hordaland and the Norwegian Intelligence Battalion in Afghanistan (Part of ISAF). All of them have been presented the ideas and possibilities promised by X-Media and all, but one has been exposed to early mock-ups and demos of more mature parts of the X-Media architecture and components.

The task of the National Lottery Board is to enforce national legislation related to lotteries and gambling. A major challenge is to monitor the on-line gambling market and assure proper conduct and compliance with rules and regulations. A major task is to identify and map sites and companies that offer gambling to Norwegian citizens. An inherent part of this is to gather and accumulate knowledge and information about the players in the market, new tendencies i.e. new games offered, stakes and prices and the type of players that are engaged. This should be stored in a central repository and be shared by different disciplines and for different purposes – in the most controversial cases for prosecution purposes. Another important task is to monitor advertisement practices, gather evidence related to law violations and provide a sound set of reports for higher government authorities as well as for strategy changes related to the task of the board itself. Part of the evidence basis includes collation of page snapshots, capture of banners, photos and other non-textual material. Another important aspect is the need to identify new trends in the market related to gambling and betting. The National Lottery Board was introduced to the status of X-Media in April 2008, and signalled their interest in trying out more of the tangibles stemming from the X-Media project. The Norwegian Lottery Board co-operates with other agencies in Northern Europe. CognIT has established an ongoing dialog with this agency.

The Norwegian Tax Agency is soliciting new technology for managing tasks similar to that of the National Lottery Board. Part of their aim is to systematically gather evidence from various sources on the Internet about improper business conduct. This includes monitoring activities on sites such as eBay and comparison with internal tax payer's records. It also includes collecting evidence about on-line businesses that might be hosted by a foreign server but systematically directed towards the national, domestic market. Evidence collection, trends, statistics, shift of behaviour and new business practices are important to spot and accumulate. Non-textual evidence gain increased importance. Image classification associated with text analysis will be very important to them. During our presentation and subsequent discussions with them it became evident that meta-data and ontological issues are very much related to that already addressed within the consortium. Norwegian Tax Agency is intrigued by the prospects of X-Media. They have been exposed to early demos during May 2008 and

are willing to try out the X-Media system and components once early tests have been documented. The FIAT (WP 13) case will be particularly important for them. As for FIAT there is a potential of saving time in browsing and searching information. The importance of collating and concentrating content that is semantically related in a shared repository is important for the tasks the agency sees as essential for itself in the future.

The work of the eBatallion in Afghanistan and the police has a restriction clause associated with it. However, it can be disclosed that a prime objective is to improve pro-activity in the way FIAT has described for their competitive intelligence. However, focus will be on anticipated events and incidents rather than new products and materials. Both have been exposed to X-Media through CognIT and have signalled strong interest in the deliveries that are yet to be made.

## 3 Exploitation by Industrial Users

### 3.1 Rolls-Royce

Between 2005 and 2009, Rolls-Royce was an industrial partner in a Department of Trade and Industry funded programme called Integrated Products and Services (IPAS). This addressed the issue of Knowledge Management for designers and service engineers.

By the time X-Media began in March 2006, Rolls-Royce were excited about semantic technologies emerging from IPAS. Furthermore, there was clear need emerging to deal more effectively with large, disparately located laboratory data, graphical plots, emails, reports, memos and component images.

Therefore, X-Media felt like a natural and essential research programme for which seven scenarios or use cases readily emerged. These included:

1. Experimental Vibration
2. Engine Health Monitoring
3. Nozzle Guide Vane Categorisation
4. Product Validation
5. Rationalised Designs for Manufacturing
6. Operations Room Support
7. Non Destructive Evaluation

Rolls-Royce held a series of discussions and workshops with the X-Media partners to review the viability of each. We wanted, of course, to choose use cases that could provide an excellent opportunity to develop, demonstrate and evaluate new semantic technologies. However, we also took considerable effort to evaluate whether the supporting resource was available (time from the use case owner and end users) and there weren't any major barriers, such as data security (material subject to government defence regulations or Intellectual Property legislation).

During phase 1, use cases 4 and 6 were combined to create use case A, and use cases 2, 3, 5 and 7 were dropped to leave those listed below. These are all well described in the work package 12 deliverables D12.1 to D12.7:

**A. Issue Resolution** – When an issue arises during new product development or whilst an engine is in service photographs are taken, many emails, memos and reports are written, laboratory experiments and data created etc. Therefore, this was clearly



large-scale (a lot of content in many corpora) and across media.

The new X-Media Issue Resolution system offered the additional benefit of guiding the engineers rigorously through the seven stage issue resolution process, from defining a issue through to identifying the root cause and, finally, implementing a fix. It could also enable the engineer to add any new learning into the system. Andy Harrison, an associate fellow from Service Engineering, was identified as a very enthusiastic and committed use case owner.

**B. Experimental Vibration** – Vibration analysis is used to calculate component life. This comes from theoretical models, laboratory experiments and whole engine testing. Huge volumes of vibration data are produced, images created, graphs plotted, emails, memos and reports written etc. Currently, a very experienced engineer analyses material from a few distinct corpora (Local Area Network and isolated databases) to determine how long a component will last.

The volume of this data certainly satisfied the need for a large-scale environment. Furthermore, this data could be linked to features in graphical plots (ZMods) and, hence, this use case clearly provided an opportunity for cross media searches.

The ZMods and images generated in the laboratory (ESPI) have consistent attributes (lighting, resolution etc) and, therefore, provided a challenging but realistic opportunity to develop image based similarity search tools.

Mark Walters, Chief of Capability Acquisition for Vibration Engineering was able to fulfil the role of use case owner. He was already focused on developing capability for his function and could be supported by a member of his team, Wye Houn Leong.

### 3.1.1 Meetings

There were internal Rolls-Royce meetings held every month to ensure each use case was progressing well technically and from a programme management point of view. During the entire X-Media programme, various partners visited the Rolls-Royce sites in Derby and Bristol to understand the use cases in more detail, develop the ontology and run the evaluations.

Rolls-Royce always had at least one representative at each plenary, usually the Principal Investigator. However, the use case owners have also attended a few plenaries. This was particularly useful in the early stages of the programme, when they were able to address the partners' technical questions.

A one day workshop was held for each use case at the end of phase one to determine the requirements for phase 2.

Throughout phase 2:

- Rolls-Royce visited the University of Sheffield every two weeks for face-to-face reviews of technical work.
- Rolls-Royce held telecons every 2 weeks to review technical development of all tools against the work package 12 plans on k@. All partners working on the Rolls-Royce use cases dialled in.

The Rolls-Royce X-Media team has also presented X-Media to many internal audiences between 2006 and 2010. These include Control Systems, Engineering Technology and the Global Knowledge Management Subject Matter Team (SMT), which includes Rolls-Royce Knowledge Management practitioners from the UK, Germany, USA and America.

The X-Media programme has delivered a set of tools that have demonstrated the

power of effective cross media searches to Rolls-Royce. Around thirty engineers have evaluated the tools and “see great potential.”

Rolls-Royce has already committed a significant budget to productionise some of the X-Media technologies within a new programme called Samulet. Specifically, we will focus on the technologies our engineers believe will deliver the greatest business benefit, namely terminology recognition, information extraction, semantic search, information visualisation (geoplot, timeline and engine plot), as well as semantic forms. The programme is funded by Rolls-Royce and the Technology Strategy Board (TSB) and will be a collaboration between K-Now, Sheffield University, Cambridge University, Loughborough University, OSYS and Tacit Connections.

In addition, the X-Media technologies will become part of another Knowledge Management programme called Siloet. This launched in February 2010 and is also be funded by Rolls-Royce and the TSB (the X-Media-related part is 100% funded by Rolls-Royce). This programme will further develop some X-Media technologies, which are include terminology recognition and search applied to meeting and discussion tracking. The current investment by Rolls-Royce on X-Media tops £600k. Many other X-Media tools have really impressed Rolls-Royce engineers, such as the ESPI similarity matcher. Currently, we have assigned these a lower priority and they will not be productionised in the near future (not before 2011).

