# Chapter 14
# Deploying and Evaluating Semantic Technologies in a Digital Library

**Ian Thurlow and Paul Warren**

**Abstract** Digital libraries have been the subject of considerable research since the 1990s. Their practical value in providing remote access to knowledge is beyond question. As they have developed, there have been numerous attempts to theorize about their future nature and role, and to lay down the challenges for further work. On a more practical level, individual organisations have developed their own digital libraries in response to their particular needs. This chapter describes research into the use of semantic technology in a Digital Library. The work draws on the technologies and tools described elsewhere in the book and puts them in the context of a particular application. The chapter also explains in some detail how user trials within the case study were used to validate our approach.

## 14.1 Introduction

Digital libraries have been the subject of considerable research since the 1990s. Their practical value in providing remote access to knowledge is beyond question. As they have developed, there have been numerous attempts to theorize about their future nature and role, and to lay down the challenges for further work. On a more practical level, individual organisations have developed their own digital libraries in response to their particular needs. One such is the BT Digital Library, which has evolved over the last dozen years in response to the needs of knowledge workers in BT. This chapter describes research into the use of semantic technology in the BT Digital Library. The lessons, however, are applicable to digital libraries in general. The work draws on the technologies and tools described elsewhere in this book and puts them in the context of a particular application. The chapter also explains in some detail how user trials within the case study were used to validate our approach.

I. Thurlow (✉)
Next Generation Web Research BT Research, Ipswich IP5 3RE, UK
e-mail: ian.thurlow@bt.com

J. Davies et al. (eds.), *Semantic Knowledge Management*,

## 14.2 The BT Digital Library

The BT Digital Library,[1] which formed the starting point for our case study, contains abstracts, and in many cases full text, of much of the literature of interest to BT's technical specialists, besides a significant amount of relevant non-technical literature. In all, this represents five million articles from over 1,200 publications, including journals, conference proceedings and IEEE Standards. This is provided in the form of two databases, Inspec[2] and ABI/INFORM from ProQuest.[3] Users search the content of the library using a keyword-based search engine. The search engine offers high recall and, although it is possible to narrow down the number of results returned through careful composition of the query, most users only formulate very simple queries; as a result being presented with a large number of search results. This was the starting point on which we hoped to build and improve through the use of semantic technologies.

### *14.2.1 The Challenges*

In a previous book (Davies et al. 2006) the authors have described the challenges facing digital libraries. Most prominent amongst these is the goal of semantic interoperability (Chen 1999). In practical terms this means providing each user with a unified view of digital objects across libraries. Other challenges included the need for improved user interfaces to navigate large information collections (Lynch and Garcia-Molina 1995) and the need to match concepts, not just search strings (NSF 2003).

For us, semantic interoperability meant the ability to integrate relevant information on the Web or corporate Intranet with information in the Digital Library.

Improving the user interface and searching on concepts were particularly prominent challenges for us, as they gave us the opportunity to address some of the weaknesses of the search tools in use in the library. Our response to these two challenges is represented by the user tools *Squirrel*, a semantically-enabled search and browse tool, and *SEKTagent*, a semantically-enabled search agent application. These tools are described elsewhere in this book and in Duke et al. (2006). Their use in the Digital Library is summarised later in this chapter.

A long-standing goal of the BT Digital Library was to be not just a repository of information, but also an enabler of communities of interest, where knowledge could be exchanged and people of similar professional interests could make contact. The BT Digital Library was not alone in this. For example, the Perseus Digital Library

[1] The BT Digital Library has been developed under the leadership of the authors' colleague, David Alsmeyer

[2] http://www.iee.org/Publish/INSPEC/

[3] http://www.proquest.com

project[4] has a commitment to "connect more people through the connection of ideas". Early in the development of the Library, the concept of information spaces was introduced (Alsmeyer and Owston 1998). On one level each information space is a collection of documents, defined by a particular query agent. On another level, the information space is the group of people who register to receive regular alerts from the agent. A companion chapter of this book describes the use of semantic technology for knowledge sharing, in the form of the *Squidz* tool, and compares this with the pre-existing information spaces.

