Semantic innovation management across the extended enterprise

Ke Ning, David O'Sullivan*, Qinxiang Zhu and Stefan Decker

Digital Enterprise Research Institute,
National University of Ireland, Galway, Ireland
E-mail: ke.ning@deri.org E-mail: dos@nuigalway.ie
E-mail: qinxiang.zhu@deri.org
E-mail: stefan.decker@deri.org
*Corresponding author

Abstract: Innovation provides the means for enterprises to survive and grow. Due to the distributed and heterogeneous characteristics, it's difficult to manage innovation across the extended enterprise. A novel vision for Semantic Innovation Management (SIM) is introduced. The system architecture combines ontology, inference and mediation technologies used to create a semantic web of innovation knowledge. A functional framework of the SIM System (SIMS) based on metadata harvesting and RDF access technologies is presented. An applied case study is explained in detail. This, in turn, facilitates all participating organisations and users in forming closer alliances based on the sharing of innovation information.

Keywords: innovation management; extended enterprise; semantic web; RDF; information integration.

Reference to this paper should be made as follows: Ning, K., O'Sullivan, D., Zhu, Q. and Decker, S. (2006) 'Semantic innovation management across the extended enterprise', *Int. J. Industrial and Systems Engineering*, Vol. 1, Nos. 1/2, pp.109–128.

Biographical notes: Ke Ning is an IT specialist at IBM in Beijing, China. His principle role is in the support of middleware. Prior to joining IBM he was a Senior Researcher at the Digital Enterprise Research Institute in Galway, Ireland. His research interests included semantic web technology applications. In 2004 he received his PhD at Tsinghua University, Beijing, China for his work in the area of Enterprise Information Systems and Business Process Modelling.

David O'Sullivan is a Research Leader at the Digital Enterprise Research Institute at the National University of Ireland in Galway. His research interest is principally in the area of semantic innovation management across the extended enterprise. He has over 80 publications to his credit, including a book entitled Manufacturing Systems Redesign. He is an Associate Editor for a number of international journals.

Qinxiang Zhu is a Researcher at the Digital Enterprise Research Institute at the National University of Ireland in Galway. His research interests are in the area of semantic innovation management across the extended enterprise. He is also interested in semantic web search, information integration and dashboard technologies.

Stefan Decker received his PhD from the University of Karlsruhe and spent several years at Stanford University and Information Sciences Institute/University of Southern California as Senior Researcher and Assistant Research Professor. He is currently Executive Director at the Digital Enterprise Research Institute, National University of Ireland, Galway. He is one of the most widely cited semantic web scientists and has served on the committees of a number of international conferences.

1 Introduction

Innovation is the process of making changes in something established by introducing something new. Within a typical organisation, there are various degrees of innovation from suggesting ideas or managing single projects to managing an entire programme of changes. The competitiveness and sustainability of an enterprise depends on the effective management of innovation (Zairi, 2004). Innovation leads to improved products and services that delight customers, more efficient processes that keep costs competitive and more agile organisations capable of meeting unforeseen challenges. The nature of innovation is changing. In the past it focused on capturing value adding ideas and making them add value for the customer. Innovation is now entering a new phase where the cycle from invention to realisation will be shorter. Innovation must now happen more quickly and more continuously (Brown and Eisenhardt, 1997). A second change in the nature of innovation is that it now needs to happen across internal functions within organisations and between organisations in an extended enterprise. A culture of collaboration and sharing of ideas is emerging (Huxham, 1996). A third change happening in the way innovation is currently being managed looks at the notion of intellectual property. Organisations that value their people and their role in the innovation process will gain higher returns and succeed in making the innovation process a powerful weapon for organisational growth (Grossman and Helpman, 1991). A recent industrial survey revealed that 90% of enterprises think innovation is a key competitive advantage (Stamm, 2004). However, organisations manage innovation poorly. Surveys indicate that up to 80% of development programmes among Fortune 500 companies fail to meet their objectives. The reasons cited are familiar: Poor Goal definition; Poor alignment of Actions to Goals; poor reporting of Results; Poor Action management; Low participation among employees in Teams (Dooley and O'Sullivan, 2000).

To support the lifecycle of innovation and improve innovation management, different kinds of innovation management systems have been developed, including systems for individual innovation e.g. (MindManager, 2005; IdeaFisher, 2005), project innovation e.g. (Teamstorm, 2005; GroupSystems, 2005), collaborative innovation e.g. (Askme, 2005; NextNet, 2005), and distributed innovation e.g. (IdeaChain, 2005; GoldfireInnovator, 2005). Previous research by some of the authors resulted in the development of a system called iTeams (Dooley and O'Sullivan, 2000), which focuses on providing a knowledge management platform to develop, communicate and control innovation in a any team. iTeams provided a structure for managing innovation in any organisation and across organisations but like all other systems reviewed, required each organisation to adopt the schema or ontology inherent in the system. This creates limitations since the nature of knowledge management systems allows each system to grow organically to reflect the expressions and terms used within the specific team. The term 'indicator', for example, in one team may evolve into the term 'measure' in another.

To date, there are still few efforts successful in helping to improve innovation management across the extended enterprise and the sharing of innovation information. An extended enterprise may be regarded as a collection of independent, heterogeneous companies working closely together in order to produce an integrated product or service, in whose commercial success they all have a vested interest (Browne, 1998). This arrangement, whereby independent legal entities have to work in increasingly coordinated relationships, is becoming a necessity in order to innovate complex systems in less time and at a lower cost. Clearly, it is more difficult to manage the innovation process across the extended enterprise than in a single organisation. In order to make different suborganisations work as a holistic system, more efforts are needed to realise the sharing of information, matching of goals, incorporation of actions, and so on. Meanwhile, from the technical viewpoint, there will be many different innovation management systems in the extended enterprise. Due to the distributed, heterogeneous and autonomic characteristics, it is not easy to integrate them to provide a seamless flow of innovation knowledge.

With the emerging and rapid development of the semantic web, it is possible to adopt novel semantic web technologies to improve innovation management across the extended enterprise. The semantic web can be envisioned as an extension of the current web, which aims to make the web more understandable to computer programs, and allows data to be shared and reused across applications, enterprise, and community boundaries, easily. There are two backbone technologies for the semantic web: RDF and OWL (Miller, 2004). They, as web standards, provide a framework for asset management, enterprise integration, and sharing and reusing data on the web. These standard formats for data sharing span applications, enterprise, and community boundaries. All users – both human and machine – can share and understand the information available on the semantic web. The foundation of RDF (2005) is built on a very simple model, but the basic logic can support largescale information management and processing in a variety of different contexts. The assertions in different RDF files can be combined, providing far more information together than they contain separately. RDF supports flexible and powerful query structures, and developers have created a wide variety of tools for working with RDF. OWL (2005) provides a language for defining structured, web based ontologies. This delivers richer data integration and interoperability among descriptive communities. Many semantic web based information systems have been created (Semantic_Web_Challenge, 2005), and have been successfully used in some industrial applications.