### 3.2 Centro Ricerche Fiat

During the project, two Fiat working applications have been built (see D13.1)

- Competitor scenario forecast (CSF)
- Noise curve analysis and evaluation (Noise)

**X-Media tools** elaborate collected data, extracted annotation and store them in a Semantic Repository.

**The Fiat Portal** is the graphical user interface that queries the Repository and presents the information to the user in the most suitable way for the task s/he has to carry out

#### **Competitor Scenario Forecast**

During the four years of the project, the CSF test bed has proven to have the potential to be a powerful means to support the activities of Marketing, Competitor Analysis and Engineering & Design aimed to continuously monitor the automotive market.

During the evaluation stage end users appreciated especially:

- “Automatic classification of non structured document allows to reduce considerably time daily employed in web search”
- “Possibility to access and manage separately Images and news text coming from the same document, enhances the effective reuse of knowledge”
- “Automation of time consuming activities like segmentation and launch calendar building increase considerably the productivity.”

During a workshop organized in CRF, other companies that showed concrete interest were CNH and IVECO.

Before the end of X-MEDIA project we plan to recruit end users from other department (i.e Aftersales) and/or other companies to participate in final evaluation to demonstrate the testbed potentialities.

After the end of the project CRF plans to install and test the as-is released prototype in Marketing to evaluate the CSF functionalities to exploit at an industrial level.

As result of this activity, a detailed plan for exploitation in a short time frame of



around one to two years of ad hoc implementations of a marketable version of the whole testbed specific project outcome will be produced.

Moreover CRF is using the current prototype as vehicle for demonstration to other companies as IVECO and CNH.

On the basis of the showed interest the possible installation and testing of the as-is released prototype and the development of feasibility study for an ad hoc implementation will be evaluated.

### **Noise curve analysis and reduction**

During the project lifetime, the Noise testbeds results has proven to have the potential of a useful mean to support the activities of OEMs functions like advanced design of vehicle performances and components.

During the evaluation the involved end users belonging to Wind Tunnel and Acoustic department appreciated:

- “The large amount of data that the system could automatically extract from Reports is really impressive”.
- “SAE prediction, Descriptor prediction and Spectra Comparison tools are very interesting and useful to improve the problem solving during a feasibility study”
- “The Noise application is very intriguing because allows to reuse all past test results in an efficient and effective way allowing a real cost and time reduction”

Before the end of the project we plan to involve people from other departments and eventually other FIAT group companies in the final evaluation of the Noise prototype.

After the end of X-Media CRF plan to

- Install and test the as-is released prototype in wind tunnel & acoustic department:
- Evaluate the Noise functionalities to exploit

A Detailed plan for exploitation in a short time frame of around one to two years of ad hoc implementations of a marketable version of the whole testbed specific project outcome will be produced.

Moreover CRF will use the current noise prototype as demonstrator to show in other FIAT Auto departments with the aim to evaluate the possible adaptation of the application to other performances (i.e. safety, etc)

## **4 Exploitation by Academic Partners**

### **4.1 University of Sheffield**

Results are being exploited both in direct participation in industrial projects and via their spin-out company K-Now (see above).

Further exploitation of the X-Media technologies is in the following Projects:

- [SmartProducts](#), an FP7 European Integrated Project on embedding proactive knowledge into Smart Products to help their design, manufacturing, servicing and use. Coordinator is SAP, users are Philips, Fiat and Eads. Two other X-Media partners are involved: Centro Ricerche Fiat and Open University (2009-2012); in this project we exploit and continue development of the knowledge acquisition and capture technologies;
- [WeKnowIt](#): an FP7 European Integrated Project on Collective Intelligence. Work carried out by our group concerns Personal Intelligence (the user-centric sharing

and retrieval of information) and Organizational Intelligence (reuse information by organizations and communities). User partners are Yahoo!, Telefonica and Sheffield City Council Emergency Services; two other X-Media Partners are involved: CERTH (coordinator) and Koblenz (2008-2011). In this project we further exploit the knowledge management and user centred technologies.

- [SEALS](#): Semantic Evaluation At Large Scale: an FP7 European Project on Evaluation of Semantic Web Technologies (2009-2012); here we exploit the search technologies.
- Samulet (mentioned above); here we further develop and exploit the search and capture technologies developed in X-Media in order to take them to Technology Readiness Level 7, the necessary certification for use in an aerospace company;
- Siloet (mentioned above), we further develop the acquisition and search technologies towards social use.
- eArchaeology: a project funded by AHRC + EPSRC, approved for the University of Sheffield. In collaboration with the University of York, Department of Archaeology document analysis technologies developed in X-Media were used for analysing 1 Million multi-media archaeology documents.

We are also currently awaiting answers from funding bodies and companies for other two industrial projects.

## 4.2 University of Koblenz

Results are being exploited in direct participation in the following projects:

- WeKnowIt: an FP7 European Integrated Project on Collective Intelligence. Work carried out by University of Koblenz concerns Mass Intelligence (i.e. learning of global behavioural patterns, recognitions of trends and events in large amounts of media, tracking provenance and estimating certainty of knowledge in large scale social networks). User partners include Yahoo!, Telefonica and Sheffield City Council. In this project we aim to exploit our X-Media results on authority ranking by tensor factorization, uncertainty management, and process tracking (WP1-WP2).
- WeGov: an FP7 STREP on analysis of processes and trends related to eGovernment in online communities, i.e. governance in the eSociety. Work carried out by University of Koblenz concerns analysis of dynamic trends and opinions in social media, as well as information exchange in online communities. We aim to adopt X-Media technologies on authority ranking by tensor factorization (TripleRank), gathering process knowledge from decentralized communication networks (process knowledge models, capturing & modelling processes from semantic email communication). User partners include Leibniz Institute for Social Sciences (GESIS) Germany, eDemocracy Programme of the Hansard Society in UK, and the non-profit Government To You initiative (Greece).
- CollabCloud: the nationally funded project (German Ministry of Education and Research) on scalable management of distributed heterogeneous data, knowledge discovery, and data analysis. Our goal is to develop a semantic information mediator which allows for easy analysis of data slices and sharing of intermediate and final results among the users. Additionally, the history of all analysis steps will be stored for later result assessment and reuse. The project will exploit X-Media achievements in the fields of meta knowledge management, uncertain knowledge, and knowledge fusion (WP1-3). User

partners include a number of German companies, such as Osthus GmbH and fluid Operations GmbH.

University of Koblenz participates in Hewlett Packard Labs Innovation Research Programme with an industrial project on advanced search and automatic generation of synthetic answer documents in large-scale corporate knowledge bases. The project will employ tensor factorization methods developed in WP4 as well as provenance tracking and meta knowledge management introduced in WP1.

We are also awaiting answers from companies for other two further industrial projects in the field of knowledge management and ontology learning (in cooperation with Siemens and Hewlett Packard departments).

### **4.3 Open University**

Further exploitation of X-Media technologies is primarily envisaged in the context of other research projects. In particular, the exploitation and development of knowledge fusion and semantic search technologies is planned within the SmartProducts project (see above).

### **4.4 University of Hildesheim**

Results are being exploited in direct participation in industrial projects :

- MyMedia is a FP7 STREP project, it aims at developing personalized recommender systems which are based on multiple multimedia sources..
- Lefos is funded by the European regional development fund EFRE. It aims at integrating recommender system components like the ones known from Amazon in line of business E-commerce shops.

### **4.5 Fondazione Bruno Kessler**

Further exploitation of X-Media technologies is primarily envisaged in the context of other research projects, e.g. the LiveMemories project (<http://livememories.fbk.eu/>) funded by the Autonomous Province of Trento.

### **4.6 KIT**

X-Media results are exploited in several projects with AIFB participation:

- Active (<http://www.active-project.eu/>), an FP7 European Integrated Project on Enabling the Knowledge-Powered Enterprise, coordinated by British Telecom In this project we deploy Semantic Search technologies developed in X-Media.
- NanOn, a project funded by the Leibnitz Gemeinschaft, Germany concerned with the semi-automatic generation of ontologies in nanotechnology. Therein we use X-Media technology for text analysis and information extraction.
- iGreen (<http://www.igreen-projekt.de/>), a project funded by the German Ministry of Education and Research. The aim of this project is to design and implement a location-based service and knowledge network for linking distributed, heterogeneous, public, as well as private sources. In this project, X-Media solutions for metaknowledge management and semantic search will be used.

