



# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Advanced knowledge system for coatings and the gas turbine MRO industry

### Conference or Workshop Item

#### How to cite:

Chandler, Peter; Hall, Wendy; Shadbol, Nigel; Alani, Harith and Szomszor, Martin (2008). Advanced knowledge system for coatings and the gas turbine MRO industry. In: International Thermal Spray Conference & Exposition: Thermal Spray Crossing Borders (DVS-ASM), 2-4 Jun 2008, Maastricht, The Netherlands.

For guidance on citations see [FAQs](#).

© 2008 The Authors

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://asmcommunity.asminternational.org/portal/site/www/AsmStore/ProductDetails/?vgnextoid=8e8288fbd8e4c110VgnVCM>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

# Advanced knowledge system for coatings and the gas turbine MRO industry

P E Chandler, W Hall, N R. Shadbolt, H Alani and M Szomszo. Southampton/UK

The growth of data generated within thermal spraying is, for many, a daunting business. Yet, this growing resource represents a largely untapped and potentially valuable asset capable of providing “knowledge” rather than just “information”.

Many companies already use a range of Web based tools. However, the Web itself is changing and the vision for the future, the “Semantic Web”, is set to revolutionise how business will be done. One important aspect of this Web “future” is that web pages will be greatly enriched and data will have additional information (tags) which help to describe it and more significantly, put the data into a context. This will enable machine readability and the use of query languages to ask direct questions.

Following on from ideas introduced at ITSC 2007, a proof of concept demonstrator has been built for thermal spray coatings used in the Maintenance Repair and Overhaul (MRO) of gas turbines. A system has been built which stores and manipulates a range of data including; aircraft deliveries, RSS feeds of aircraft sales, engine types, MRO business details, thermal spray coatings and market dynamics. This paper presents the development of this system and discusses its future potential.

## 1 Introduction

Like most business afflictions the problem of “data-overload” can be looked at as an important and even significant commercial opportunity. In order to make this turn-around happen a method is needed which enables the consolidation of data held in a wide range of formats and sources plus the ability to interrogate this vast data source in order to extract knowledge. Given the exponential growth of data with time any system must also be fully expandable and, of course, be capable of automatic operation!

The initial reaction to this issue may be to ask “why not just use conventional spread sheets, databases and web tools?” This is understandable but to just “Google it” does not actually solve a basic problem namely how does one piece of data relate to another? There is no knowledge behind a simple word search. A common example often used to show this is to search on the word “apple”. The results will include references to computers and fruit in order of commercial sponsorship, and a mix of citations, number of hits, meta data and word searches. The only real knowledge involved will lie with the user who has to refine the search terms. What is needed is a way to add some knowledge to the data itself.

In fact this aspect is currently being worked on by developers of what is seen to be the next generation of the Web, the Semantic Web [1,2]. The new vision will be to enrich data by incorporating additional information (tags) which both help to describe it but more significantly, put the data into a context. Furthermore, this Semantic Web content will be fully machine readable and allow the use of query languages to ask direct questions. In essence, this future generation of the Web will contain knowledge as well as documents. The term “semantic” simply means the meaning of things. In this context it is used

to define the meaning of data but more importantly how they interrelate.

Establishing interrelationships however is in itself a complex field of study. A useful starting point may be the development and use of taxonomies which aim define and organise objects. The ASM use a taxonomy as a browse feature “to provide relevant results from all of content sources in the ASM Community” [3]. Adding further information about relationships goes one step further and creates what is called an ontology. This idea was introduced at ITSC 2007 [4] and is a central feature of the system discussed in this paper. A more detailed treatment of this field may be found elsewhere [5]. Some examples of some defined relationships within an aviation ontology are provided in table 1.

Table 1: Some examples of ontology classes and relationships, instance triples in aviation and MRO shops.

Ontology classes and relationships		
Aero-Space Companies	Manufacture	Aircraft
Aero Engines	Power	Aircraft
Airlines	Operate	Aircraft
Gas Turbine	Is a	Aero Engine
Examples of instance triples - aviation		
Boeing	Manufacture	777-200
Airbus	Manufacture	Airbus 320
BA	Operate	777-200
777-200	Is a	Aircraft
IAE	Manufacture	Aero Engines
V2500	Is a	Aero Engine
Examples of instance triples – MRO Shops		
MRO shops	Overhaul	Aero Engines
MTU, Zhuhai	Operate	An MRO shop
MTU, Zhuhai	Is located in	Far East
MTU, Zhuhai	Overhaul	V2500

Instances arising from these sorts of ontological relations can be used to infer various sorts of knowledge. For example, it can be inferred that Boeing is an aerospace company from the fact it makes the 777-200, the 777-200 is an aircraft and aerospace companies make aircraft.

