Next Generation Search Interfaces – Interactive Data Exploration and Hypothesis Formulation

Jane Hunter[1], Katya Falkovych[2], and Suzanne Little[3]

[1] DSTC, Brisbane, Australia 4072
jane@dstc.edu.au
[2] CWI, Amsterdam, Netherlands
katya@cwi.nl
[3] ITEE, The University of Qld, Brisbane, Australia 4072
slittle@dstc.edu.au

Abstract. To date, the majority of Web search engines have provided simple keyword search interfaces that present the results as a ranked list of hyperlinks. More recently researchers have been investigating interactive, graphical and multimedia approaches which use ontologies to model the knowledge space. Such systems use the semantic relationships to structure the assimilated search results into interactive semantic graphs or hypermedia presentations which enable the user to quickly and easily explore the results and detect previously unrecognized associations. More recently, the proliferation of eResearch communities has led to a demand for search interfaces which automate the discovery, analysis and assimilation of multiple information sources in order to prove or disprove a particular scientific theory or hypothesis. We believe that such semi-automated analysis, assimilation and hypothesis-driven approaches represent the next generation of search engines. In this paper we describe and evaluate such a search interface which we have developed for a particular eScience application.

1 Introduction

Traditionally Web search engines have provided simple keyword search interfaces which retrieve relevant documents and present the results as a list of hyperlinks which the user has to click through one at a time [1]. The Semantic Web [2] is beginning to enable more interactive, graphical and multimedia search-and-browse interfaces which leverage semantic relationships between retrieved information objects. Technologies such as machine-processable semantic annotations, ontologies and semantic inferencing rules and engines are enabling automated reasoning about complex relationships and a shift towards automated integration and analysis of retrieved documents and data. Researchers are developing systems that can assimilate, structure and present large amounts of mixed-media, multi-dimensional data and information as interactive semantic graphs or hypermedia presentations – greatly enhancing the capacity of domain analysts to process information and mine new knowledge.

In addition, the recent proliferation of eScience communities has led to a demand for more sophisticated search interfaces which assist users to interpret experimental

R. Heery and L. Lyon (Eds.): ECDL 2004, LNCS 3232, pp. 86–98, 2004.

data in order to prove or disprove a particular scientific theory or hypothesis. We believe that eScience will drive the next generation of search engines – interactive 'hypothesis refinement' interfaces which automatically retrieve, process, assimilate and present relevant information in such a way that the user can see whether there is sufficient evidence to corroborate their hypothesis or if it needs refinement. Assuming the hypothesis refinement process does produce promising results, the next requirement is to be able to capture and store the hypothesis and its associated provenance data and body of evidence. This will enable future collaborative sharing, discussion and defense of new theories and help prevent duplication of analytical or experimental activities.

The research that we describe in this paper, focuses on the design, prototyping and evaluation of a data exploration and hypothesis-driven search interface called FUSION, that supports indexing and querying of complex semantic relationships and is driven by notions of information trust and provenance and the interactive investigation, development and capture of hypotheses. Although we have developed FUSION for a particular eScience application, the optimization of fuel cells by fuel cell experts, the research described here is applicable across any domains (e.g., science, engineering, homeland security, social sciences and health) that are attempting to solve complex problems through the analysis and assimilation of large-scale, mixed information and data sets.

The remainder of this paper is structured as follows. The next section describes related work, the background and objectives. Section 3 describes the architectural design of the system and the motivation for design decisions that were made. Sections 4 and 5 describe the interactive data exploration and hypothesis generation interfaces respectively. Section 6 describes the results of evaluating the system on real fuel cell data and images. Section 7 contains concluding remarks and plans for future work.

2 Related Work and Objectives

Hypothesis formulation involves finding local interrelations (hypotheses) among attributes within large databases of high dimensionality [3]. Since finding all possible interrelations is an infeasible task for many such databases, current research is concerned with the problem of finding potentially promising hypotheses, which can be further verified. This problem is tackled by a number of technologies including statistical data mining [4], data visualization, clustering and image processing of visualized data. In this paper we focus on a novel interactive visualization approach.