#### 4.7 University of Ljubljana

We plan to use the techniques developed within the X-Media project in the following industrial and research projects:

- An industrial project with one of the largest Slovenian companies (Mercator), where the task is to develop a tool that would suggest new places (nation-wide and internationally) for their expansion.
- A research project on intelligent tutoring systems, where the main focus will be on automatic conceptualisation of domains (development of cognitive expert models), which will require analysis of cross-media data (text and raw data).

Moreover, most of the developed tools are also included in the Orange, a data-mining suite used by many researchers around the world.

#### 4.8 CERTH

The scientific results of the activities carried out by CERTH within the X-Media project are being exploited in several on-going projects with both industrial and research objectives:

- [WeKnowIt](#): Develops novel techniques for exploiting multiple layers of intelligence from user-contributed content, which together constitute Collective Intelligence, a form of intelligence that emerges from the collaboration and competition among many individuals, and that seemingly has a mind of its own. To this end, the scientific results of X-Media in cross media analysis drive the technologies investigated in WeKnowIt for analyzing and combining inputs from various sources.
- [JUMAS](#): Addresses the need to build an infrastructure able to optimise the information workflow in order to facilitate later analysis. New models and techniques for representing and automatically extracting the embedded semantics derived from multiple data sources are being developed. The technology developed in X-Media for semantically driven image analysis has been considered in this project.
- MEDUSA: Combines different types of sensors supporting operations within urban environments. The essential improvement that the MEDUSA project implements is the provision of a means to address major existing capability shortfalls: (i) in the fusion of data from multiple diverse types of sensors and (ii) in the data representation, by means of a consolidated and integrated view, including overlaying across 3D GIS (Geographical Information System) models to facilitate decision-making by commanders of control and support operations. The expertise gained from studying technologies for cross media analysis has been exploited for deciding the optimal multi-sensor fusion techniques.
- Recommender System Development via Automatic Semantic Induction: a project funded by Motorola with the scope to develop a distributed semantic framework for personalised recommendation using advanced textual analysis and user profiling techniques, knowledge-based reasoning and collaborative filtering to enable targeted recommendation of textual content, for any domain for which there is textual data even when little or no expert knowledge is available.
- PESCaDO: Develops strategies for user-centred environmental service orchestration and multilingual information delivery in the context of active decision support for environment conditioned tasks.
- [VERGE](#): a hybrid interactive video retrieval system, which is capable of searching into video content by integrating different search modules that employ visual and

textual-based techniques

#### 4.9 Labri

Scientific results of LaBRI obtained within the X-Media project will be used in two projects.

- POLINUM(2010-2012): this project aims to develop a platform devoted to the scanning of old documents. The ambition of Polinum is to provide a global processing chains from the digitization of a document to the visualisation on different kind of devices. The image classification and image retrieval techniques will be used in this projects. Height partners (six industrials and two academics) are involved in this project for a global budget of 4 millions of euros.
- DIGIBOOK(Submitted to the french ANR CONTINT) : The project DIGIBOOK is define and implement a new format called “DIGIBOOK” for the representation of the quality of document scanning. This format requires the extraction of local features of a decomposition of image in regions. The segmentation techniques based on knowledge can provide more robust segmentations than usual techniques. Knowing, the kind of document we plan to segment it using the segmentation methods developed in the X-MEDIA project. DIGIBOOK involved 4 French laboratory and the BNF (French National Library).
- IMMED : Multimedia Indexing for Monitoring of Dementia Diseases" - 2009 - 2012 supported by French Nationala Research Agency. The use of know - how in descriptors and training methods.

## 5 Conclusions

In this deliverable we have discussed the dissemination and exploitation routes taken as part of the X-Media project over the four years, as well as the future exploitation of the project results. We believe that X-Media has been highly successful in engaging the community, in obtaining relevant results and in exploiting them. The project has been able to attract several thousands of Euros in further investments both from research councils and industrial partners. Patents have also been applied for and spin-out companies have been created to further exploit some of the technologies.

## 6 Appendix A: Details of the exploitation over 4 years

### 6.1.1 Invited Talks

- Fabio Ciravegna was Invited keynote speaker at the 3rd STI Roadmapping Workshop "Charting the next generation of semantic technology", June 2009 co-located with ESWC 2009: title of talk: "Six Impossible Things".
- Fabio Ciravegna was Invited keynote speaker at the 3rd Asian Semantic Web Conference, Bangkok, Thailand, 2-5 February 2009.
- Fabio Ciravegna was Invited speaker at roundtable discussion: "How can the UK incubate the next generation of innovators?" held in London on the 6th of November 2008, in occasion of the 25th anniversary of the Toshiba Fellowship Program.
- Fabio Ciravegna was Invited keynote speaker at the 11th International Conference on Business Information Systems, BIS 2008, Innsbruck, Austria 5-7 May 2008
- Fabio Ciravegna was Invited Speaker at the Norwegian Semantic Days 2008, Stavanger, Norway, April 2008.
- Ivan Bratko was Invited speaker at the main slovenian conference on computer science and informatics (Dnevi slovenske informatike), title: "Current directions in artificial intelligence", April 2009. (In this talk Bratko presented also the X-media project).
- Lars Schmidt-Thieme was Invited Speaker at the eEvolutions-Workshop, 2010, Hildesheim, Germany.
- Lars Schmidt-Thieme was invited speaker at the German-Polish Symposium on Data Analysis and its Applications (GPSDAA), Aachen, Germany, October 8 to 10, 2009.
- Lars Schmidt-Thieme was invited speaker at the IFCS 2009 (International Federation of Classification Societies), Dresden, Germany
- Lars Schmidt-Thieme was Invited Speaker at the eEvolutions-Workshop, 2009, Hildesheim, Germany.
- Lars Schmidt-Thieme was Invited Speaker at NEC Research Labs, 2008, Heidelberg, Germany.
- Lars Schmidt-Thieme was Invited Speaker at the Workshop on Data Mining for Enterprises, 2008, Hildesheim, Germany.
- Lars Schmidt-Thieme was Invited Speaker at the eEvolutions-Workshop, February 29, 2008, Hildesheim, Germany.
- Enrico Motta was invited speaker at the 6th European Summer School on Ontological Engineering and the Semantic Web (SSSW-2008), Cercedilla, Spain, July 2008
- Sebastian Blohm was Invited Speaker at IBM Almaden Research Center by back in August 2008. In course of a 3 month collaboration visit on X-Media related topics.
- Yiannis Kompatsiaris from CERTH gave a talk on "Extracting Collective Intelligence from Social Content", Workshop on Cross-media information analysis, extraction and management at SAMT 2008, Koblenz, Germany, 3-5 Dec. 2008.
- Yiannis Kompatsiaris from CERTH gave a talk on "Combining visual and knowledge processing for semantics extraction", Workshop on Semantic User



Descriptions and their influence on 3D graphics and VR, Lausanne – Switzerland, November 13, 2008