What makes the Semantic Web so appealing is that these instances of objects and relations in the ontology can be expressed in Web standard languages such as the Resource Description Framework (RDF) [6] that allows the data to be accessed universally.

In terms of application, this emerging technology could potentially be applied to any field. In order to demonstrate its potential in the surface engineering sector a small demonstrator has been built which uses data on aircraft, airlines, engines, coatings and markets to provide a business tool for MRO shops. Interestingly, Rolls-Royce are working at the other end of this market and have recently demonstrated its Shop Visit Planner [7, 8]. In this case “Web agents” provide a planning and scheduling system for aero-engine repair and overhaul. The agents, representing fleet managers and other stakeholders, adaptively negotiate to schedule engine shop visits. Clearly these emerging technologies will eventually embrace the whole manufacturing, customer and supply chain within this sector.

This paper sets out a brief introduction to the semantic technology being used and how it works, a description of the demonstrator that has been built and a presentation of some of the results and capabilities. Finally, suggestions for how this technology could be developed are also discussed.

## 2 Semantic Technology

This section provides an overview of the technology used to build the MRO-Assist demonstrator. Data describing aircraft types, the engines they use and the MRO shops that can repair them, is available from a number of public data sources. Manufacturers such as Airbus and Boeing provide detailed information on the number of planes ordered in formats, such as Excel and PDF. The list below outlines the data sets used in the MRO-Assist application, including the source data format, and the information they contain:

- Airbus-Orders: Airlines, aircraft, number ordered / delivered / operational (Excel).
- Boeing-Deliveries: Airline, country, region, model, engine fitted, order date, number ordered (Excel).
- Engine Yearbook 2007: Aircraft, model, type, number of engines, compatible engines
- Engine Yearbook 2007: MRO, Company Name, Address, Approvals (pdf).
- CIA Fact-book: Country coordinates, flags, demographic / economics (rdf).

The data was then processed in three-stages which: create a suitable ontology, convert the data to RDF (a number of bespoke scripts were developed to convert these data sources to RDF) and link it to other relevant sources. These steps are described briefly below.

### Ontology Construction

The first stage was to create a suitable ontology to capture the semantics and structure of each data source. While much of the data is replicated between different sources, both in terms of instances (e.g. that American Airlines is an airline from the U.S.A) and concepts (e.g. aircraft, engine), subtle differences in the granularity of information and conceptualisation makes the construction of a domain ontology different sources, both in terms of instances (e.g. that American Airlines is an airline from the U.S.A) and concepts (e.g. aircraft, engine), subtle differences in the granularity of information and conceptualisation makes the construction of a domain ontology infeasible. For example, the Boeing deliveries data set specifies the date when an order was made, which airline made the order, for which model of plane the order was made, how many were ordered, and which engine was fitted.



Figure 1 Ontology map

With the Airbus orders data set, only summaries are provided for each airline stating how many of each model has been ordered, how many have been delivered, and how many are currently operational. No information is provided by Airbus regarding the engine types that were fitted. While it is possible to use data from each data set to estimate the number of engines currently in operation, the calculations will differ subtly for each data source. Therefore, small ontologies were created, with low complexity, for each individual data set that can be linked together at the ontology level. Figure (1) provides a visual representation of the Boeing deliveries ontology.

### RDF Generation

With suitable ontologies in place to describe each data source, the next stage was to convert the raw data to RDF. For data sets provided in Excel format, the content was first exported to a text-based representation (comma separated values in this case) to facilitate compatibility with programming languages. Small Java programs were then written, making use of the Jena API, to parse the source data into appropriate concept instances and properties. These in-memory models were then exported to RDF. Conversion of PDF documents to RDF followed a similar pattern but required more manual intervention to massage the source text to a machine-readable format.