Visualization of large data sets is not new – a large amount of research has been undertaken on the application of visualization to data mining and knowledge discovery. Visualization can provide a qualitative overview of large and complex datasets, summarize data, find patterns, correlations, clusters or exceptions in data sets and greatly assist with exploratory data analysis. A comprehensive overview of data visualization techniques can be found in [5]. These approaches mainly apply to purely numerical data, do not support heterogeneous data and mixed-media objects (e.g., images, audio, video, text) and don't employ Semantic Web technologies to infer or visualize semantic associations.

Systems like Flamenco [6], Topia [7], CS AKTive Space [8] are examples of early efforts at blending specific information exploration goals with well-associated contextual information. Information from multiple heterogeneous sources are combined and presented to provide an integrated view of a multi-dimensional information space. Polyarchy visualizations [9] and mSpaces [10] are two formalisms recently employed to visualize semantic relationships between multiple information objects. Other examples include work by researchers at DSTC [11] and CWI [12] who have been working on automatic generation of multimedia presentations based on the semantic relationships between mixed-media information objects. The common objective of all of these systems is to provide interactive browsers which present semantically-associated information visually. The methods for visualizing relationships between database attributes or information objects varies depending on the nature (i.e., types, formats, size, granularity, dimensionality and subject) of the data and information objects. These information objects may be spatial, temporal, spectral, visual, audio, textual, 3D, numerical, arrays, matrices, web pages or scholarly publications. Graphs (2D and 3D), animations, virtual reality, hypermedia, map interfaces and combinations of these, have all been employed to visually represent knowledge bases or information structures.

The objectives of our work are to enable scientists to solve particular scientific or engineering problems by presenting the relevant data in an integrated, synchronized and coherent way that facilitates the discovery of new relationships or patterns that would not be possible through traditional search interfaces. More specifically we wanted to develop a system that combines visualization techniques with semantic inferencing and applies them to both multimedia information and multi-dimensional data. Consequently our objectives were to:

- Provide a search, browse and data exploration interface which allows users to interactively formulate hypotheses.
- Enable users to define their own mappings from semantic relationships between objects and data to preferred spatio-temporal presentation modes.
- Provide a hypothesis testing interface which allows users to quickly and easily specify their hypotheses, see whether there was any evidence to support this theory and modify or refine it based on the visual/graphical feedback
- Determine standardized methods for defining, recording and exchanging hypotheses (e.g., RuleML) and their associated, corroborative, evidential and provenance data which have been aggregated within a multimedia object (e.g., SMIL + 3D).
- Enable storage, search and retrieval of past hypotheses. This captures domain expert knowledge, enables its re-use and refinement, reduces duplication and provide evidence and provenance for experimental results.
- Enable annotations of stored presentations – particularly those that reveal new or interesting trends. Semantic annotations (based on domain-specific ontologies) of presentations enable their retrieval and re-use for further knowledge mining.
- Test and evaluate the system within the context of a particular eScience application. In our case, we have chosen ‘fuel cell optimization’ because it is a typical scientific problem involving a large number of variables and data types.

2.1 eScience Example Scenario

Fuel cells offer an alternative, clean, reliable source of energy for residential use, transport and remote communities. Their efficiency is dependent on the internal structure of the fuel cell layers and the interfaces between them. Electron microscopy generates images of cross-sectional samples through fuel-cell components that reveal complex multi-level information. Simple macro-level information such as the thickness of the cell layers, surface area, roughness and densities can be used to determine gas permeation of the electrode materials. Nano-level information about the electrode's internal interface structure provides data on the efficiency of exchange reactions. Figure 1 illustrates the range of image data obtainable.