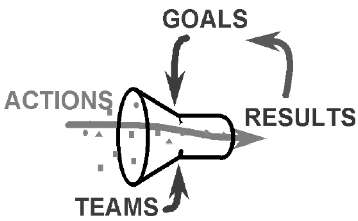
Based on our practice in extended enterprise innovation management, as well as our knowledge in semantic web technologies, we believe that using semantic web technologies will produce great benefits for innovation management across extended enterprises. Thus, the vision of Semantic Innovation Management (SIM) emerges. Inspired by this vision, we have designed the system architecture and a functional framework for semantic innovation management, and developed an applied SIMS based on commercially available semantic software products. Though it is still in its infancy, SIMS has shown good potential.

This paper is organised as follows. Firstly, a novel vision for Semantic Innovation Management is introduced. Secondly, the system architecture for SIM, which combines ontology, inference and mediation technologies, is discussed. The semantic features of SIM are also analysed. Thirdly, a functional framework of the SIM System based on metadata harvesting and RDF data access technologies is presented. We define five modules for SIMS: Metadata Provider, Metadata Harvester, Central Repository, Service Provider, and Semantic Innovation Dashboard. Fourthly, an applied SIMS case study is presented, which uses available semantic web technologies, including RDF Gateway (Intellidimention, 2005), Kapow HTML wrapper (2005), Microsoft Dot Net Framework, and so on. We explain this technology and its integration. Finally, conclusions are drawn and statements are made on future research and development.

2 Vision of SIM

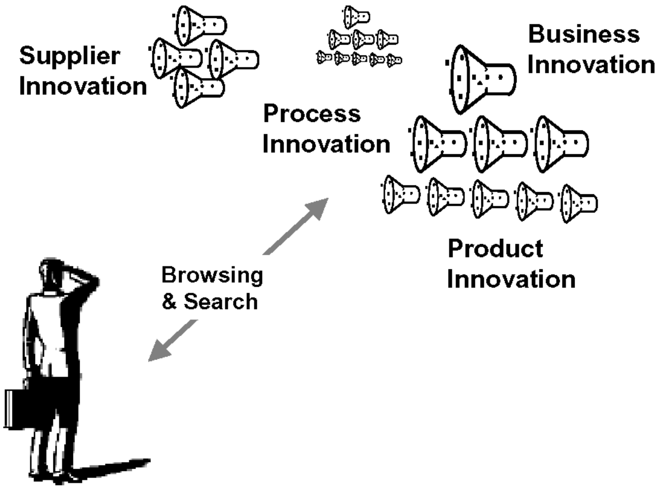
The vision of SIM is to manage innovation information and to seamlessly operate innovation processes across the extended enterprise. In our former work, we have developed a concept model for innovation management in an organisation – the innovation funnel (see Figure 1). Innovation in any organisation or any business unit within an organisation can be visualised as a funnel. The funnel identifies Goals, Actions, Teams and Results as the key elements of innovation, which when integrated, provide the essential framework for promoting sustainable development in organisations. Goals and Teams constrain the neck of the funnel. Actions represent ideas and projects that enter the mouth of the funnel. Those actions that map well with goals and the availability of teams flow into the funnel to become goal centred projects. Goals and actions require periodic review to continuously manage goal alignment and to stimulate and regulate new ideas and projects entering the mouth of the funnel. Such a funnel can help us to structure the innovation process, and we have created functional software to support this model called iTeams. This system has now been successfully used in over 40 types of organisations for product, process and service innovation management.

Figure 1 iTeams innovation funnel



Based on this innovation funnel model, let us imagine the big picture of SIMS. From the user's point of view, the big picture of semantic innovation management across the extended enterprise will look like Figure 2. There will be many innovation funnels within a single organisation and many more across the extended enterprise. Users will be able to access information from any innovation funnel regardless of its location, information type or semantics. Users can pose all kinds of questions to SIMS about innovation processes, such as: What projects use welding robots? Which performance indicators are most common in Manufacturing? Which Warranties are most common? What projects do we need to pay more attention to? And so on. The questions asked may concern innovation systems in any part of the extended enterprise.

Figure 2 Innovation funnels across an extended enterprise

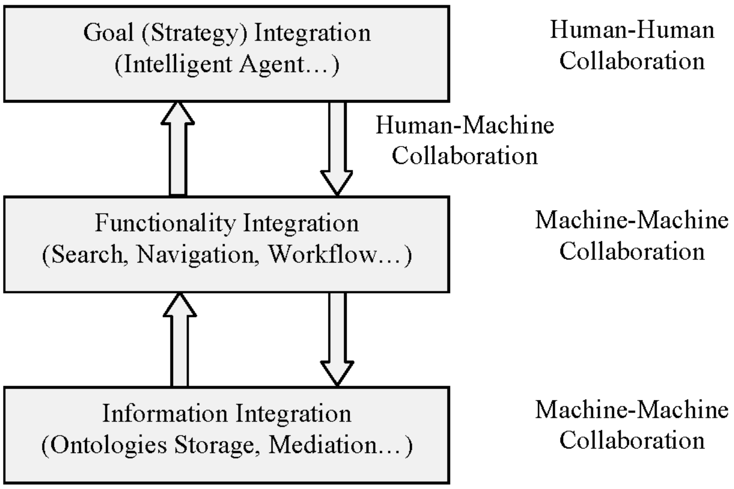


Our objective is to implement such a SIM system, which will: (1) improve link capability between separate data sources, providing an easy way to navigate all the innovation systems of the extended enterprise, (2) support extensibility to easily add new innovation data sources in different contexts, (3) provide the ability to integrate different innovation database resources with structured data, semistructured data and unstructured data.

3 System architecture for SIM

To bring the vision of SIM into reality, we have designed the system architecture based on semantic web technologies, as shown in Figure 3.