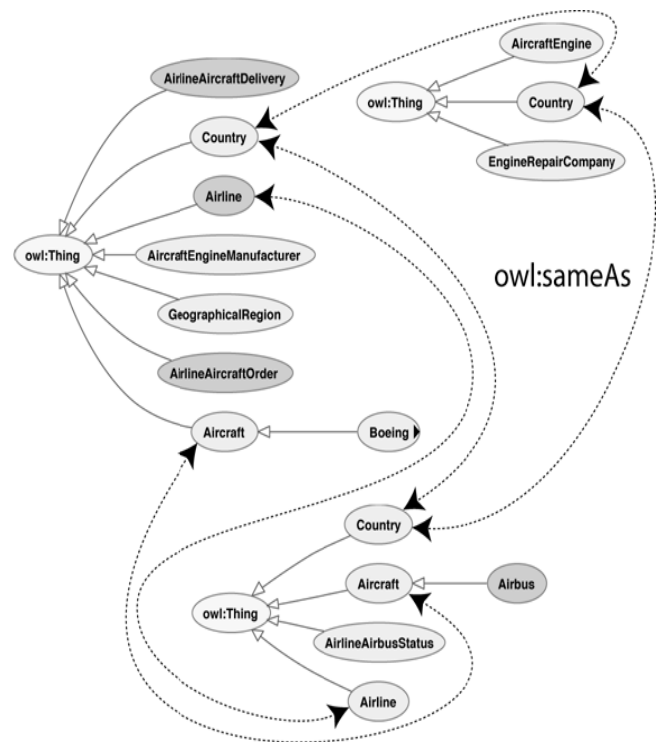
(Note that languages such as RDF represent their content as triples consisting of a subject, a relationship operator and an object. Some examples of these are found in Table 1. A “triple store” storage system is used to provide easy access to the data [9].

### Data Linking

Once the source data had been converted into RDF, concept instances from each data set that refer to the same entity were linked using a Web Ontology language, OWL [10]. In this language, the owl:sameAs property is used to link identical terms found within different datasets as illustrated in Figure (2).

This linking provides the power to query over multiple data sources simultaneously. For example, by linking instances of airlines and countries between the Boeing and Airbus data sets it is possible to query the knowledge base for all orders made by region (e.g. North America or Europe), by country, or by airline. To achieve this linking, a number of mapping files were manually created, as summarised in the list below:

- airbus\_engine\_mapping.rdf
- boeing\_manufacturers\_mapping.rdf
- cia\_country\_mapping.rdf
- overhaul\_engine\_mapping.rdf



**Figure 2** Links between ontologies

### The Overall system

Once the RDF data, ontologies, and mapping files were created, all RDF was imported into the Triple Store. This provided an end-point through which query languages could be used to extract information. In this case the query language used was **SPARQL** [11] which was able to provide summaries as well as the statistics needed for the forecasting algorithm. A PHP+ Javascript front-end was developed to provide the application interface using HTTP to query the triple store. A number of caching procedures were used to improve performance since many queries required large amounts of processing.

## **3 The MRO Assist Demonstrator**

There are three principal objectives underlying the MRO-Assist demonstrator:

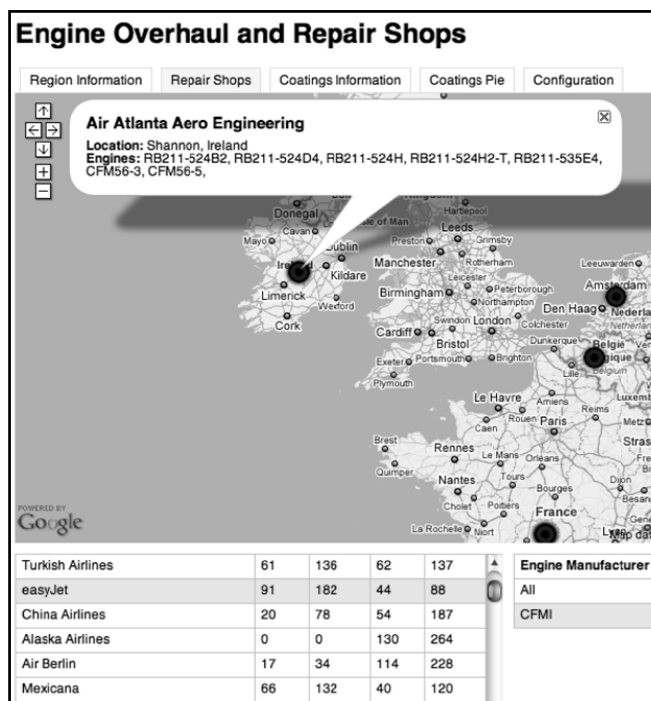
- To consolidate data from multiple heterogeneous data sources into a single representation that facilitates the extraction of information that is not possible from the individual sources.
- To provide a graphical interface to view and explore data.
- To forecast future MRO re-coating business by estimating the number of planes and engines in operation at a regional level.