### *14.2.2 Understanding User Needs*

Before beginning development we sought the views of users to validate our ideas. This is described in some detail in Davies et al. (2006), but a brief overview is given in this section.

Initially a questionnaire was used to obtain a general view of user requirements. There was then a focus group drawn from a small number of the Library users. Again, the questions posed to the group were quite general. We learned that users wanted improved ways of searching; and that they wanted searches to take into account their profile of interest.

A second questionnaire was used to test users' reactions to our proposed functionality. This confirmed the value of the planned functionality and gave us an understanding of the users' perceived priorities. The majority of the features valued by users were subsequently incorporated into *Squirrel*, a semantic search and browse application.

A final technique for understanding users' requirements was that of user preference analysis.[5] We used this to understand the trade-off, from the user's perspective, between precision and recall. Semantic technology can help us improve both. Precision can be improved by the capability to specify more precisely the entity sought, for example, that the string "BT" represents a company in the telecommunications sector. Recall is enhanced by the ability to recognise different words or phrases as synonyms. When the results of our preference analysis were analysed there were two clusters of users. In one cluster users had a clear preference for precision over recall. In the other cluster users gave equal weight to precision and recall. This led, in our user trials, to the adoption of a different measure of information quality, specifically oriented to the end-user perspective.

In all, the combination of questionnaires, focus group and user preference analysis provided an understanding of user requirements to guide the development of the end-user tools.

---

[4] http://www.perseus.tufts.edu

[5] The user preference analysis was undertaken with the guidance and assistance of Professor Tom Bösser, of kea-pro GmbH.

## 14.3 Squirrel – A Semantic Search and Browse Application

*Squirrel* was designed to improve access to digital library content, and to overcome some of the limitations presented by the search engine that is currently in use in BT's digital library. *Squirrel* combines free-text search with the means to query and browse ontology-based semantic meta-data in the form of a Web Ontology Language (OWL)[6] shared ontology. A collection of bibliographic records from Inspec and ABI is integrated with information sourced from the Web using a common topic ontology.

The user interface to *Squirrel* is shown in Fig. 14.1. A complete description of the main functions are described in a companion chapter in this volume. Users enter their