Figure 3 System architecture for semantic innovation management



The system architecture has three levels of integration – information, functionality and goal. In the Information Integration level, relevant information about innovation is first 'marked up' using the innovation ontology. Semantic web applications can better understand the semantics, and therefore more intelligently locate and integrate data for a wide variety of organisations. Storage, mediation, and evolution management allows the ontology to facilitate the smooth flow of relevant information among different parts of the innovation process. The Functionality Integration level is used to realise semantic innovation management functional requirements (e.g., semantic inference, context aware search and navigation, workflow management, and so on) and allow 'machine-machine' interrogation and interpretation of innovation related data. To encourage the different levels of users to participate in the interactive innovation process, the prerequisite is that there are shared goals or interdependent goals among these users. The Goal Integration level is the guarantee for the realisation of low level information and functionality integration. The goals alignment in distributed innovation is conducted by the human, intuitively and manually. However, in the semantic innovation process, intelligent agents specified in terms of the respective goals and obligations are employed to cooperate with one another to realise goal directed collaboration. Goal integration corresponds to human-human collaboration, functionality and information integration corresponds to machine-machine collaboration. The interaction between goal and functionality corresponds to human-machine collaboration. The ultimate goal is to improve goal integration by effective and efficient machine-machine collaboration.

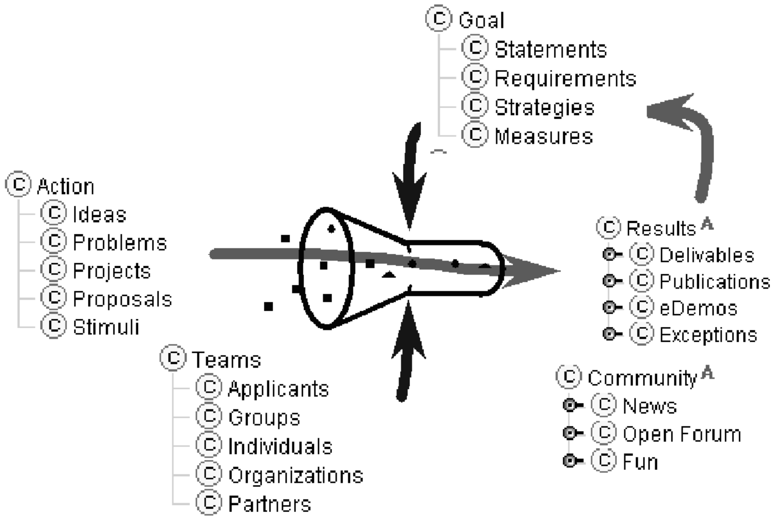
In this architecture, Shared Ontology, Semantic Inference, Semantic Mediation, and Semantic Interface are key technologies, and each is now discussed in more detail.

3.1 Shared ontology

Ontology is a general conceptualisation of a specific domain in both human and machine readable format. In general, it consists of classes, properties, relationships and axioms. Ontology forms the backbone technology for the semantic web. Ontology promises a shared understanding of a domain that can be communicated between people and software applications (Staab and Studer, 2004). To realise semantic innovation, it is necessary to build a formally and explicitly expressed ontology about innovation related information. Such ontology needs to include all the necessary information for innovation management. It should also provide the conceptual framework for making innovation data machine readable. Different organisations have different concepts for their work activities and also have a different set of innovation related information for populating their internal and shared innovation ontology. The semantics of an organisation's specific working context is captured by its local or private ontology which serves the purposes of the particular organisation. Public or shared ontologies are used to articulate the common or generic semantics about a domain, which need to be shared and communicated across multiple local contexts. The main role of shared ontology is to support the common subject related interoperation across the extended enterprise. Semantic innovation is concerned with shared ontology and the common subject refers to innovation management.

Based on the innovation funnel, we have created an innovation ontology for the extended enterprise, which is called 'Iteams Ontology', as shown in Figure 4.

Figure 4 Iteams innovation ontology



Among this ontology, all the information about innovation management is classified as five categories and specified as five abstract classes, i.e., Goals, Actions, Teams, Results and Community. Goals are defined as the objectives of an organisation's effort. It can be embodied as Requirements, Strategies, Measures, and so on. Actions are defined as expenditure of the effort to achieve the goals, and which materialise as Proposals, Projects, and so on. To realise the alignment of goals and actions, the concept of Goals needs to be linked with the concept of Actions. For example, a project is aligned with various strategies. Teams describes the human elements of the organisation that interact within the innovation process and permit the effective management of their involvement. It includes a number of subclasses, including Individuals, Groups and Organisations and so on. All individuals also need to be linked to the goals of the innovation process through an inbuilt performance appraisal system. Results specify the extensive yet concise reporting of outcome for all goals and actions, which are refined as subclasses, i.e., Publication, Deliverables, Exceptions, etc. The final concept, Community, provides team members with different collaboration spaces for their interests. Currently, it is divided into News, Forum, and other subclasses which are useful additional tools for effective communication and collaboration.

With this ontology, computers and people will be able to work cooperatively and efficiently in the innovation process. It also serves as the foundation for the design and implementation of specific tools for application in semantic innovation management.

3.2 Semantic inference

The explicit description of innovation related information makes their formal analysis feasible (e.g., goals or goal-action matching among different parts of the extended enterprise, detecting partially defined and possibly inconsistent goals, actions, or the relationships between them). This automated process reduces human intervention in innovation management. Ontology inference or reasoning, which can improve the efficiency of query, and processing of innovation related instance data, will play the role in realising the alignment analysis among different goals, actions, and results existing in different funnels. Until now, several semantic specification languages, such as RDF and OWL, have been proposed for the semantic web. Correspondingly, inference systems for semantic innovation can adopt Description Logic (Baader et al., 2003), F-Logic (Kifer et al., 1995), or Horn Logic (Sintek and Decker, 2002) as its theoretical foundation for the reasoning of innovation related information.

3.3 Semantic mediation

Different organisations can make statements about the same innovation process using different languages. Semantic mediation facilitates the interaction between shared ontology and local ontology. Semantic mediation (Ding and Foo, 2002) includes ontology consolidation, ontology mapping and alignment. Through ontology consolidation, different parts of a local ontology or several local ontologies are merged as a new shared ontology supporting a specific domain or topic (e.g., innovation management). Ontology mapping and alignment allow different terminologies and modelling styles to be linked together by creating bridges between separated pieces of knowledge in some domain specific ontology. This consolidated ontology and the bridges to domain specific ontology are then used to perform cross context information search and retrieval of innovation related data.

Semantic innovation is an adaptive complex system of goals, actions, teams and results for an entire organisation and is necessary for the extended enterprise comprising suppliers, dealers and other strategic partners. In this system, anyone, anywhere, in any organisation can have access to time critical innovated related information, that will inform the decision making process. Fast, efficient and informed decision making offers an organisation the best chance of increasing its innovation effectiveness and offers the individual and the team, the power to make change happen.