The first aim was realised using the Semantic Web technologies described above. The second two were

achieved through the development of a graphical, web based application that exploits the integrated and stored data.

## Data Presentation

In order to aid visualisation, a Google map was used for the interface. This provides a convenient way of accessing and interrogating the system. For example, Figure 3 shows the locations of MRO shops. Each of these can be selected to view additional information such as the company details and repair approvals held.



**Figure 3** Display Interface showing MRO shops engine capabilities/approvals

A tabular data browser (much like contemporary RDF browsers such as mSpace [12] and Tabulator [13]) is also supplied to show further information. An example of this, shown in the screen shot in figure 3, which has a summary of airlines and the number of engines and planes they own and which aircraft manufacturer supplied them. Going deeper by highlighting a particular airline, the engine inventory can be viewed showing types and numbers currently in service.

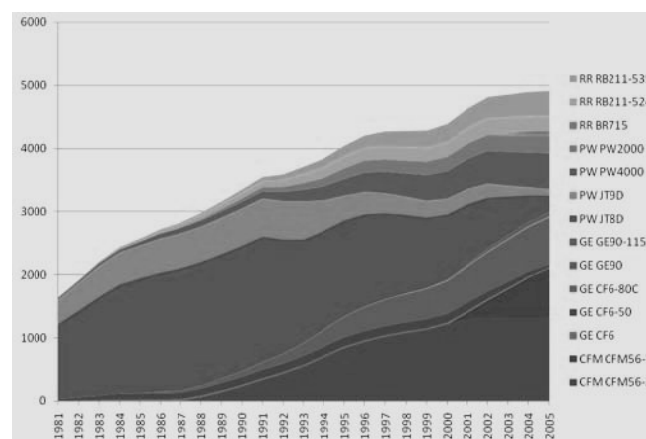
The maps and tabular browser interface are just examples of how the stored data can be presented but are actually based on questions posed by the user to provide specific answers to some detailed questions about the MRO industry. The demonstrator currently has the following capabilities:

**Regional Information.** Markers for every country in the dataset, using coordinate data extracted from the CIA factbook. On selecting a particular country,

additional information (such as the country's flag, population, number of airports, etc.) is displayed. Note that it would be easy to add financial data such as GDP per head which is an indicator of numbers of the population who take flights.

**Repair Shops** This map view displays a marker for each repair shop. By clicking on a repair shop it is possible to view the company name, it's address, and the engine approvals.

**Coatings Information** This tab will place a small graph for each region on the map showing the total MRO business generated for engine re-coating. A user may click on the graph snap-shot to obtain a full size, more detailed version.



**Figure 4** Historical growth in MRO business by engine type (1981-2005)

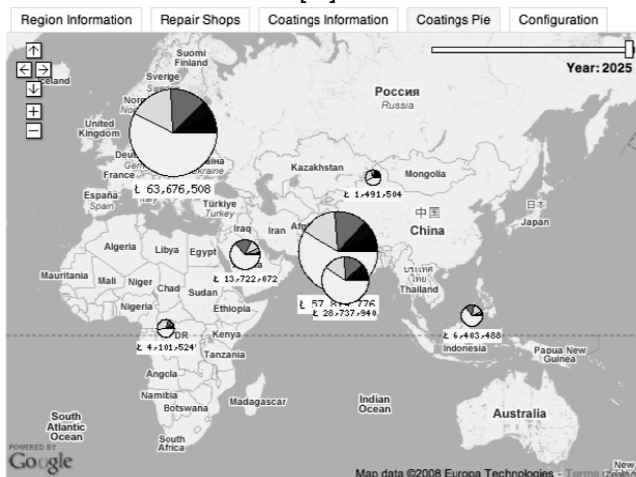
**Coatings Pie** This tab is used to display the total value of the MRO re-coating business generated in each region. The size of each chart is relative so the user can easily see how the business in each of the regions compares to each other. A time slider is also displayed, to select a year from 1965 – 2025.

A configuration screen is also provided to enable the user to tweak the parameters used in the forecasting algorithm.