- Fabio Ciravegna was Invited Speaker at the open day 'Call 3 in Motion', organized by INFOSO.E2 Content and Knowledge, European Commission, Luxembourg, December 2007.
- Steffen Staab gave the keynote 'Ontology Learning' at the inaugural workshop of the Language, Interaction and Computation Lab at the Centre for Mind / Brain Sciences (CiMEC) at the University of Trento, Rovereto, May 29, 2007
- Steffen Staab gave an invited talk 'Querying OWL Ontologies', Learning Lab Lower Saxony, University of Hannover, May 25, 2007
- Steffen Staab gave an invited talk 'Das Semantic Web: Wissensstrukturierung von allen für alle' Studium Generale, University of Marburg, May 23, 2007
- Fabio Ciravegna was invited speaker at NWO-CATCH Meeting: 'Television archives and the Semantic Web', Nederlands Instituut voor Beeld en Geluid, February 9 2007, Hilversum, The Netherlands. Talk about X-Media with title: Automatic and Semi-Automatic Document Annotation for Knowledge Acquisition.
- Philipp Cimiano gave an invited talk on 'Ontologies and text mining for biomedicine' at an internal workshop at Böhringer Ingelheim, May 2007.
- Fabio Ciravegna Invited speaker at KAMC 07, workshop on Knowledge Acquisition from Multimedia Content, organised within the 2nd International Conference on Semantics And Digital Media Technologies (SAMT), Genoa, Italy, 5-7 December 2007.
- Fabio Ciravegna Invited Speaker at the Doctoral Symposium of the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Busan, Korea, 11-15 November 2007.
- Fabio Ciravegna was Invited Speaker at the Seminar on Information Logistics (SOIL), University of Jönköping, Jönköping, Sweden, October 2007.
- Yiannis Kompatsiaris from CERTH gave a talk on 'Convergence of multimedia and knowledge technologies in the framework of the EC funded IPs and STREPs: AceMedia, Aim@Shape, BOEMIE, MESH, X- Media, K-Space, Vitalas and VICTORY' at CIVR 2007 Practitioner Day ACM International Conference on Image and Video Retrieval, July 9-11 2007 University of Amsterdam, Amsterdam, The Netherlands
- Philipp Cimiano gave an invited talk at the International Conference on Terminology and Artificial Intelligence (TIA).
- Lars Schmidt-Thieme was Invited Speaker at Statistics under one Umbrella (DagStat 2007), Bielefeld, Germany.
- Lars Schmidt-Thieme had an invited talk at the 14th meeting of the special interest group on information technology (2007), Hildesheim, Germany.
- Yiannis Kompatsiaris from CERTH gave a talk on "Ontologies and Machine Learning for Semantic Multimedia Analysis", Workshop on Knowledge Management and Semantic Web for Engineering Design, 28 October 2007 in Whistler, British Columbia, Canada, in conjunction with K-CAP 2007
- Lars Schmidt-Thieme was Invited Speaker at the International Conference on Computational Statistics (CompStat 2006), Roma, Italy.
- Lars Schmidt-Thieme was Invited Speaker at ISM Symposium on Large-scale Data Linkage, Data Mining and Statistical Methods, Tokyo, Japan, October 3, 2006

- Lars Schmidt-Thieme had an invited talk at the Knowledge Media Institute, The Open University, Milton Keynes, UK, 18.5.2006.
- S. Rudolph, Description Logic Rules: Virtual and Real Gain in Expressivity at No (Additional) Cost. Invited talk at the International Center for Computational Logic (ICCL), TU Dresden, Germany, July 2008.
- S. Rudolph, Why Infinity is Your Friend. Invited talk at the Knowledge Representation and Reasoning Group, Oxford University Computing Laboratory, March 2009.
- S. Rudolph, OWL Reasoning. Invited lecture at the advanced course on Semantic Web of the Netherlands research school for Information and Knowledge Systems, Utrecht, Netherlands, September 2009.
- S. Rudolph, OWL 2 Rules. Invited lecture at the advanced course on Semantic Web of the Netherlands research school for Information and Knowledge Systems, Utrecht, Netherlands, September 2009.
- S. Rudolph, Conjunctive Query Entailment in Description Logics With Inverses, Counting, and Nominals. Invited talk at the Knowledge Based Systems Group, Vienna University of Technology, November 2009.
- S. Rudolph, Rules for OWL 2. Invited talk at the Knowledge Based Systems Group, Vienna University of Technology, November 2009

### 6.1.2 Tutorials

- Fabio Ciravegna gave an Invited Tutorial at the 7th European Summer School on Ontological Engineering and the Semantic Web (SSSW-2009), Cercedilla, Spain, July 2009.
- Fabio Ciravegna gave an Invited Tutor at the Summer School on Multimedia Semantics '09, Koblenz, August 2009
- Yiannis Kompatsiaris gave an invited tutorial on "Detecting, Understanding and Exploiting Web communities" 18th Intl. World Wide Web Conference, WWW 2009.
- Yiannis Kompatsiaris gave an invited tutorial on "Managing and Modeling of Multimedia and User Generated Content in Web 3.0", 4th Summer School for Multimedia Semantics (SSMS), 2009.
- Fabio Ciravegna gave an Invited Tutorial on Introduction to the Semantic Web at the 7th International Semantic Web Conference, Karlsruhe, Germany. Dates: October 26-30, 2008. Other tutors are Jim Hendler, Asun Gomez Perez, Aldo Gangemi, Natasha Noy, Jerome Euzenat, John Domingue, and Enrico Motta.
- Fabio Ciravegna was invited tutor at the 6th European Summer School on Ontological Engineering and the Semantic Web (SSSW-2008), Cercedilla, Spain, July 2008
- Fabio Ciravegna gave Invited Tutorial on Semantic Web Technologies in Large Distributed Enterprises at the Norwegian Semantic Days 2008, Stavanger, Norway, April 2008.
- Fabio Ciravegna gave a two hour tutorial on exploitation of X-Media results at Altran Italia, a major European software house, in Milano Italy.
- Fabio Ciravegna gave an Invited Tutorial on Introduction to the Semantic Web at the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Busan, Korea, 11-15 November 2007. Other tutors are Jim Hendler, Asuncion Gomez Perez, Aldo Gangemi, Natasha Noy, Jerome Euzenat, John Domingue, and Enrico Motta. Fabio is giving the lecture on Automatic



Techniques for Extracting Semantic Data.

- Fabio Ciravegna gave a Tutorial on Semantic Web Technologies for Knowledge Management in Large Distributed Organisations at the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference, Busan, Korea, 11-15 November 2007.
- Fabio Ciravegna gave a Tutorial at the 4th European Semantic Web Conference on Semantic Web Technologies for Knowledge Management in Large Distributed Organisation, Innsbruck, June 2007
- Fabio Ciravegna was Invited tutor at the 5th European Summer School on Ontological Engineering and the Semantic Web (SSSW-2007), Cercedilla, Spain, July 2007. Steffen Staab gave an invited tutorial 'Ontologies', at Summer School on Multimedia Semantics. Analysis, Annotation, Retrieval and Applications. Glasgow, UK, July 15-21, 2007
- Philipp Cimiano gave a lecture on Ontologies and Natural Language Processing at ESSLLI 2007 in Dublin
- P. Hitzler, M. Krötzsch, and S. Rudolph, Semantic Web Ontology Languages - Tutorial at Informatik 2009, Lübeck, Germany, September 2009.
- P. Hitzler, M. Krötzsch, and S. Rudolph, Knowledge Representation and Reasoning for the Semantic Web - OWL 2 and Rules - Tutorial at KI2009, Oldenburg, Germany, September 2009
- M. Krötzsch and S. Rudolph, Ontology Modelling Languages - Course at the 21st European Summer School in Logic, Language and Information (ESSLLI), Bordeaux, France, July 2009.
- P. Hitzler, M. Krötzsch, and S. Rudolph, Semantic Web Modelling Languages - Tutorial at the 21st International Joint Conference on Artificial Intelligence, Pasadena, CA, USA, July 2009.
- P. Hitzler, M. Krötzsch, and S. Rudolph, OWL 2 Rules - Tutorial at the 6th Annual European Semantic Web Conference (ESWC2009), May 2009

### 6.1.3 Conference and Workshop organisation

- Fabio Ciravegna was General Chair of 6th European Semantic Web Conference, Heraklion, Greece, 2009
- Fabio Ciravegna was Sponsorship chair for the 8th International Semantic Web Conference, Washington DC, 25-28 October 2009.
- Yiannis Kompatsiaris was Program Chair of the 4th International Conference on Semantic and Digital Media Technologies (SAMT 2009), Graz, Austria, Dec 2-4, 2009.
- Yiannis Kompatsiaris was Co-Chair of the CIVR2009 ACM International Conference on Image and Video Retrieval, July 8-10, 2009, Santorini, Greece.
- The Workshop on High-Level Information Extraction at the ECML PKDD 2008 in Antwerp, Belgium was co-organized by Sebastian Blohm (UKARL).
- The Workshop on Cross-Media Information Analysis, Extraction and Management at SAMT 08 in Koblenz, Germany, was co-organized by Spiros Nikolopoulos (CERTH) and Joao Magalhaes (USFD).
- X-Media partners also co-organized the KI 2008 Workshop on Ontology-Based Information Extraction Systems, September 2008.
- Steffen Staab was track chair for Wirtschaftsinformatik 2007 (German Conf. for

Information Systems) – Track chair for 'Knowledge Management', 28.02. - 02.03.2007