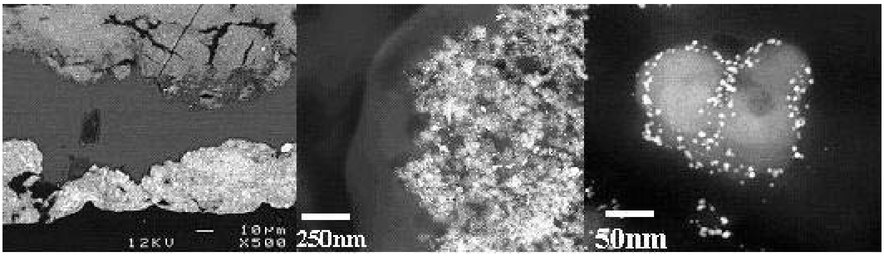

Fig. 1. Microscopic images of a fuel cell at 3 different magnifications

By digitising the images and applying image processing techniques (MATLAB) to them, the amount of information expands even further to levels where human processing is not possible and more sophisticated means of data mining are required. In addition to the microstructural information revealed by the images, there are the manufacturing conditions and processing parameters used to produce the cell configurations. Finally, for each cell configuration, performance data is available and the crux of the project is to marry the microstructural data with manufacturing and performance data to reveal trends or relationships which could lead to improvements in fuel cell design and efficiency. Table 1 shows the range of parameters we are dealing with, in addition to the fuel cell images captured at different magnifications.

Table 1. Fuel Cell Parameters

Fuel cell characteristics	**Performance**	**Manufacturing**
Layer thickness	Strength	Wt% Y2O3 - ZrO2
Composition	Density	Wt% Al2O3
Density	Conductivity graph	Wt% Solvent
Particle Size and Shape	Efficiency	Solid Content
Nearest neighbors	Lifetime	Viscosity
Surface area		Tape Speed and Thickness
Porosity		Drying Temperature and Time
Surface roughness		Cost

The aim of the work described here was to build and test an interactive interface which will enable fuel cell experts to quickly and easily explore the fuel cell images and data in order to determine associations or patterns between parameters, formulate and validate hypotheses, and save hypotheses and associated corroboratory evidence to share with others and keep as a historical record that tracks past investigations.

3 System Architecture

Figure 2 illustrates the overall architecture and major components of the FUSION system.

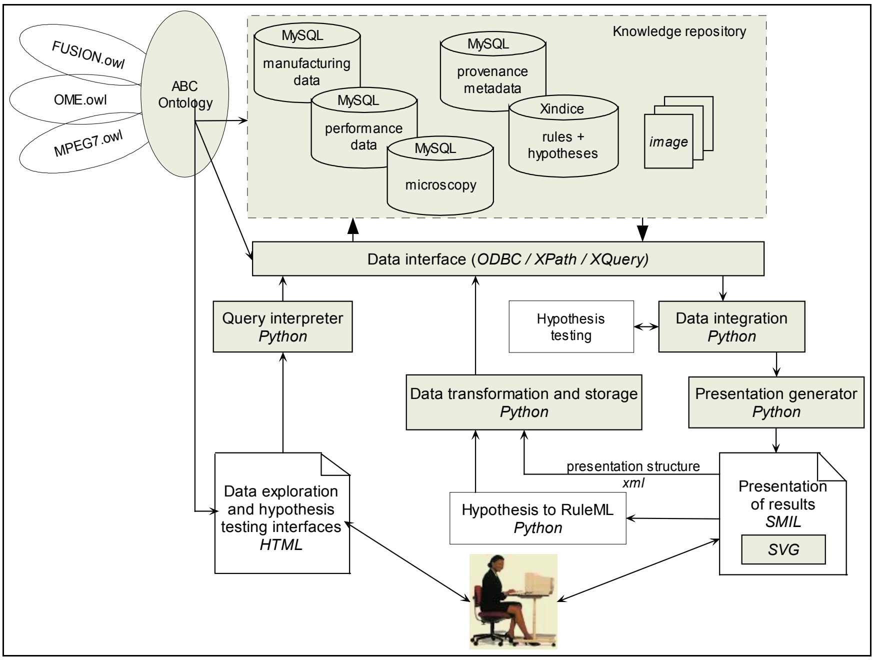

Fig. 2. System Architecture

Users access multiple distributed repositories through a Web browser (Microsoft Internet Explorer) and an ODBC interface – no specialized software is required on the client side except IE 5.5 or higher and an SVG plug-in. The Microsoft implementation of SMIL, HTML+TIME [13], is used to build the multimedia presentations. It allows spatio-temporal relationships between information objects as well as visual effects (such as fading between images) to be implemented. Dependencies between values are represented graphically using SVG [14]. Presentations and graphs are dynamically generated using Python scripts. The HTML forms and pull-down menus presented to the user are generated from domain-specific (OWL[15]) ontologies, described in earlier work [16] and which are specified during system configuration. The architecture is extremely flexible and can quickly and easily be adapted to any domain by connecting to different backend ontologies and knowledge repositories.