3.4 Semantic interface

According to Hendler (2001), the integration of agent technology and ontology will significantly affect the use of web services and the ability to extend programs to perform tasks for users more efficiently and with less human intervention. Agents act at the interface for the human-human and human-machine collaboration in the semantic innovation process (in fact, the software agent is also a generally accepted user interface of the semantic web). The effectiveness of such software agents will increase exponentially as more machinereadable web content and automated services (including other agents) become available. In the scenario of semantic innovation, different intelligent software agents work together in anticipating user's information requirements and thus avoid manual browsing for common information gathering tasks. The shared ontology allows for the development of search tools and intelligent agents that can automatically find any information requested by the user, and thus avoid inefficient or manual 'surfing'. With semantic innovation, any user will have instant access to all of the innovation going on anywhere within the organisation, regardless of language, structure, or location of the information. These semantically rich applications will be capable of realtime event monitoring and handling of alerts, ensuring that the stakeholder is constantly engaged and provided with priority data.

3.5 Analysis of semantic features

The semantic web technologies play an important role in reusing innovation across the extended enterprise. The semantic web is about adding machine 'processable' semantics to data. The computer can 'understand' the information and therefore process it on behalf of the human user. Semantic innovation employs the semantic related technologies to improve the collaboration capability across the extended enterprise. Every organisation or individual uses the innovation ontology to describe him/her/themself and publish the information in a trust worthy semantic web. This allows software to process these descriptions, as part of an automated search engine, to discover information about you and the communities of which you are a member. Semantic innovation has the following features.

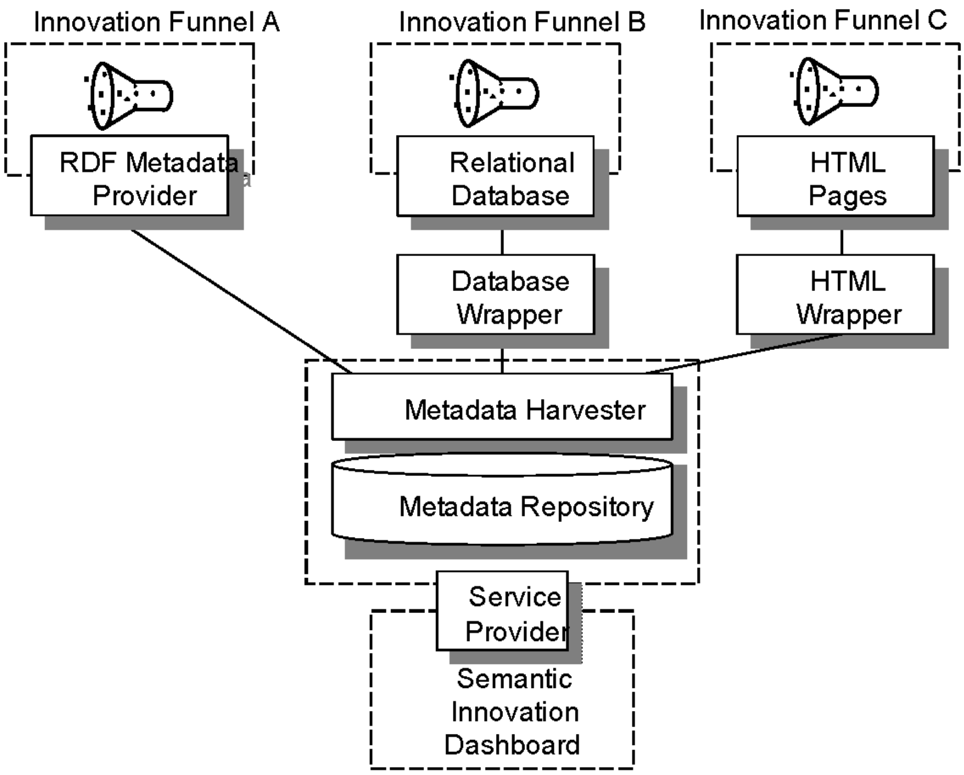
- *Goal directed*. In the extended enterprise, different individuals, projects or departments will have different goals. To reduce the uncertainty that exists in the extended enterprise, innovation ontology gives an explicit and formal specification of the goals in terms of its mission, requirements, strategies, performance indicators and so on. Then the process of formulating goals that reflect a holistic perspective of the organisation and horizontal goals matching are done through the support of an intelligent agent. It allows the organisation's overall direction to be communicated not only to the human but also the application or machine, and empowers them to participate in the innovation process, which assures the high level alignment among different but overlapping goals.
- *Action based*. The innovation process can be looked on as the temporal sequence of activities or events that occur in developing and implementing new ideas. To support the correct flow of actions through the various stagegates of the innovation process, innovation ontology based semantic mediation and inference are employed in the semantic innovation to ensure a low level of consistency within the decision making (about plans and their execution) over time. Ontology based context aware application is implemented to support the decision making process concerning selection of actions or plans for project execution.
- *Team centred*. On the one hand, employees of the extended enterprise are provided functionality by ontology enabled intelligent agents to capture minimum critical information relating to innovation. On the other hand, the innovation process requires much creativity, problem solving and teamwork, which requires a diverse skill base. Each action of the innovation process must have an efficient number of personnel (with corresponding knowledge and skills) assigned to them with a leader being ultimately responsible for action achievement. The individual member in the extended enterprise is described in terms of role (academic, engineer or manager), task description (which person is responsible for which goal or action), core competency (knowledge and skills), and relevant area of the task in the innovation process. Through the ontology inference and mediation, semantic matching is realised among goals, capability requirements of action execution, as well as roles and the expertises from different organisations.

- *Result oriented*. The reporting of outcome for all goals and actions are specified extensively and concisely in the innovation ontology, which facilitates the agent which represents the corresponding responsible manager in identifying all the actions and goals that require different kinds of attention, due to their status.

4 Functional framework of SIMS

This section gives a functional framework for SIMS. The idea of designing SIMS is to harvest innovation information (innovation metadata) from distributed and heterogeneous innovation funnels across the extended enterprise. From this harvested information, we can build a central metadata repository, which will provide an integrated semantic innovation management platform for end users. The functional framework for SIMS is illustrated in Figure 5 and comprises five main modules.

Figure 5 Framework of SIMS



4.1 Metadata provider (MP)

The main functionality of Metadata Provider is to export innovation metadata from distributed and heterogeneous innovation funnels. We have provided three kinds of solutions for different kinds of innovation funnels.