- Steffen Staab was workshop chair for SOA-Tag 2007, 'Sicherheit und Beherrschbarkeit von SOA in virtuellen Organisationen', Koblenz 28.9.2007
- CERTH is the organiser of the 8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007). WIAMIS is one of the main international fora for the presentation and discussion of the latest technological advances in interactive multimedia services. (<http://mkg.iti.gr/wiamis2007/index.html>)
- Yiannis Kompatsiaris was Co-Chair of the First international conference on Semantics And digital Media Technology (SAMT 2006), December 6-8, 2006, Athens, Greece

#### 6.1.4 Book

- P. Hitzler, M. Krötzsch and S. Rudolph, Foundations of Semantic Web Technologies. CRC Press/Prentice Hall, 2009

#### 6.1.5 Papers and Posters

- Dadzie, A.-S. & Ciravegna, F. (2009). Flexible Knowledge Generation and Re-Use in Aerospace Engineering using the XMediaBox. At the European Semantic Technology Conference (ESTC 2009) - Session: Semantics in the Enterprise.
- Petrelli, D., Mazumdar, S., Dadzie, A.-S., and Ciravegna, F. (2009). Multi visualization and dynamic query for effective exploration of semantic data. In Bernstein, A., Karger, D. R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., and Thirunarayan, K., editors, The Semantic Web - ISWC 2009, Proceedings, 8th International Semantic Web Conference, volume 5823 of LNCS, pages 505–520.
- Dadzie, A.-S., Lanfranchi, V., and Petrelli, D. (2008). Towards AI usability: Problems, strategies and practicals. In CHI 2008 Workshops and Courses: Usable Artificial Intelligence Session.
- Dadzie, A.-S. and Petrelli, D. (2009). Visual knowledge exploration and discovery from different points of view. In Proc., IEEE Symposium on Visual Analytics Science and Technology, VAST 2009, pages 227–228.  
Philipp Cimiano, Sebastian Rudolph, Helena Hartfiel Computing Intensional Answers to Questions - An Inductive Logic Programming Approach Data & Knowledge Engineering, 69, (3), pages 261-278, March, 2010
- Dadzie, A.-S. and Petrelli, D. (2009). Visual knowledge exploration and discovery from different points of view. In Proc., IEEE Symposium on Visual Analytics Science and Technology, VAST 2009, pages 227–228.
- Rowe, M & Ciravegna, F. Disambiguating Identity Web References using Web 2.0 Data and Semantics. To Appear in the Journal of Web Semantics Special Issue on "Bridging the Gap" Data Mining and Social Network Analysis for integrating Semantic Web and Web 2.0, 2010
- J. Iria, S. Nikolopoulos, M. Mozina, A. Lavelli, C. Giuliano, L. Romano, D. Kuznar and I. Kompatsiaris. Cross-Media Knowledge Extraction in the Car Manufacturing Industry. In Proceedings of the 21st IEEE International Conference on Tools with Artificial Intelligence (ICTAI'09), New Jersey, USA, November 2009.

- J. Iria. Automating Knowledge Capture in the Aerospace Domain. In Proceedings of the 5th ACM International Conference on Knowledge Capture (K-CAP'09), Redondo Beach, California, September 2009.
- J. Iria and J. Magalhaes. Exploiting Cross-Media Correlations in the Categorization of Multimedia Web Documents. In Proceedings of the IJCAI'09 Workshop on Cross-Media Information Access and Mining, Pasadena, USA, July 2009.
- Dadzie, A.-S., Lanfranchi, V., Petrelli, D. (2009). Seeing is Believing: Linking Data With Knowledge. In the Special issue of the Information Visualization Journal on Human-Centered Information Visualization, **8:3**, pp.197-211
- A. Dadzie, J. Iria, D. Petrelli and L. Xia. The XMediaBox: Sensemaking through the Use of Knowledge Lenses. In Proceedings of the 6th European Semantic Web Conference (ESWC), 2009.
- J. Iria and J. Magalhaes. Web News Categorization using a Cross-Media Document Graph. In Proceedings of the ACM International Conference on Image and Video Retrieval (CIVR), 2009.
- J. Iria. A Core Ontology of Knowledge Acquisition. In Proceedings of the 6th European Semantic Web Conference (ESWC), 2009.
- Chapman, S. K-Now: Enabling Your Knowledge, Now. Knowledge Management Market Place 2009, Bath.
- Thanh Tran, Sebastian Rudolph; Philipp Cimiano; Haofen Wang: Efficient Computation of Formal Queries from Keywords on Graph-Structured RDF Data. 25th International Conference on Data Engineering (ICDE), Shanghai, China, 2009.
- Christoph Ringelstein, Steffen Staab, "DIALOG: Distributed Auditing Logs," Web Services, IEEE International Conference on, pp. 429-436, 2009 IEEE International Conference on Web Services, 2009
- Ullmann, T., Uren, V., and Nikolov, A. The SemSearchXplorer - Exploring Semantic Search Results with Semantic Visualizations, Workshop: AST 2009 - Applications of Semantic Technologies at the Informatik 2009, Luebeck, Germany., Informatik 2009 - Im Focus das Leben. GI Edition Lecture Notes in Informatics, eds. Gesellschaft fur Informatik e.V.; Fischer, Stefan; Maehle, Erik; Reischuk, Rudiger, Kollen Verlag.
- Fawad Nazir, Victoria Uren, Andriy Nikolov: Algorithms for Generating Ontology Based Visualization from Semantic Search Results, To be presented at 1st International Workshop on Modelling and Visualization of XML and Semantic Web Data (MoViX'09) August 31, 2009
- M. Fernandez, V.Lopez, M.Sabou, V.Uren, D.Vallet, E.Motta, P.Castells. Using TREC for cross-comparison between classic IR and ontology-based search models at a Web scale . Workshop on Semantic Search at the 18th International World Wide Web Conference (WWW 2009). Madrid, Spain, April 2009.
- Ansgar Scherp, Thomas Franz, Carsten Saathoff and Steffen Staab: F---A Model of Events based on the Foundational Ontology DOLCE+ Ultra Light. International Conference on Knowledge Capture, K-CAP, 2009.06.24
- Thomas Franz and Ansgar Scherp and Steffen Staab: Are Semantic Desktops Better? International Conference on Knowledge Capture, K-CAP, 2009.06.24
- Thomas Franz and Antje Schultz and Sergej Sizov and Steffen Staab: TripleRank: Ranking Semantic Web Data By Tensor Decomposition. International Semantic Web Conference (ISWC), 2009.08.14

- Fernando Silva Parreiras and Carsten Saathoff and Tobias Walter and Thomas Franz and Steffen Staab: APIs a gogo: Automatic Generation of Ontology APIs'. International Conference on Semantic Computing (ICSC), 2009.06.24
- Ansgar Scherp and Thomas Franz and Carsten Saathoff and Steffen Staab: A Common Model of Events and Objects based on DOLCE+ Ultra Light. ESWC Poster Session, 2009.06.29
- Ansgar Scherp and Thomas Franz and Carsten Saathoff and Steffen Staab: A Model of Events based on a Foundational Ontology, Tech Report Universität Koblenz-Landau, 2009.07.10
- Giuliano, Claudio, Alfio Massimiliano Gliozzo and Carlo Strapparava: Kernel Methods for Minimally Supervised WSD, Computational Linguistics, vol. 35, no. 4, 2009
- Hendricks, Iris; Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid O' Seaghdha, Sebastian Pado', Marco Pennacchiotti, Lorenza Romano and Stan Szpakowicz: SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals. In Proceedings of the NAACL HLT 2009 Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-2009), Boulder, Colorado, 2009
- Claudio Giuliano: Fine-Grained Classification of Named Entities Exploiting Latent Semantic Kernels. In Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL-2009) Boulder, Colorado, 2009
- Tonelli, Sara, Claudio Giuliano: Wikipedia as Frame Information Repository. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP 2009), Singapore, Malaysia, 2009
- Busche, Andre, Schmidt-Thieme, Lars, Ferraro, Massimo, Biasuzzi, Christian (2009): Action-Based Linking of Distributed Heterogeneous Semantic Web Tools. Submitted to ESCW2010
- Christine Preisach, Lars Schmidt-Thieme (2009): /Relational Time Series Classification/. Submitted to ADAC (Advances in Data Analysis and Classification) Journal
- Nikolopoulos, S., Papadopoulos, G. T., Kompatsiaris, I., and Patras, I. 2009. An Evidence-Driven Probabilistic Inference Framework for Semantic Image Understanding. In Proceedings of the 6th international Conference on Machine Learning and Data Mining in Pattern Recognition (Leipzig, Germany, July 23 - 25, 2009). P. Perner, Ed. Lecture Notes In Artificial Intelligence, vol. 5632. Springer-Verlag, Berlin, Heidelberg, 525-539.
- Elisavet Chatzilari, Spiros Nikolopoulos, Eirini Giannakidou and Ioannis Kompatsiaris, "Leveraging Social Media For Training Object Detectors", 16th International Conference on Digital Signal Processing (DSP'09), Special Session on Social Media, 5-7 July 2009, Santorini, Greece.
- Spiros Nikolopoulos, Christina Lakka, Ioannis Kompatsiaris, Christos Varytimidis, Konstantinos Rapantzikos, Yannis Avrithis, "Compound Document Analysis by Fusing Evidence Across Media," Content-Based Multimedia Indexing, International Workshop on, pp. 175-180, 2009 Seventh International Workshop on Content-Based Multimedia Indexing, 2009
- Spiros Nikolopoulos, Elisavet Chatzilari, Eirini Giannakidou, Ioannis Kompatsiaris, "Towards fully un-supervised methods for generating object detection classifiers using social data," Image Analysis for Multimedia Interactive