The two main system components, that are described in the next two sections, are:

1. Data Exploration;
2. Hypothesis Formulation.

3.1 Data Exploration

The data exploration process consists of four stages, as shown in Figure 3.

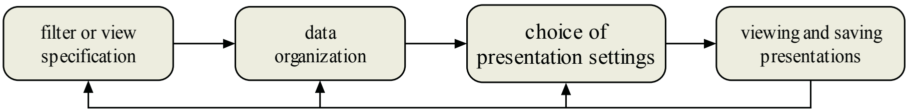

Fig. 3. Four stages of the data exploration process

Initially, the user chooses the aspects of the fuel cell data which they are interested in by selecting the parameters from the data set to be viewed. For example, *porosity*, *efficiency* and *cost*. The Query Interpreter (Figure 2) transforms the user's selection from the HTML form into a format that the Data Interface can process. The Data Interface refers to the metadata structure (harmonized ontologies) to determine which knowledge repositories should be queried. In this example, the performance and manufacturing data repositories are queried. The retrieved results include the unique IDs for the fuel cells matching the query. A request is sent to the image database to retrieve images of the fuel cells matching these IDs. The results of the search are submitted to the HTML interface for the execution of stages 2 and 3.

The HTML interface allows users to specify his/her preferences for displaying the retrieved results. Users can specify the following display preferences:

- an ordering parameter for structuring and presenting the results;
- selection of any additional parameters to be displayed;
- the type of presentation mode required (time-based or static);
- preferred data presentation formats (values displayed graphically, or in a list);
- any special effects to be applied to the presentation (e.g., fading etc.).

Figure 4 shows the user interface in which the user has specified that they wish to "order retrieved fuel cell data by increasing efficiency" and "also display values for porosity and cost". These additional specifications are submitted to the Query Interpreter, which reformulates the request to the Data Interface ("for previously retrieved fuel cell IDs retrieve corresponding values for efficiency, porosity and cost"). The results of the reformulated query are processed by the Data Interpreter, which transforms them into the necessary format for the Presentation Generator. The Presentation Generator makes decisions about the spatio-temporal layout of the result sets based on the users' preferred presentation mode, format and any special effects.

Figure 5 shows the user interface for specifying presentation preferences. There are three possible presentation modes to choose from: a time-based mode (slide-show) or one of two possible static modes – interactive and thumbnail views. When a large number of images are retrieved as a result of a search, the slide-show mode will display the images automatically without the need to click the next button. The default speed for a slide-show is 2 seconds per screen, but this can be adjusted. Alternatively, a static mode (interactive or thumbnail tiled view) can be selected to allow viewing of the results without any time restrictions. The interactive mode requires the users to press the next button to move to the next fuel cell image.

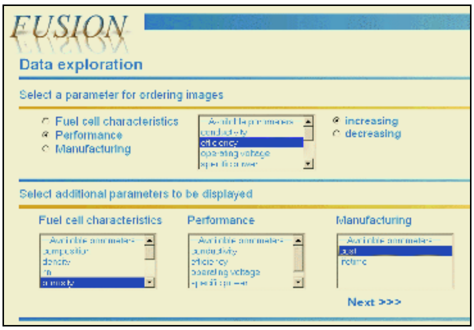

Fig. 4. Data Organization

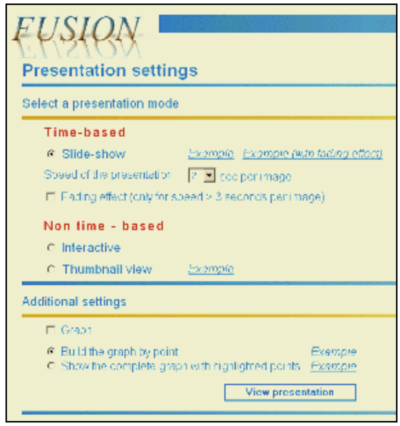

Fig. 5. Presentation Settings

The slide-show mode also displays an animated graph together with the images and any additional parameter values chosen by the user. The SVG graph plots one or more parameter values against time, and is generated dynamically in synchronization with the fuel cell images. This enables users to relate visual features in the images to manufacturing and/or performance data. In addition, fading effects can be applied to the images in the slide-show. This is helpful for distinguishing differences across sequential images.