- *RDF metadata*. If an innovation funnel has provided functionality of exporting RDF metadata, then the Metadata Harvester (MH) can harvest directly. Hopefully, in the future we can have more and more of such information sources, which is also the foundation of success for the semantic web. If the RDF metadata from the innovation funnel are consistent with the Iteams innovation ontology, MH simply stores data into the metadata Central Repository. If the RDF metadata are in other ontology formats, MH not only harvests the metadata, but also harvests the new ontology. Moreover, that ontology must have been related with the Iteams innovation ontology. Ontology mapping and alignment is useful in this situation. Ontology mapping and alignment allow different terminologies and modelling styles to be linked together by creating bridges between separated pieces of context knowledge in domain specific ontologies.
- *Relational database*. Many of the innovation funnels are built on relational database technology. In the extended enterprise, it is feasible to provide an SQL interface from one innovation funnel to other systems, thus we can develop a Database Wrapper which retrieves innovation information from the database directly, and then transfers the data into RDF metadata, consistent with the Iteams innovation ontology.
- *HTML pages*. Most of the innovation funnels can be accessed only through their HTML web pages. In this situation, we need a HTML Wrapper to wrap the HTML content into the RDF format we want. There are many commercial HTML Wrapper softwares, for example, Kapow RoboSuite, which offers comprehensive and fully automated collection of dynamic information and content from any online web source. Based on this kind of software, we can easily develop our own wrapper which can export RDF metadata consistent with the Iteams ontology.

Some NLP technology can be used to retrieve information from flat texts (Pllelt, 1999), but this process can yield significant errors in translating the flat texts to RDF compatible files.

4.2 Metadata harvester (MH)

The Metadata Harvester is built in the Central Repository Server side. The task of MH is to harvest RDF metadata from different MPs, and to store the metadata in the Central Repository. Based on different requirements, different harvesting strategies can be adopted for MH. It can harvest periodically, for instance, one time per day; it can harvest manually at users designated times; it can harvest based on the updating notification of MP; and so on.

It is important to maintain the consistency of metadata between Central Repository and different innovation systems. To realise this functionality, collaboration between MP and MH can vary. Some research is already well underway in this area, for example, the Open Archive Initiative (OAI_PMH, 2005) has developed a protocol, named OAI-PMH, to support metadata harvesting. This protocol had been successfully applied in areas such as the digital library. Although it transfers metadata based on XML language, there are still some helpful ideas in using this protocol as reference for RDF metadata harvesting.

4.3 Central repository (CR)

The Central Repository is the core of SIMS. It is a buffer between the Metadata Provider and Service Provider. Innovation Metadata from different innovation funnels are harvested continuously and placed into this repository. Meanwhile, Service Provider will perform various queries on it.

Not all the innovation information is harvested into CR. We only store innovation metadata here. It is like a catalogue, which will help us to organise the innovation information across the extended enterprise. Actually, it is often not clear which information are data and which are metadata in all situations. We use the term metadata here to emphasise that we harvest only crucial information, but not all information. For example, unstructured detailed documents will not be harvested nor will the detailed contents of the various concepts stored in the innovation funnels.

As mentioned earlier, a consensual metadata schema (ontology) is not required in CR. We only need to provide certain mechanisms to allow different schema to communicate with the Iteams innovation ontology, in order for us to access innovation information in a holistic way.

Many methods can be used to build Central Repository for SIMS, e.g., file system, relational database, etc., but we prefer an RDF Database. One of the reasons is that we want to be able to store data in different RDF schemas. In a relational database, if you want to store data in different schema, you have to create new tables. However, by using a RDF database, all the RDF triples can be stored in one table, no matter what schema is being used. So when a new organisation is added to the extended enterprise (i.e., a new funnel), changes are seldom needed. There are many kinds of RDF databases available currently, e.g., RDFDB (2005), Redland (2005), etc., although they are not robust enough for industrial use, due to the limitations in mass data processing.

4.4 Service provider (SP)

The Service Provider provides the functionalities for the Semantic Innovation Dashboard to browse and search innovation information from the Central Repository. It is important to use an efficient query technology here. Another important consideration is, how to present the query results to the Semantic Innovation Dashboard. It needs to be an integrated approach, so that these two parts can communicate seamlessly. RDF access technology is required to search from the central repository. There are currently a number of RDF Query tools, e.g., RDQL and RDFQL. These tools are improving; however they are not currently practical for high speed queries of large amounts of triples. W3C has just released a working draft: SPARQL (2005) Query Language for RDF. This work will promote the development of sophisticated RDF access technology greatly.

4.5 Semantic innovation dashboard (SID)

Semantic Innovation Dashboard is the front end of SIMS to users. SID not only organises and presents query results from SP, but also connects users to different innovation systems or funnels. Users can access SID in different ways: web browser, PDA, and other intelligent devices. By accessing the Semantic Innovation Dashboard, Users can retrieve innovation information and manage innovation processes that are related to them, whenever they want, wherever they are.

5 Implementation case study

We now present a SIMS for organisations in the Health Care Sector and below we will introduce the key aspects of this system to give readers a concrete image of the implementation of SIMS.

5.1 Multiple types of MP

There are three types of innovation funnels in the Health Care Sector Case Study. For each type, we have developed a corresponding MP. The first type of innovation funnel was created recently and is based on RDF technologies. It can provide RDF metadata automatically, and the vocabularies it uses are Iteams ontology. As we have discussed in Section 4.1, harvesting metadata from this kind of innovation funnel is straightforward.

The second one is created based on Microsoft .Net Framework, whose data are stored in Microsoft SQL Server. Data stored in a relational database are structured, so it is not too difficult to transfer them to RDF format. We have created a MP to retrieve data directly from the database by SQL query, and then wrap and export it to the MH in RDF format.

The third one is created based on IBM Lotus Notes Software (Lotus, 2005). Regarding this kind of innovation funnel, we can only access data from the web pages. So we use the Kapow Robosuite to extract information from the HTML web pages, and export the information as an XML file. Then we create a web server as MP, who uses a XSLT to transfer the XML file into RDF file. Figure 6 is one of the processes to extract information from HTML pages, which is defined in Kapow Robosute RoboMakers. Figure 7 is part of the XSLT file for transferring the XML files into RDF files.