Services, International Workshop on, pp. 230-233, 2009 10th Workshop on Image Analysis for Multimedia Interactive Services, 2009.

- Nadejda Nikitina, Sebastian Rudolph, Sebastian Blohm Refining Ontologies by Pattern-Based Completion In Eva Blomqvist, Kurt Sandkuhl, Francois Scharffe, Vojtech Svatek. Proceedings of the Workshop on Ontology Patterns (WOP 2009), CEUR, 516, Oktober, 2009
- ESPI Image Indexing and Similarity Search in Radon Transform Domain, Remi Vieux, Jenny Benois-Pineau, Jean-Philippe Domenger, Achille Braquelaire. 2009 Seventh International Workshop on Content-Based Multimedia Indexing (accepted CBMI'09)
- Multiclass Object Detection And Categorization using SVM Region Remi Vieux, Jenny Benois-Pineau, Jean-Philippe Domenger, Achille Braquelaire Classification and Relaxation Labelling, submitted to IEEE Conference on Computer Vision and Pattern Recognition, 2010. Denny Vrandecic, Frank Dengler, Sebastian Rudolph, Michael Erdmann RDF syntax normalization using XML validation In Lalana Kagal, Ora Lassila, Tim Finin, Proceedings of the Workshop Semantics for the Rest of Us at ISWC2009, CEUR-WS, Washington, DC, Oktober, 2009
- Uta Lösch, Sebastian Rudolph, Denny Vrandecic, Rudi Studer Tempus Fugit - Towards an Ontology Update Language In Lora Aroyo et al., 6th European Semantic Web Conference (ESWC 09), pages: 278--292, Springer-Verlag, Lecture Notes on Computer Science, 5554, Juni, 2009
- Duc Thanh Tran, Haofen Wang, Peter Haase Hermes: Data Web search on a pay-as-you-go integration infrastructure Journal of Web Semantics, 7(3), 2009
- Haofen Wang, Thomas Penin, Kaifeng Xu, Junquan Chen, Xinruo Sun, Linyun Fu, Qiaoling Liu, Yong Yu, Duc Thanh Tran, Peter Haase, Rudi Studer Hermes: a travel through semantics on the data web SIGMOD, 2009
- Haofen Wang, Qiaoling Liu, Thomas Penin, Linyun Fu, Lei Zhang, Duc Thanh Tran, Yong Yu, Yue Pan Semplore: A Scalable IR Approach to Search the Web of Data Journal of Web Semantics, 7(3), 2009
- Duc Thanh Tran, Peter Haase, Rudi Studer Semantic Search - Using Graph-Structured Semantic Models for Supporting the Search Process 17th International Conference on Conceptual Structures (ICCS'09), 2009
- Krisztian Buza, Lars Schmidt-Thieme (2009), Motif-based Classification of Time Series with Bayesian Networks and SVMs, in Proceedings of 32nd Annual Conference of the Gesellschaft für Klassifikation (GfKI 2008), Hamburg, Germany.
- Steffen Rendle, Christine Preisach, Lars Schmidt-Thieme (2009), Learning to Extract Relations for Relational Classification, in Proceedings of the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009), Bangkok, Springer.
- Lopez, V., Nikolov, A., Fernandez, M., Sabou, M., Uren, V., and Motta, E. Merging and Ranking Answers in the Semantic Web: The Wisdom of Crowds, The 4th Asian Semantic Web Conference, Shanghai, China, December 2009 (best paper award)
- Lopez, V., Sabou, M., Uren, V., and Motta, E. Cross-Ontology Question Answering on the Semantic Web –an initial evaluation, Knowledge Capture Conference, 2009, California, September 2009
- Fernandez, M., Overbeeke, C., Sabou, M., and Motta, E. What makes a good ontology? A case-study in fine-grained knowledge reuse, The 4th Asian Semantic

Web Conference, Shanghai, China, December 2009

- Nikolov, A., Uren, V., Motta, E., and de Roeck, A. Overcoming Schema Heterogeneity between Linked Semantic Repositories to Improve Coreference Resolution, 4th Asian Semantic Web Conference (ASWC 2009), Shanghai, China, December 2009
- Nikolov, A., Uren, V., Motta, E., and de Roeck, A. Towards data fusion in a multi-ontology environment, Workshop: Linked Data on the Web at 18th International World Wide Web Conference (WWW 2009), Madrid, Spain, April 2009
- Nikolov, A., Uren, V., Motta, E., and de Roeck, A. Refining instance coreferencing results using belief propagation, 3rd Asian Semantic Web Conference (ASWC 2008), Bangkok, Thailand, February 2009
- B. Strle, M. Možina, I. Bratko. Qualitative approximation to Dynamic Time Warping similarity between time series data, 23rd Annual Workshop on Qualitative Reasoning, June 2009
- D. Kužnar, M. Možina, I. Bratko. Curve prediction with Kernel Regression, Workshop on Learning from Multi-Label Data at ECMLPKDD 2009.
- J. Žabkar, M. Možina, I. Bratko, Demšar. Discovering Monotone Relations with Pade, Learning Monotone Models from Data at ECMLPKDD 2009.
- M. Možina, M. Guid, J. Krivec, A. Sadikov, I. Bratko. Fighting knowledge acquisition bottleneck with argument based machine learning, ECAI 2008C. Giuliano, M. Možina, I. Bratko. Argument Based Machine Learning from Examples and Text. In Proceedings of the 1st Asian Conference on Intelligent Information and Database Systems, Dong Hoi City, Vietnam, 2009.
- D. Tomas, C. Giuliano. A semi-supervised approach to question classification. In Proceedings of the 17th European Symposium on Artificial Neural Networks: Advances in Computational Intelligence and Learning, Bruges, Belgium, 2009.
- L. Romano, K. Buza, C. Giuliano, L. Schmidt-Thieme. XMedia: Web People Search by Clustering with Machine Learned Similarity Measures. In Proceedings of the 2nd Web People Search Evaluation Workshop (WePS 2009), 2009. Krisztian Buza, Christine Preisach, Andre Busche, Lars Schmidt-Thieme, Wye Houn Leong, Mark Walters (2009), Eigenmode Identification in Campbell Diagrams, in International Workshop on Machine Learning for Aerospace.
- L. Xia, J. Iria, F. Ciravegna. An Approach to Modeling Heterogeneous Resources for Information Extraction. In Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), Marrakesh, Morocco, 26-30 May, 2008.
- M. A. Greenwood, J. Iria, F. Ciravegna. Saxon: An Extensible Multimedia Annotator. In Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), Marrakesh, Morocco, 26-30 May, 2008.
- Z. Zhang, J. Iria, C. Brewster, F. Ciravegna. A Comparative Evaluation of Term Recognition Algorithms. In Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), Marrakesh, Morocco, 26-30 May, 2008.
- Iria, J. Relation Extraction for Mining the Semantic Web. In "Semantic Web - Concepts And Applications", Ravi Kumar Jain B (editor), Icfai University Press.
- A. Dadzie, R. Bhagdev, A. Chakravarthy, S. Chapman, J. Iria, V. Lanfranchi, J. Magalhaes, D. Petrelli, F. Ciravegna. Applying Semantic Web Technologies to Knowledge Sharing in Aerospace Engineering. To appear in Special issue of the

Journal of Intelligent Manufacturing on Knowledge Discovery and Management in Engineering Design and Manufacturing, 2008.