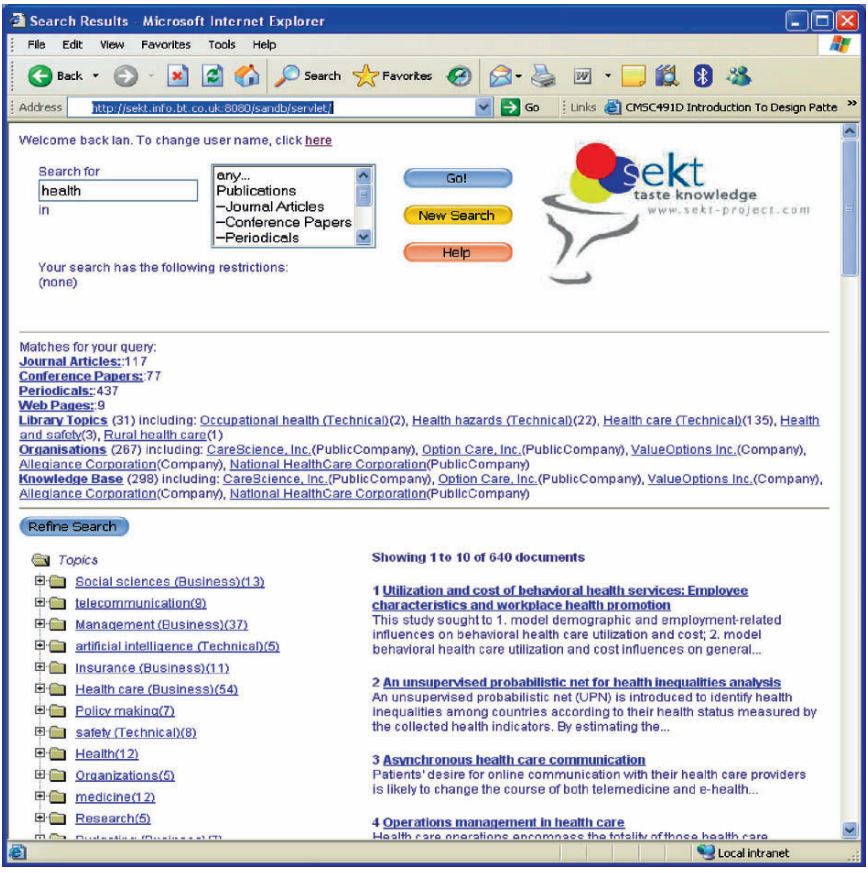



**Fig. 14.1** The user interface to Squirrel

[6] http://www.w3.org/2004/OWL

search query in the search box, and optionally select the type of information they need from the drop-down menu, for example, users can search for publications authored by a particular person, for publications of a particular type (e.g., a conference paper), or for items which mention a specific company or one of its aliases. Alternatively, for users wanting to invoke a broader search, the search application offers a search which covers all types of content held in the digital library. The user's query is formed using concepts from the ontology. Search results are ranked, taking into account user profiles to give a degree of context to the user's search (the user profile is constructed from an analysis of a user's interaction with BT's digital library and other Web information sources, and is described through a set of interests that are defined in the ontology). Users can refine their query by selecting from a set of related topics returned with the results of a search.

Squirrel identifies and highlights *named entities* within the search results, for example, the names of people, names of companies, names of organisations, location names, etc. Supplementary information concerning those entities is presented to the user, for example, for an entity of type *company*, additional information such as the location of its headquarters, its web address, the sector it operates in, and details of its key personnel are presented to the user. The rationale here is that the identification and presentation of named entities should give the end user information additional to that in the original text, and help the user to more swiftly identify the context of the results from a search, thereby enabling the results to be filtered based on entities of interest.

Users can refine the search by specifying the type of document (e.g., journal article, conference paper, periodical, web page), library topic, or name of company identified in the document. Details of any organisations matching the search criteria are also presented to the user. In addition to searching against the metadata that describes the information entities, the user can also browse the topic hierarchy to find relevant documents. Users are able to navigate up or down the topic hierarchy, expanding or refining their search as they go, for example, should a search on a particular topic give no useful result, users can make use of super topics of the original topic to broaden their search. Conversely, if there are too many results, users can make use of subtopics to narrow their search.

Besides searching and browsing by topic, the user can also browse the library using other characteristics of a document, for example, by requesting other documents by the same author, or papers published in the same conference proceedings. The refine search option enables a set of results to be filtered by searching specific fields of the records, for example, the title or the abstract. A user can also filter the results set by date or by author name.

## 14.4 SEKTagent – A Search Agent Application

*SEKTagent* collects relevant content from a pre-indexed set of documents on behalf of a user. This set includes the ABI and Inspec abstracts (and full-text where available), Web pages and RSS items. The application highlights named entities identified

within the search agent's results, and links those entities to relevant supplementary information in the knowledge base. The *SEKTagent* application uses an API provided by the KIM platform (Kiryakov et al. 2003). End users configure each agent with a semantic query making use of the BT digital library ontology. Some example *SEKTagent* queries, shown in natural language, are given below:

Find all intranet pages that contain a named person holding a particular position within a user-specified organisation.

- Find all conference papers authored by a particular person.
- Find all articles where the author is affiliated with a named organisation.
- Find all publications mentioning a named company that is active in a particular industry sector.

This mode of searching for types of entity is complemented with a full text search, allowing the user to specify terms that should occur in the text of the retrieved documents. The results page for a *SEKTagent* that is searching for any *person* that holds the *job position* of *analyst* within an *organisation* that is *located* in *any location* is shown in Fig. 14.2.