The thumbnail presentation mode lays out thumbnail images for all of the retrieved fuel cells in a tiled structure ordered by the chosen parameter. In addition any requested parameter values are displayed below each corresponding thumbnail. Users can click on a thumbnail image to view it in full-size with all requested parameters and values listed below it.

Figure 6 illustrates the results of a slide-show presentation – the final stage of the data exploration process. The data exploration interface has been designed to enable a user to interactively explore large mixed-media, multidimensional data and information sets. By enabling users to choose the presentation style which best suits them, and to focus on the range and scope of data sets that most interest them, the system maximizes the potential and speed at which domain experts can discover new interdependencies or trends within the data or develop hypotheses. If the user finds an interesting pattern or association they can save the HTML+TIME+SVG presentation, together with the associated metadata (Unique ID, Date/Time, Creator, Settings, Objective) or move on to the next stage, the hypothesis testing interface, which is described in the next section.

All screenshots in this paper are also available at [17].

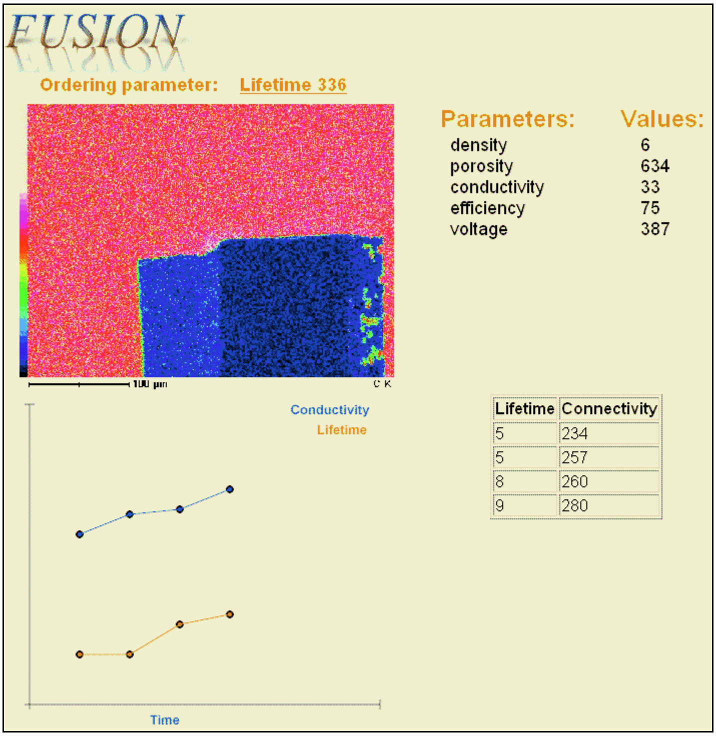

Fig. 6. Screenshot of a slide-show presentation with animated graphs

3.2 Hypothesis Testing

The design of the data exploration interface was based on a user-needs analysis and user feedback, as well as certain assumptions regarding the usefulness of particular features and presentation modes for displaying large amounts of heterogeneous data and information. The design of the hypothesis testing interface, however, was based on an analysis of the cognitive process of hypothesis testing and scientific discovery. Hypothesis formulation does not occur spontaneously but is an interactive, evolutionary process which grows out of background experience and assumptions which lead to ideas about relationships within the data, which the scientist wants to verify. Research into the process of conducting tests and experiments has shown that the hypothesis formulation workflow depends on whether results are expected or unexpected [18, 19]. If results conform to a particular hypothesis, then the work continues forward with the verification of further hypotheses. If an unexpected result occurs, further testing is done in order to explain the result. An unexpected result may occur due to an erroneous primary assumption or methodological errors. Otherwise, unexpected results may lead to a new discovery. Taking this into account, we developed an interface that enables users to specify their hypotheses, define prerequisites for validating and testing hypotheses and attach explanations to the results which are obtained. The workflow for the hypothesis testing process is shown in Figure 7.