- Ravish Baghdev, Sam Chapman, Fabio Ciravegna, Vita Lanfranchi and Daniela Petrelli: Hybrid Search: Effectively Combining Keywords and Semantic Searches in Proceedings of the 5th European Semantic Web Conference, ESWC 08, Tenerife, June 2008
- A. Gliozzo, C. Giuliano. Instance-based ontology population exploiting named-entity substitution. In Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008), Manchester, UK, 2008.
- Bernhard Schueler, Sergej Sizov, Steffen Staab, Duc Thanh Tran: Querying for Meta Knowledge. 17th International Conference on World Wide Web (WWW), Beijing, China, 2008: pp. 625-634.
- Christine Preisach and Lars Schmidt-Thieme, Ensembles of Relational Classifiers, Knowledge and Information Systems Journal 2008
- Steffen Rendle, Lars Schmidt-Thieme (2008), Active Learning of Equivalence Relations by Minimizing the Expected Loss Using Constraint Inference, in Proceedings of 8th IEEE International Conference on Data Mining (ICDM 2008), Pisa.
- Yan Liang, Haofen Wang, Qiaoling Liu, Thanh Tran, Thomas Penin, Yong Yu: Efficient Index Maintenance for Frequently Updated Semantic Data. 3rd Asian Semantic Web Conference (ASWC), Bangkok, Thailand, 2008: pp. 182-196
- Thomas Penin, Haofen Wang, Thanh Tran, Yong Yu: Snippet Generation for Semantic Web Search Engines. 3rd Asian Semantic Web Conference (ASWC), Bangkok, Thailand, 2008: pp. 493-507
- Haofen Wang, Kang Zhang, Qiaoling Liu, Thanh Tran, Yong Yu: [Q2Semantic] A Lightweight Keyword Interface to Semantic Search. 5th European Semantic Web Conference (ESWC), Tenerife, Spain, 2008: pp. 584-598
- Stephan Bloehdorn, Marko Grobelnik, Peter Mika, Duc Thanh Tran: Proceedings of the Workshop on Semantic Search (SemSearch) at the 5th European Semantic Web Conference (ESWC), Tenerife, Spain, 2008. CEUR Workshop Proceedings 334, CEUR-WS.org
- Sebastian Blohm, Philipp Cimiano. Scaling up Pattern Induction for Web Relation Extraction through Frequent Itemset Mining. In Benjamin Adrian, Günter Neumann, Alexander Trousov, Borislav Popov, Proceedings of the KI 2008 Workshop on Ontology-Based Information Extraction Systems. September 2008.
- Jürgen Umbrich, Sebastian Blohm. Exploring the knowledge in Semi Structured Data Sets with Rich Queries. Tenerife, Spain, June 2008.
- Sebastian Blohm, Markus Krötzsch, Philipp Cimiano. The Fast and the Numerous – Combining Machine and Community Intelligence, 2008.
- M. Poesio, E. Barbu, C. Giuliano, L. Romano, Supervised relation extraction for ontology learning from text based on a cognitively plausible model of relations., In Proceedings of the ECAI 2008 3rd Workshop on Ontology Learning and Population, Patras, Greece, 2008.
- C. Giuliano, M. Mozina, I. Bratko. Arguments Extracted from Text in Argument Based Machine Learning: A Case Study. In Proceedings of the SAMT 2008 workshop on Cross-media information analysis, extraction and management, Koblenz, Germany, 2008.
- M. Giordanino, C. Giuliano, D. Kuznar, A. Lavelli, M. Mozina, L. Romano. Cross-Media Knowledge Acquisition: A Case Study. In Proceedings of the SAMT

2008 workshop on Cross-media information analysis, extraction and management, Koblenz, Germany, 2008.

- A. Gliozzo, C. Giuliano. Kernel Methods for Minimally Supervised WSD. *Computational Linguistics* 35(4).
- Steffen Rendle, Lars Schmidt-Thieme: Scaling Record Linkage to Non-Uniform Distributed Class Sizes, in *Proceedings of the 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2008)*.
- M. Cammarata, 'Knowledge Management System per il Web semantico: studio di un' interfaccia per il progetto X-Media in FIAT Auto', degree thesis discussed at Università degli studi di Torino, February 2008
- Steffen Rendle, Lars Schmidt-Thieme (2008): Scaling Record Linkage to Non-Uniform Distributed Class Sizes, in *Proceedings of the 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2008)*, Osaka, Springer.
- Christine Preisach, Steffen Rendle, Lars Schmidt-Thieme (2008), Relational Classification Using Automatically Extracted Relations by Record Linkage, in *Proceedings of the High Level Information Extraction Workshop at the European Conference on Machine Learning (ECML-WS 2008)*.
- Vieux R., Benois-Pineau J., Domenger J.P., Braquelaire. "Image classification with user defined ontology" *EUSIPCO 2007*, page(s) 723-727 – 2007
- Vieux R., Benois-Pineau J., Domenger J.P., Braquelaire "A Multidimensional Clustering Index for Image Retrieval". *The 5th IET Visual Information Engineering 2008 Conference (VIE'08)* page(s): 152-157 Chine (2008)
- Jose Iria, Lei Xia and Ziqi Zhang WIT: Web People Search Disambiguation using Random Walks. In *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*
- J. Iria, V. Uren, A. Lavelli, S. Blohm, A. Dadzie, T. Franz, J. Magalhaes, S. Nikolopoulos, I. Kompatsiaris, C. Preisach, P. Slavazza. Enhancing Enterprise Knowledge Processes via Cross-Media Extraction. In *Proceedings of the Fourth International Conference on Knowledge Capture (KCAP 2007)*.
- Vitaveska Lanfranchi, Ravish Bhagdev, Sam Chapman, Fabio Ciravegna, Daniela Petrelli. Extracting and Searching Knowledge for the Aerospace Industry. In *Proceedings of 1st European Semantic Technology Conference, Vienna 31 May - 1 June 2007*.
- Lei, Y., Uren, V., and Motta, E. A Ranking-Driven Approach to Semantic Search. Poster at 3rd Asian Semantic Web Conference (ASWC 2008), Bangkok, Thailand, February 2009 (best poster award)
- T. Franz, S. Staab, R. Arndt. The X-Cosim Integration Framework for a Seamless Semantic Desktop. In: *K-CAP 2007 – Proceedings of the Fourth International ACM Conference on Knowledge Capture*. 28-31 October 2007, Whistler, BC.
- B. Schüler, S. Sizov, S. Staab. Management of Meta Knowledge for RDF Repositories. In: *IEEE ICSC-2007 – Proc. of 1st IEEE International Conference on Semantic Computing*. September 17-19, 2007. Irvine, California.
- Sebastian Blohm, Jürgen Umbrich, Philipp Cimiano, York Sure: Iterative Learning of Relation Patterns for Market Analysis with UIMA. *UIMA Workshop at the GLDV Frühjahrstagung, Tübingen, April 2007*.
- Sebastian Blohm: Using the Web to Reduce Data Sparseness in Pattern-based Information Extraction. In *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)*. Springer ,



Warsaw, Poland, September 2007.