5.2 SIMS server based on RDF gateway

We use the RDF Gateway as the infrastructure of the SIMS server, which includes the MH, CR, SP and SID modules. RDF Gateway provides (1) RDF database to store RDF triples, (2) the RDFQL language to query RDF and to execute serverside tasks, (3) certain inference capability (by RULEBASE command) to support RDFS and other customised rules, (4) the RSP language to generate web content by integrating static content with serverside scripting, (5) the web server to hold the SIMS.

Figure 8 shows the key components of the server. We will discuss each part in detail below. To simplify the description, we will only focus on two innovation funnels in the extended enterprise: one is a supplier 'Supplier A'; another is a manufacturing factory 'Manufacturing Factory B'.

Figure 6 Process to extract information from HTML pages

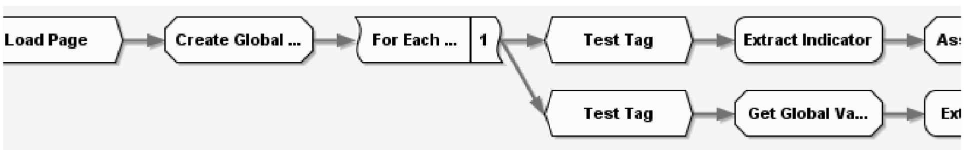
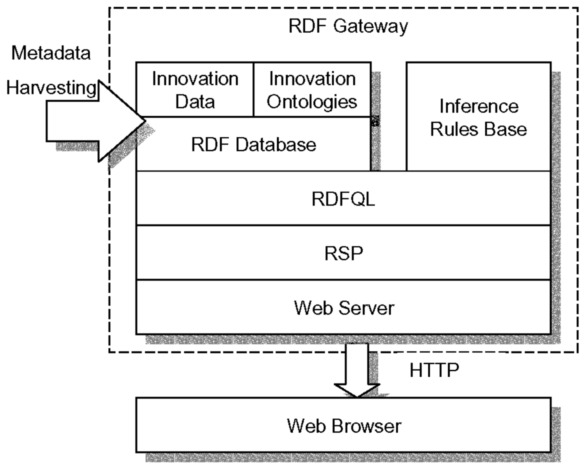


Figure 7 Part of the XSLT for transferring the XML files into RDF files

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
   <xsl:template match="/">
      <rdf:RDF xmlns:iteams="http://km.deri.org/iteams-schema#"
             xmlns:dc="http://purl.org/dc/elements/1.1/"
             xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
             xml:base="http://iteams.nuigalway.ie/trials/gamma.nsf/">
         <iteams:Company rdf:about="">
            <dc:title>Health Care</dc:title>
         </iteams:Company>
         <xsl:apply-templates select="/objects"/>
      </rdf:RDF>
   </xsl:template>
   <xsl:template match="/objects">
      <xsl:for-each select="indicator">
         <iteams:Goal>
            <xsl:attribute name="rdf:about">
               <xsl:value-of select="title"/>
            </xsl:attribute>
            <dc:title><xsl:value-of select="title"/></dc:title>
            <iteams:status><xsl:value-of select="status"/></iteams:status>
            <iteams:belong_to_company
rdf:resource="http://iteams.nuigalway.ie/trials/gamma.nsf/" />
         </iteams:Goal>
   ……
</xsl:stylesheet>
```

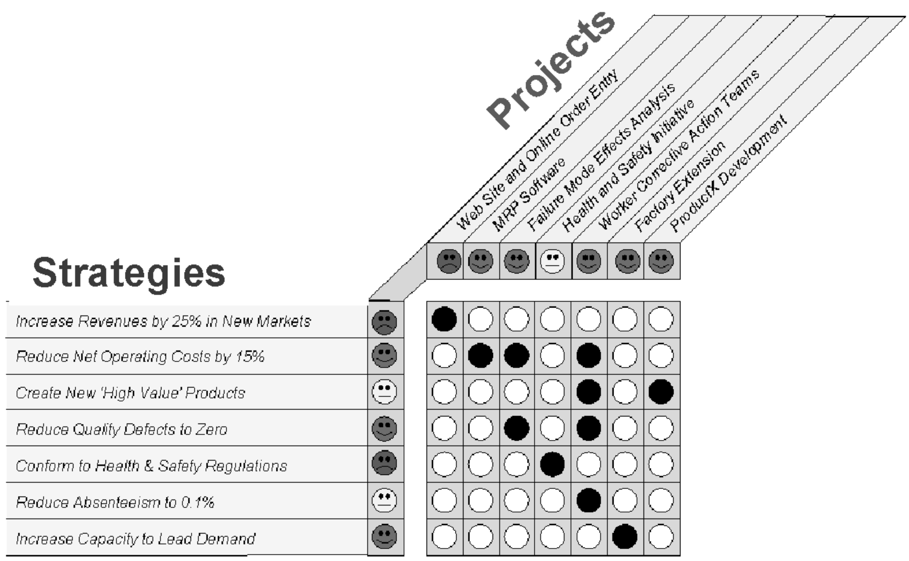
Figure 8 Key components of the SIMS server



Both companies have their own Innovation funnels, and an independent Metadata Provider has been developed for each of them. Each MD will provide RDF triples to Central Repository in its own schema: The schema of 'Supplier A' is called Iplanner; the schema of the 'Manufacturing Factory B' is the Iteams ontology.

Iteams ontology is the shared ontology of SIMS, but what about Iplanner? The main idea of Iplanner is a Strategies and Projects matrix (Iplanner, 2005), as shown in Figure 9. Each project, strategy, and their relationship are measured by qualitative indicators. Thus, projects, strategies, and their relations form a 'matrix', which is an excellent channel to monitor in order to achieve innovation efficiency and effectiveness. In fact that Iplanner is a subset of Iteams ontology: e.g., a 'project' in Iplanner is a kind of 'action' in Iteams, and 'strategy' in Iplanner is a kind of 'goal' in Iteams. Based on these recognitions, we can align Iplanner to Iteams easily. We will discuss it in detail later on.

Figure 9 Iplanner matrix



We used the RDF Gateway to harvest innovation metadata from the MDs. Two strategies are adopted: One is harvesting periodically, which is supported by a 'Timer' procedure; another is harvesting manually, users can decide when to harvest. Figure 10 shows some of the key source codes for harvesting functionality: Old information is deleted, and then new information is retrieved from a different Data Source and inserted into the database. These codes are written in RDFQL, an RDF query language provided by RDF Gateway.

Figure 10 Source codes of metadata harvester in RDFQL