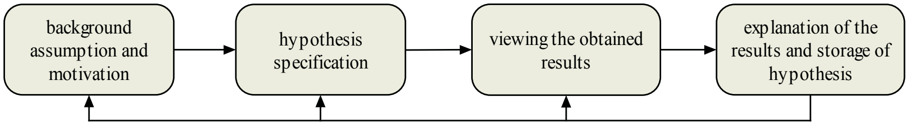

Fig. 7. Four stages of the hypothesis testing process

Consider the following example. The user wants to test the hypothesis: “IF substrate width is greater than 12µm AND density value lies between 5–10 particles/µm^2 THEN efficiency is greater than or equals 80%”. A HTML interface generated from the back-end ontologies enables the user to specify such a hypothesis. Figure 8 illustrates the user interface which supports the first and the second phases in Figure 7. Figure 8 consists of the following sections:

- Descriptive Metadata – for each hypothesis, the system generates the following metadata: Unique ID, Date/time and the Creator/Researcher ID.
- A free text field that enables researchers to record the background motivation for this particular hypothesis.
- A searchable list of all previously conducted experiments. Documenting past work enables sharing of earlier hypotheses, the layering of new hypotheses and helps reveal conflicts between hypotheses and prevent duplication of research.
- The bottom left section provides the interface for entering a hypothesis. It is divided into two parts: the conditional part of a hypothesis (if statement) and the consequence part of a hypothesis (then statement). Multiple sub-statements can be combined within the if or then statements using logical connectors (AND/OR). The bottom right part of the window contains dynamically filled lists of then statements that match the specified if statement on the left hand side. This mechanism helps to indicate dynamically, whether a hypothesis with the same if statement has previously been tested and what the outcome of this investigation was. If a matching previously-tested hypothesis is found, then the user can click on the then statement and retrieve a complete record of the results of that investigation.

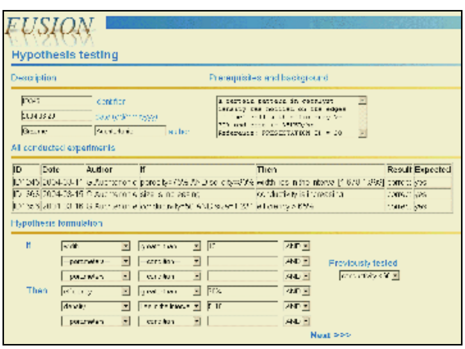

Fig. 8. Hypothesis specification

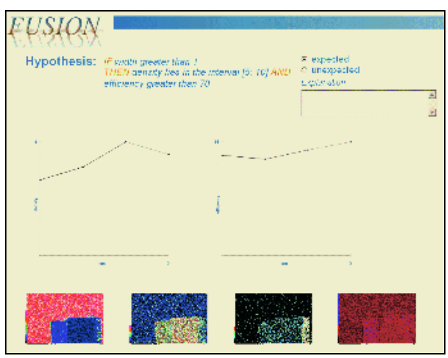

Fig. 9. Results of Hypothesis testing

The process of translating a hypothesis into a set of queries and retrieving the relevant data is identical to the process described in Section 3.1. The values/ranges that are retrieved for the specified parameters are sent to the Hypothesis Testing component, which attempts to verify or refute the hypothesis. The verification results together with the hypothesis itself, are passed through the Data Interpreter and Presentation Generator to produce a presentation. Dependencies between specified parameters are displayed within dynamically-generated SVG graph(s). The hypothesis itself is transformed into RuleML [20] format that users can choose to save to a Xindice [21] repository of stored hypotheses.

Figure 9 illustrates a presentation that was generated following the specification and submission of a hypothesis to the knowledge repository. The hypothesis statement is displayed at the top of the screen. The set of graphs displayed beneath the hypothesis statement, depict the dependencies between parameters specified in the hypothesis and provide feedback to the user on whether or not there is any evidence to support their hypothesis. For the example hypothesis given above, two graphs are generated. One plots density against substrate width. The other plots efficiency against substrate width. Users are able to either save this hypothesis (with an explanatory annotation) or go back and make changes to the original hypothesis and resubmit this to the knowledge base. A complete record of the saved experiment/hypothesis consists of:

- The metadata for the hypothesis: unique ID, date, and author;
- The motivation or background for the hypothesis;
- The hypothesis itself in the form of an if-then statement;
- The results of applying the hypothesis to the knowledge repository – an HTML+TIME+SVG presentation;
- An outcome attribute specifying whether the results were positive or negative;
- An annotation field, entered by the user, which contains a possible explanation for the results that were obtained.