- Thanh Tran, Holger Lewen, Peter Haase: On the Role and Application of Ontologies in Information Systems In Proceedings of 5th IEEE International Conference on Research, Innovation & Vision for the Future (RIVF'07), Hanoi, Vietnam. March 2007.
- Sebastian Blohm, Philipp Cimiano, Egon Stemle: Harvesting Relations from the Web - Quantifying the Impact of Filtering Functions In Proceedings of the 22nd Conference on Artificial Intelligence (AAAI-07). Association for the Advancement of Artificial Intelligence (AAAI), July 2007.
- Anupriya Ankolekar, Denny Vrandečić, Markus Krötzsch, Thanh Tran: The Two Cultures: Mashing up Web 2.0 and the Semantic Web In Proceedings of the 16th International Conference on the World Wide Web (WWW), Banff, Canada. May 2007.
- Claudio Giuliano, Alberto Lavelli, Daniele Pighin, Lorenza Romano FBK-IRST: Kernel Methods for Semantic Relation Extraction. In Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)
- Claudio Giuliano, Alfio Gliozzo and Carlo Strapparava FBK-irst: Lexical Substitution Task Exploiting Domain and Syntagmatic Coherence In Proceedings of the 4<sup>th</sup> International Workshop on Semantic Evaluations (SemEval-2007)
- Claudio Giuliano, Alfio Gliozzo Instance Based Lexical Entailment for Ontology Population In Proceedings of the 2007 Joint Meeting of the Conference on Empirical Methods on Natural Language Processing (EMNLP) and the Conference on Natural Language Learning (CoNLL)
- Claudio Giuliano, Alberto Lavelli, Lorenza Romano. Relation Extraction and the Influence of Automatic Named Entity Recognition ACM Transactions on Speech and Language Processing, 5(1), 2007.
- Anna Corazza, Alberto Lavelli, Giorgio Satta. Phrase-Based Statistical Parsing. In Proceedings of the EVALITA Workshop, Rome, Italy, 10 September 2007.
- Magalhães, J. and S. Rüger (2007). Information-theoretic semantic multimedia indexing. ACM Conference on Image and Video Retrieval. Amsterdam, The Netherlands.
- Yuanguai Lei, Andriy Nikolov, Victoria Uren, and Enrico Motta, Detecting Quality Problems in Semantic Metadata without the Presence of a Gold Standard, presented at 5th International EON Workshop on Evaluation of Ontologies and Ontology-based tools, (EON2007) held at the 6th International Semantic Web Conference (ISWC), 12 Nov.2007. Busan, Korea.
- Lei, Y., Uren, V. and Motta, E. (2007) SemEval: A Framework for Evaluating Semantic Metadata, KCAP 2007, The Fourth International Conference on Knowledge Capture Whistler, Canada, 28-31 Oct. 2007.
- Lei, Y., Uren, V. and Motta, E. (2007) SemRef: An Evaluation Reference Model for Semantic Metadata, Poster at 4th European Semantic Web Conference
- Andriy Nikolov, Victoria Uren, Enrico Motta and Anne de Roeck, Using the Dempster-Shafer theory of evidence to resolve ABox inconsistencies, presented at the Uncertainty Reasoning for the Semantic Web Workshop (URSW), held at the 6th International Semantic Web Conference (ISWC), 12 Nov.2007. Busan, Korea.
- Andriy Nikolov, Victoria Uren and Enrico Motta, KnoFuss: A Comprehensive Architecture for Knowledge Fusion, Poster at KCAP 2007, The Fourth International Conference on Knowledge Capture Whistler, Canada, 28-31 Oct. 2007.

- Lei, Y., Lopez, V., Motta, E., Uren, V. An infrastructure for building semantic web portals. *Journal of Web Engineering*, vol.6 (4), 283-308, (2007).
- Uren, V., Lei, Y., Lopez, V., Liu, H., Motta, E. and Giordanino, M. The usability of semantic search tools: a review, *Knowledge Engineering Review*, (2007), 22:361-377
- Steffen Rendle and Lars Schmidt-Thieme, Information Integration of Partially Labeled Data, to appear in Proceedings of 31th Annual Conference of the Gesellschaft fuer Klassifikation (GfKI) 2007, Freiburg.
- X-Media poster at the workshop 'Knowledge and Experience Management', Halle, Germany, September 2007.
- X-Media poster at the workshop 'Security and Governance in Virtual Organizations', Koblenz, Germany, September 2007.
- Christine Preisach, Lars Schmidt-Thieme (2006), *Relational Ensemble Classification*, in Proceedings of 6th IEEE International Conference on Data Mining (ICDM) 2006, Hong Kong.
- Gilardoni, L., Biasuzzi, C., Ferraro, M., Fonti, R., Slavazza, P. LKMS/Mnemosyne: Semantic Web technologies and Markup techniques for a Legal Knowledge Management System, V Legislative XML Workshop 14-16 June 2006 – Fiesole.
- Slavazza, P., Fonti, R., Ferraro, M., Biasuzzi, C., Gilardoni, L. Towards a Knowledge Ecosystem, In Steffen Staab, Vojtech Svátek (Eds.) *Managing Knowledge in a World of Networks*, 15th International Conference, EKAW 2006, Pödebrady, Czech Republic, October 2-6, 2006, Proceedings. Lecture Notes in Computer Science 4248 Springer 2006, ISBN 3-540-46363-1.
- Pallen, I. and Fullerton, S. Bridging the Traditional-Semantic Divide: Extending Enterprise Search using Semantic Technology. In Proceedings of the 6th European Semantic Web Conference (ESWC), 2009.
- Fullerton, S. and Pallen, I. Extending Traditional Enterprise Search using Semantic Techniques. In Proceedings of the 3rd Annual European Semantic Technology Conference (ESTC), 2009.

### 6.1.6 Flyer

A flyer has been created that describes the project goals. The flyer has been distributed at conferences and events, as well as directly to potential interested parties (e.g. potential customers or users). It is available on the project web site.

### 6.1.7 Sponsorships

To create awareness of X-Media in the relevant communities from, X-Media sponsored the following events in the area of semantic Web (we list the conferences sponsored in 2008-2009 only):

- 6<sup>th</sup> European Summer School on Ontological Engineering and the Semantic Web (SSSW-2008), Cercedilla, Spain, July 2008;
- 7<sup>th</sup> European Summer School on Ontological Engineering and the Semantic Web (SSSW-2009), Cercedilla, Spain, July 2009;
- 7<sup>th</sup> International Semantic Web Conference (ISWC), Karlsruhe, Germany, October 26-30, 2008.

- 6<sup>th</sup> European Semantic Web Conference (ESWC), Heraklion, Greece, 2009
- 8<sup>th</sup> International Semantic Web Conference, Chantilly, Virginia, October 2009
- Semantic Search 2009 Workshop Located at the 18th Int. World Wide Web Conference (WWW2009 April 21st, 2009 (Workshop day), Madrid, Spain
- 16<sup>th</sup> International Conference on Knowledge Engineering and Knowledge Management Knowledge Patterns (EKAW) 2008, Acitrezza, Italy. September 29<sup>th</sup>-3<sup>rd</sup> October 2008
- Summer School on Multimedia Semantics 2009 (SSMS), Koblenz, Germany, 23-28<sup>th</sup> August 2009
- 7<sup>th</sup> International Workshop on Content-Based Multimedia Indexing (CBMI), Chania, Crete, 3<sup>rd</sup>-5<sup>th</sup> June 2009
- European Semantic Technology Conference 2009 (ESTC), Vienna, Austria, 2<sup>nd</sup>-3<sup>rd</sup> December 2009

### 6.1.8 Community Building and Clustering

X-Media director Fabio Ciravegna has become service manager for road-mapping for STI International (<http://www.sti2.org/>), a major think tank on semantic technologies. X-Media established a connection with CHORUS project, a Coordination Action funded recently (FP6 call 4.1) in Audiovisual Search Engines strategic objective. Its main target is to coordinate activities of other projects in the same call. More specifically, X-Media participation is targeted to the following action lines:

- WG1: Audio-visual content indexing and retrieval technologies, since the cross-media analysis techniques developed within the project could be useful for indexing and retrieval applications
- WG6: Use-Cases and New services. X-Media in collaboration with the FIAT and Rolls-Royce industrial partners has well defined use-cases for knowledge extraction and management. Moreover as a dissemination activity, a poster describing X-Media was presented at the CHORUS workshop on “use cases and new services for multimedia content search” that was organized at the Inria Rocquencourt premises (<http://www.ist-chorus.org/rocquencourt--mar-13--14-07.php>).

X-Media is now also a member of SmaRT (Semantic Multimedia Research and Technology) networking cluster which aim is to bring together the leading researchers from the field of Semantic Web, Multimedia and Signal Analysis. The goal of the cluster is to identify and address emerging research challenges in Semantic Multimedia (<http://www.smart-network.eu/>).