In this study a range of data sources have been linked together using the newly developed semantic tools (ontologies, triple store, links etc). In this system the linked data can be easily interrogated (for example by using SPARQL query language [11]) to extract the required knowledge. For example by querying both the Airbus and Boeing datasets for airplane orders restricted to a particular region, it is possible to obtain an estimate for the total number of planes in operation for that region. By using the Aircraft-Engine ontology to examine which engines are fitted to these planes, it is possible to determine the number and types of engines in operation and therefore estimate the MRO business that would be generated.



[a.]



[b.]

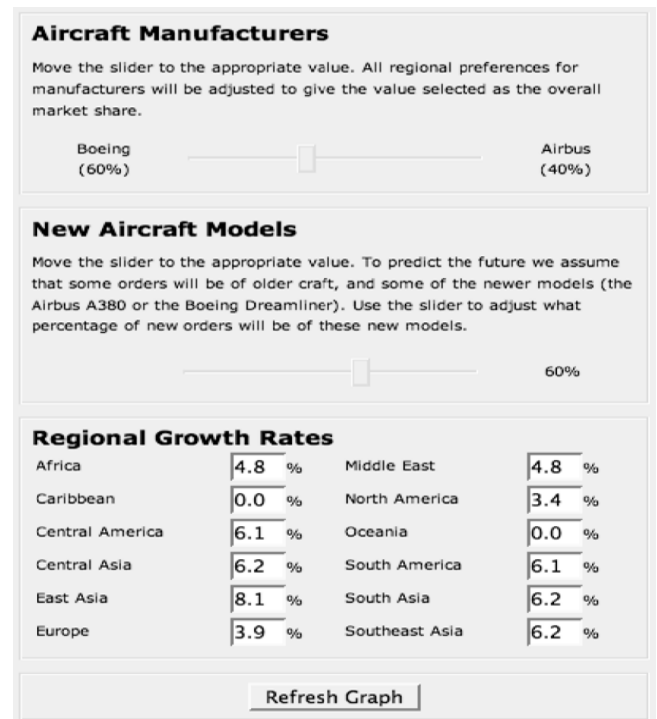
**Figure 5** Coating business (a) 2007 and (b) 2025

For values in the future, a forecasting algorithm is used to estimate the MRO coatings business generated.

#### 4. The MRO Coating Business

As already mentioned, this system is both flexible and scalable. Any amount of further data sets can be easily integrated by adding extra ontological terms and links where necessary. So, for example, adding specific components to engines such as fan casings, combustors and turbine blades and details of coatings used on these components can readily be added and used to answer questions about market size. To estimate the MRO re-coating business that would be generated in the future, we use a forecasting algorithm to estimate the number of planes that will be in operation in a particular region at a given date in the future. From this, we can infer the number of engines that will require overhaul, and therefore predict the business generated.

The screen-shot of figure 6 shows the configuration screen with all the parameters that are used by the forecasting algorithm. These parameters are split into three categories:

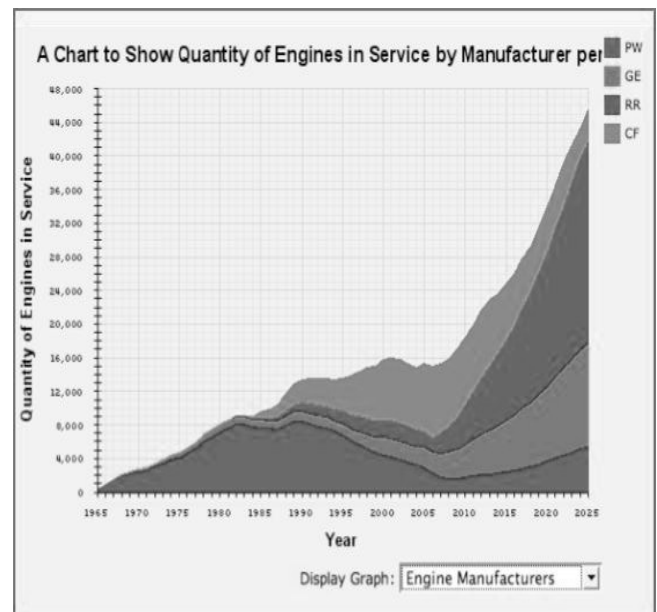


**Figure 6** Configuration Screen

**Aircraft Manufactures** The slider bar in the upper left allows the user to modify the percentage of the market that would be held by Airbus and Boeing in the future.