An XML metadata record and the RuleML representation for the example hypothesis given at the start of this section can be found at [17].

4 Evaluation

User testing of the system has been carried out by fuel cell scientists from The University of Queensland's Centre for Microscopy and Microanalysis. Feedback from the users to date has indicated the following:

- The user interface design and incorporation of domain-specific ontologies [16] allowed users with little knowledge of the domain, to quickly and easily explore the data and gain an understanding of the knowledge space;
- Different users carry out research activities differently. Being able to customize or personalize the mode, scope and focus of the assimilated data presentations and the hypothesis refinement process was very beneficial for individual productivity;

- The different presentation modes enabled faster processing and interpretation of large data sets and images by the fuel cell scientists than was possible manually and expedited the hypothesis generation and refinement process;
- Slide shows of images synchronized with animated graphs that plot corresponding requested parameter values, were the most popular method of data exploration and hypothesis formulation;
- The fading effect was useful for detecting subtle image differences;
- Static presentations and graphs can be incorporated directly into scholarly publications, reducing the time required to disseminate research results;
- Being able to record, browse and retrieve past investigations and hypotheses, reduced duplication and enabled existing hypotheses to be refined or new hypotheses to be developed based on past work. It also provides a way of capturing and sharing tacit domain expert knowledge, explicitly, in the form of rules;
- Existing automatic hypothesis testing techniques (e.g., statistical analysis) only work on quantitative data. A major advantage of The FUSION system's approach is that it applicable across a range of data and media types.
- The saving of evidential and provenance data with hypotheses, enables the validity of earlier hypotheses or assumptions to be assessed by other scientists – who are able to attach their own opinions in the form of annotations;
- The use of semantic web technologies such as ontologies, annotations and inferencing rules, provide a consistent, machine-processable way for describing, capturing, re-using and building on the domain knowledge. It also enables better collaboration between distributed research laboratories and industry through improved sharing of knowledge and data.

An on-line demonstration of the prototype system is available at [22]. Users need to be using IE 5.5 or higher and an SVG plug-in, such as Adobe's plugin [23].

5 Conclusions and Future Work

In this paper we describe a search interface that enables scientists to interact with a knowledge base through a hypothesis-driven approach that combines data exploration, integration, search and inferencing – enabling more complex analysis and deeper insight. We believe that such interfaces represent the next generation of search engines and that they are will be increasingly in demand and applied across many domains including science, engineering, homeland security, social sciences and health, to solve complex problems and provide decision support tools based on the analysis and assimilation of large-scale, mixed-media, multi-dimensional information and data sets.

Plans for future work include:

- Further testing and refinement of the system, particularly within a real-world industrial environment. We plan to deploy it within a fuel-cell manufacturing company to facilitate the exchange of knowledge between university research and industry organizations in this domain;

- Integrating statistical data analysis methods and applying them to hypotheses formulated through our system, to fit more precise mathematical models to relationships between parameters;
- Investigating how the empirical modeling approach described here can be combined with the physical modeling approach to generate a more accurate predictive model for simulating fuel cell behaviour.
- Testing the portability, flexibility and scalability of the system by applying it to other domains, such as environmental modeling and bioinformatics.

Looking even further into the future, we envisage that instead of users interactively submitting hypotheses to such a system, there will be pro-active systems which are: constantly dynamically assimilating new information; using existing, stored hypotheses to automatically detect anomalies, problems, or exceptional events; inferring new hypotheses and knowledge; and notifying users by returning actionable information. But we still have a long way to go before such intelligent or sophisticated systems become widely available.

Acknowledgements

The work described in this paper has been funded by the Cooperative Research Centre for Enterprise Distributed Systems Technology (DSTC) (through the Australian Government's CRC Programme) and through a University of Queensland Development Grant. Thanks to CWI for supporting Katya's secondment to DSTC and to John Drennan and the University of Queensland's Centre for Microscopy and Microanalysis for providing the fuel cell images and data and for their valuable user feedback.