```
……

DELETE {?p ?s ?o} FROM innovation USING innovation WHERE{?p
?s ?o};

var S_A = new DataSource("inet?url=http://www.supplier_a.net/
iplanner_rdf.aspx&parsetype=rdf");

var MF_B = new DataSource("inet?url=http://
www.manufacturing_factory_b/iteams_rdf.rdf&parsetype=rdf");

INSERT {?p ?s ?o} INTO innovation USING #S_A #MF_B WHERE {?p
?s ?o};

…...
```

The harvested metadata will be stored in a RDF Database. Two tables have been created in the database for innovation metadata storage: Innovation (stores innovation metadata) and Innovation_Ontology (stores ontology information). We use two tables here, because their update strategies are different from each other: Innovation Data is updated frequently, while Innovation Ontology which is very stable is updated less frequently. Figure 11 shows the table of Innovation information.

Figure 11 Innovation data table

tables/innovation

Predicate	Subject	Object
[rdf:type]	[http://www.manufacturing_factory_b/iteams_rd	[http://km.deri.org/ipla
[http://purl.org/dc/elements/1.1/tit	[http://www.manufacturing_factory_b/iteams_rd	Manufacturing Factory B
[http://purl.org/dc/elements/1.1/tit	[http://www.manufacturing_factory_b/iteams_rd	Investigation of the Sto
[rdf:type]	[http://www.manufacturing_factory_b/iteams_rd	[http://km.deri.org/itea
[http://km.deri.org/iplanner-schema#	[http://www.manufacturing_factory_b/iteams_rd	11/09/2005

To support the integration of the Iplanner and Iteams ontologies, some additional information has to be added to the definition of Iplanner Ontology to allow it to communicate with Iteams, while also allowing its own vocabularies to remain unchanged. This is called mediation. Figure 12 shows some of the Iplanner Ontology after the update (the italics are newly added contents). For example, by this statement:

<rdfs:subClassOf rdf:resource="&iteams;Goal"/>

RDF Gateway can infer that Iplanner:Strategy is a sub class of Iteams:Goals.

Figure 12 Updated schema of Iteams

```
……
<rdfs:Class rdf:about="&iplanner;Strategy"
        rdfs:label="Strategy"
        rdfs:comment="Strategy of A company.">
        <rdfs:subClassOf rdf:resource="&iteams;Goal"/>
</rdfs:Class>
<rdfs:Class rdf:about="&iplanner;Project"
        rdfs:label="Project"
        rdfs:comment="Project of A company.">
        <rdfs:subClassOf rdf:resource="&iteams;Action"/>
</rdfs:Class>
……
```

To realise the functionality of inference, we have created some inference rules based on the definition RDFS Definition in RDFQL, e.g., 'rdfs:subClassOf' has been modelled as a rule, as shown in Figure 13. These rules will be used by every query request in the RDFQL language.

After this work is finished, users can query what they want from the system, no matter where the information comes from i.e., 'Supplier A' (in schema Iplanner) or from 'Manufacturing Factory B' (in schema Iteams). For instance, if we want to know 'What actions is 'Red' in the extended enterprise?' The Service Provider will send this RDFQL (see Figure 14) to the Central Repository.

Figure 13 Rule of rdfs:subClassOf

```
RULEBASE schema
{
   INFER {[rdf:type] ?s ?class} FROM
      {[rdf:type] ?s ?subclass} AND {[rdfs:subClassOf] ?subclass ?class};
};
```

Figure 14 Query by RDFQL