**New Aircraft Models** This sliders bar dictates what percentage of planes purchased in the future will be new models (such as the Airbus A380 or the Boeing Dreamliner).

**Regional Growth Rates** Growth rates are used to forecast the number of planes sold in each region.



**Figure 7** Configuration Screen

The first step in the forecasting algorithm extracts historical plane order data for each region from the triple store using SPARQL queries. These results are used to calculate the number of planes and engines currently in operation, as well as the regional preferences towards a particular engine manufacturer. The second step estimates the number of planes that would be ordered in the future using the three parameters (outlined above) set by the user. The regional growth rate is used to predict the number of new aircraft that will be ordered. Using 5 years of historical data, and the user parameters that set the number of new aircraft that would be ordered and the Airbus – Boeing market share, an estimate for the aircraft models that would be ordered is generated. The final step uses the regional preferences towards a particular engine manufacturer to estimate the engines that would be fitted to the newly ordered aircraft. From this figure, it is possible to predict the total number of engines that will be in operation and therefore the MRO coatings business generated.

## 5. Summary and Discussion

The MRO-Assist application has shown how a number of datasets can be integrated to provide an intelligent knowledge base for the civil engine MRO business. That this can be expanded to cover any aspect of the business has been demonstrated for one particular activity i.e. coatings. However, there is potential for using this technology as a platform for any number of applications.

Companies such as Airbus already have a simple system for stock control whereby parts are bar-coded. When used they are scanned and this information is passed to the supply chain in order to optimise stock control. Using this new technology much more ambitious systems could be realised. Companies could integrate their own data about parts, materials, processes, quality, approvals, commercial detail and management systems and generate a common system for the whole company and/or supply chain. This could form web based tool which was expandable in real time.

Another, more general example could be the generation of a system for Surface Engineering. There is a wealth of data available on surface engineering processes, materials, coatings, wear, corrosion etc which lies in a disparate set of sources. Increasingly, data from processing is being generated as well as in-service data. In the latter case, data mining is becoming an increasingly important diagnostic tool.

The ability to integrate all of these aspects into a single, intelligent system would provide designers and users of surface engineering and coatings with a very important tool.

## 6. References

1. Berners-Lee, T., Hendler, and O. Lassila, "The Semantic Web", Scientific Am., May 2001, pp. 34–43.
2. Shadbolt, N., T. Berners-Lee, T. and W. Hall, W. , 2006 The Semantic Web Revisited. IEEE Intelligent Systems, May/June 2006.
3. ASM. <http://asmcommunity.asminternational.org/portal/site/asm/ResearchLibrary/>
4. Chandler P. E., "Semantic technology as a core business activity for thermal sprayers". In Proc. ITSC 2007, Beijing. 14-16 May 2007. Poster Session.
5. Guarino, N. Formal Ontology and Information Systems. Proceedings of FOIS'98, Trento, Italy, 1998
6. RDF. [www.w3.org/TR/rdf-primer](http://www.w3.org/TR/rdf-primer).
7. Stranjak, A., Dutttag, P.S., Ebdon, M., Rogers, A., Vytelingum, P. "A Multi-Agent Simulation System for Prediction and Scheduling of Aero Engine Overhaul", to be presented at AAMAS-08.
8. Rolls-Royce. Demonstration of "Shop Visit Planner" made at ARGUS II Conference, BERR Conference Centre, London. 28 February 2008.
9. Steven Harris S. And Gibbins N., Triple-Store: Efficient bulk RDF storage. Proc. 1st Int. Workshop on Practical and Scalable Semantic Systems (PSSS'03), Sanibel Island, FL, USA, 2003.
10. OWL. <http://www.w3.org/TR/owl-features/>
11. SPARQL. <http://www.w3.org/TR/rdf-sparql-query/>
12. schraefel, m. c., Karam, M. and Zhao, S. (2003) "mSpace: interaction design for user-determined, adaptable domain exploration in hypermedia. In: AH 2003: Workshop on Adaptive Hypermedia and Adaptive Web Based Systems, August 26, Nottingham, UK
13. Berners-Lee T., Chen Y., Chilton L., Connolly D., Dhanaraj R., Hollenbach J., Lerer A. and Sheets D.. "Tabulator: Exploring and Analyzing linked data on the Semantic Web". 3rd International Semantic Web User Interaction Workshop, Collocated with the 5th Int. Semantic Web Conf., Athens, Georgia, USA, 2006.