References

1. Google, (http://www.google.com).
2. Hendler, J., Berners-Lee, T., and Miller, E., Integrating Applications on the Semantic Web. Journal of the Institute of Electrical Engineers of Japan. Vol 122(10): p. p. 676-680., 2002.
3. Amir, A., Kashi, R., and Netanyahu, N. S. "Efficient Multidimensional Quantitative Hypotheses Generation." in Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM). Melbourne, Florida, USA, November 19-22, 2003.
4. Smyth, P., Data Mining at the Interface of Computer Science and Statistics, Invited Chapter in Data Mining for Scientific and Engineering Applications. 2001, Kluwer. p. 35-61.
5. Fayyad, U., Grinstein, G. G., and Wierse, A., Information Visualization in Data Mining and Knowledge Discovery. 2001: Morgan Kaufmann.
6. Hearst, M., et al., Finding the Flow in Web Site Search. Communications of the ACM. 45(9): p. pp.42-49., 2002.
7. Topia, (http://topia.demo.telin.nl/).
8. Shadbolt, N. R., et al., CS AKTive Space or how we stopped worrying and learned to love the Semantic Web. IEEE Intelligent Systems. 2004.
9. Robertson, G., et al. "Polyarchy visualization: visualizing multiple intersecting hierarchies", in Conference on Human Factors in Computing Systems. Minneapolis, Minnesota, USA.

10. Gibbins, N., Harris, S., and Schraefel, M. “Applying mSpace Interfaces to the Semantic Web”, in Submitted to Proceedings of World Wide Web Conference 2004. New York, USA.
11. Little, S., Geurts, J., and Hunter, J. “The Dynamic Generation of Intelligent Multimedia Presentations through Semantic Inferencing”, in ECDL2002. Rome, Italy, September 2002.
12. Geurts, J., et al. "Towards Ontology-Driven Discourse: From Semantic Graphs to Multimedia Presentations", in International Semantic Web Conference (ISWC2003). Sanibel Island, Florida, USA, October 2003.
13. Microsoft, HTML+TIME 2.0 Reference, (http://msdn.microsoft.com/workshop/author/behaviors/reference/time2_entry.asp).
14. W3C, Scalable Vector Graphics (SVG) 1.1 Specification, W3C Recommendation, 14 January 2003, Edited by Jon Ferraiolo and Fujisawa Jun and Dean Jackson, (http://www.w3.org/TR/SVG/).
15. W3C, OWL Web Ontology Language Reference, W3C Candidate Recommendation, 18 Aug 2003, Edited by Mike Dean and Guus Schreiber, (http://www.w3.org/TR/owl-ref/).
16. Hunter, J., Drennan, J., and Little, S., Realizing the Hydrogen Economy through Semantic Web Technologies. IEEE Intelligent Systems Journal - Special Issue on eScience. 2004.
17. Falkovych, K., Little, S., and Hunter, J., Appendices and Examples Screenshots, (http://metadata.net/sunago/fusion/visualisation/ecdl2004.html).
18. Dunbar, K., How Scientists Build Models InVivo Science as a Window on the Science Mind, in Model-Based Reasoning in Scientific Discovery, L. Magnani, N.J. Nersessian, and P. Thagard, Editors. 1999, Kluwer Academic/Plenum Publishers: New York. p. 85-99.
19. Okada, T. and Simon, H. A. "Collaborative Discovery in a Scientific Domain", in Proceedings of the 17th Annual Conference of the Cognitive Science Society, 1995.
20. Boley, H., Tabet, S., and Wagner, G. "Design Rationale of RuleML: A Markup Language for Semantic Web Rules", in Semantic Web Working Symposium (SWWS), 2001.
21. Apache Software Foundation, Apache Xindice, (http://xml.apache.org/xindice/).
22. Falkovych, K., Little, S., and Hunter, J., Scientific Data Exploration and Hypothesis Testing OnLine Demo, 2004, (http://metadata.net/sunago/fusion/visualisation/intro.html).
23. Adobe, Adobe SVG plugin, (http://www.adobe.com/svg/).