```
SELECT ?company_id ?company ?action
USING innovation innovation_schema
RULEBASE schema
WHERE
  {[rdf:type] ?action_id [iteams:Action]}  AND
  {[iteams:title] ?action_id ?action} AND
  {[iteams:belong_to_company] ?action_id ?company_id} AND
  {[iteams:status] ?action_id "Red"} AND
  {[iteams:title] ?company_id ?company};
```

The results are shown in Figure 15. We can see that there are two actions whose status is 'Red'. One is from 'Supplier A', another is from 'Manufacturing Factory B'. Although they are represented in different schema, they all can be retrieved by a uniform query request.

Figure 15 Query results

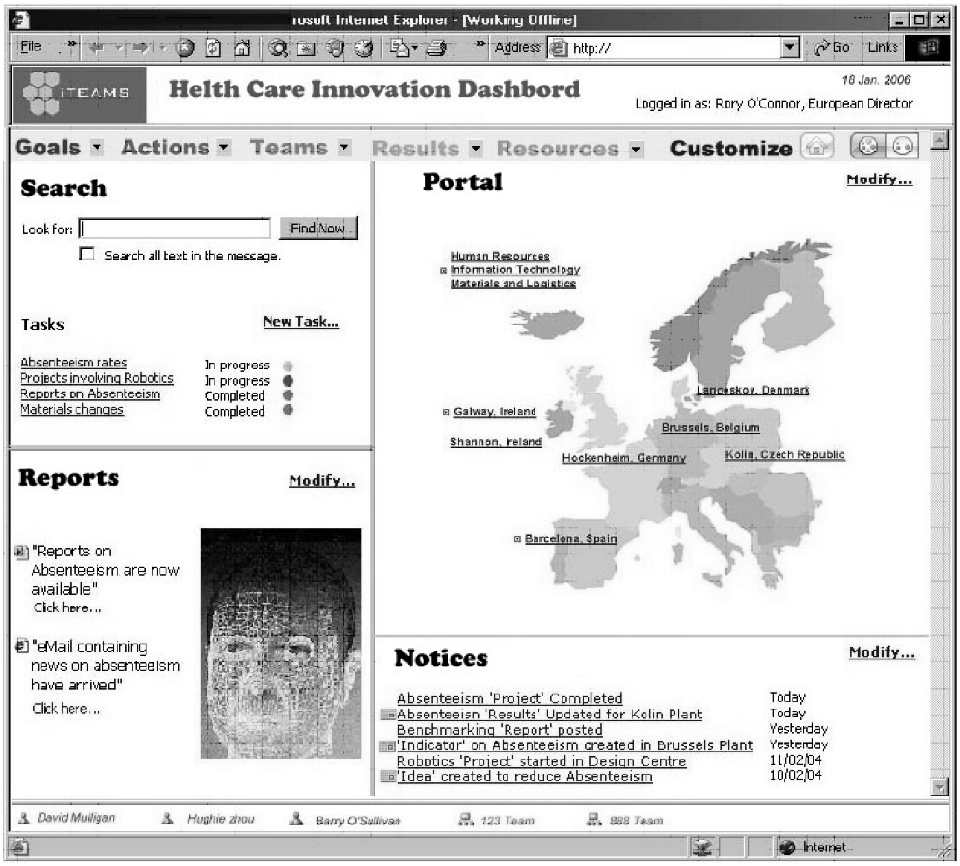
```
company                          action                status
--------------------------------------------------------------------
'Supplier A'                     'Staff retraining'    'Red'
'Manufacturing Factory B'        'Re-cycling sheme'    'Red'
```

Finally, the query results will be sent to the front end to present to users. We used RSP (RDF Server Pages) to create the front end user interface, which operates RDFQL to retrieve query results from RDF Database, and then provide it to end users in a friendly way through HTTP. Incidentally, the term 'red' manifests itself as a red traffic light to indicate that a particular action requires attention.

Figure 16 shows one instantiation of the interface for SIMS. There are four major modules in the SID for end users.

- *Search*. In this module, users can search what they want from among the innovation information from the extended enterprise.
- *Portal*. It is an entry point to different innovation funnels across the extended enterprise. Users can go into the special funnel to look at detailed innovation information.
- *Reports.* This module provides some customised search agents to retrieve the innovation information which concerns the related users.
- *Notices*. This module uses push technology to provide the emergent information to users.

Figure 16 Interface of SID



6 Conclusions

Innovation management is crucial for enterprises to succeed in competitive environments. To implement innovation management in the extended enterprise, we need to deal with the problems of integration of distributed and heterogeneous information systems. Semantic web technology is a promising direction for information integration. In this paper, we introduce a vision for Semantic Innovation Management. We use semantic web technologies, especially RDF data access technologies and metadata harvesting technology, to build a semantic innovation management system for the extended enterprise. The current system is being tested in a multinational organisation and shows some promise in reducing costs associated with data integration and automatic alerting of users. Validation work is ongoing and full tests will require deployment in a truly extended enterprise scenario.

Currently, the implementation of SIMS still has some disadvantages. The query function is limited due to the software we use. And data protection problem may occur because it is difficult for a data source to just provide part of its data. We will continue to work to solve these problems. We are also considering developing more services for

SIMS. One direction is matching of innovation elements in different abstraction levels (or hierarchies). In any extended enterprise there will be hundreds of goals and thousands of actions which will be related to each other through complex child-parent hierarchies. In addition, goals and actions will overlap. Since the goals and actions are created separately in the extended enterprise, inconsistencies will evolve. Another direction is building more services at the user interface that can push information to the user through the innovation dashboard. The research team is currently experimenting with social networking concepts where a user's internet profile can be used to increase collaboration with other users in the system. And OWL will be used as the ontology language instead of RDF Schema. By adding more vocabulary for describing properties and classes, OWL will facilitate greater machine interpretability of web content.

Acknowledgements

This research is supported in part by the Science Foundation of Ireland (www.sfi.ie). Our colleague Jianqiang Li has also made a contribution to the publishing of this research.

References

Askme (2005) http://www.askmecorp.com/.

Baader, F., Calvanese *et al*. D. (2003) *The Description Logic Handbook*, Cambridge University.

Brown, S. and Eisenhardt, K. (1997). 'The art of continuous change: linking complexity theory and time-paced evolution in relentlessly shifting organizations', *Administrative Science Quarterly*, Vol. 42, pp.1–34.

Browne, J. (1998) 'The extended enterprise', Molina, A., Sanchez, J.M. and Kusiak, A. (Eds.): *Handbook of Life Cycle Engineering Concepts, Tools and Techniques*, Chapman & Hall, London, pp.3–29.

Ding, Y. and Foo, S. (2002) 'Ontology research and development, Part 2 – A review of ontology mapping and evolving', *Journal of Information Science*, Vol. 28, No. 5, pp.375–388.

Dooley, L. and O'Sullivan, D. (2000) 'Systems innovation manager', *International Journal of Production, Planning and Control*, Vol. 11, No. 2, pp.369–379.

GoldfireInnovator (2005) http://www.goldfire.com/.

GoupSystems (2005) http://www.groupsystems.com/.

Grossman, G. and Helpman, E. (1991) *Innovation and Growth in the Global Economy*, The MIT Press, Cambridge, Mass.

Hendler, J. (2001) 'Agents and the semantic web', *IEEE Intelligent Systems Journal*, Vol. 16, No. 2, pp.30–37.

Huxham, C. (1996) *Creating Collaborative Advantage*, SAGE, London.

IdeaChain (2005) http://www.imaginatik.com/web.nsf/docs/prod_ideachain.

IdeaFisher (2005) http://www.ideafishing.com/.

Intellidimention (2005) RDF Gateway, http://www.intellidimension.com/.

Iplanner (2005) http://www.innovationchannel.net/.

Kapow (2005) *Kapow Robosuite*, http://www.kapowtech.com/.

Kifer, M., Lausen, M.G. *et al.* (1995) 'Logical foundations of object oriented and frame-based languages', *Journal of the Association of Computer Machinery (JACM)*, Vol. 42, No. 4, pp.741–843.

Lotus (2005) http://www-306.ibm.com/software/lotus/.

Miller, E. (2004) 'The W3C's semantic web activity: an update', *IEEE Intelligent Systems: Semantic Web Chanllenge*, Vol. 19, No. 3, pp.95–96.

MindManager (2005) http://www.mindjet.com/.

NextNet (2005) http://www.generalideasinc.com/.

OAI_PMH (2005) http://www.openarchives.org/.

OWL (2005) http://www.w3.org/2004/OWL/.

Pllelt, D. (1999) 'An introduction to information extraction', *Artificial Intelligence Communications*, Vol. 12, No. 3, pp.161–172.

RDF (2005) RDF Primer, http://www.w3.org/TR/rdf-primer/.

RDFDB (2005) http://www.guha.com/rdfdb/.

Redland (2005) http://librdf.org/.

Semantic_Web_Challenge (2005) http://challenge.semanticweb.org/.

Sintek, M. and Decker, S. (2002) 'TRIPLE – a query, inference, and transformation language for the semantic web', *Proceeding of International Semantic Web Conference (ISWC)*, Sardinia, Italy, pp.364–378.

SPARQL (2005) *SPARQL Query Language for RDF*, http://www.w3.org/TR/rdf-sparql-query/.

Staab, S. and Studer, R. (2004) *Handbook on Ontologies*, Springer.

Stamm, B.v. (2004) *Innovation Management Survey*, http://www.imaginatik.com/web.nsf/docs/idea_reports_wp3.

Teamstorm (2005) http://www.teamstormsolutions.com/.

Zairi, M. (2004) *Best Practice: Process Innovation Management*, Oxford, Butterworth-Heinemann.