The title of the Web page, or title of the abstract/full-text for ABI/Inspec records, and a short summary of the content relevant to the query are displayed for each result. The summary highlights the occurrences of the named entities that satisfy the query. Other recognised named entities are also highlighted. In a similar way to the *Squirrel* search and browse application, the user is able to access further information about the entity from the knowledge base by placing their mouse pointer over any of the named entities that are highlighted.

## 14.5 Evaluation Process

The *Squirrel* search and browse and *SEKTagent* search agent applications were subjected to a three-stage user-centred evaluation process. The first two stages of the process, which comprised a *heuristic evaluation* (Nielsen and Molich 1990) and a *cognitive walkthrough* (Wharton et al. 1992), aimed to identify and correct any bugs or navigational problems with the applications. The final stage of the process comprised a set of *field tests* that aimed to evaluate the applications against the search engine currently in use in BT's digital library.

### *14.5.1 Heuristic Evaluation*

An informal heuristic evaluation of the user interface was undertaken, with the primary objective to assure that the functions of the applications worked as intended. The aim was to identify and correct all significant defects and user interaction problems in the early stages of development.

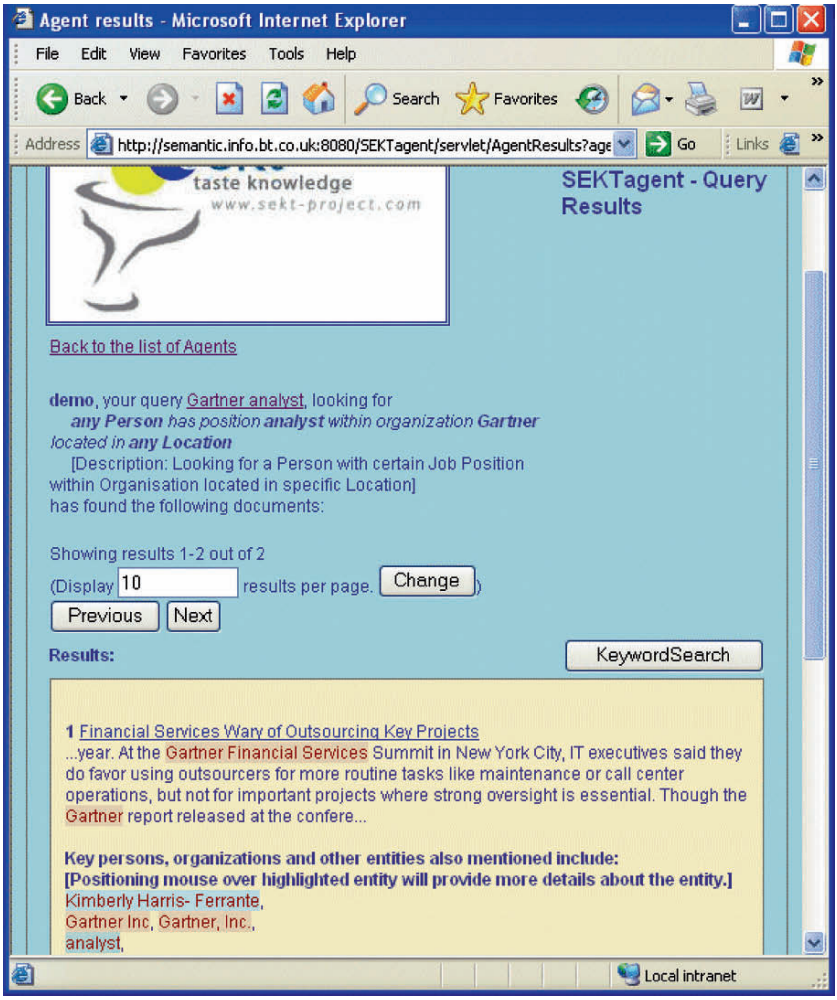



**Fig. 14.2** SEKTagent results page

A group of five researchers, acting as usability experts, judged whether the user interface of early prototypes of *Squirrel* and *SEKTagent* adhered to a list of established usability principles known as the *usability heuristics*. The usability heuristics were based on a checklist adapted from the Xerox heuristic evaluation system checklist.[7] Each expert inspected the prototype during a session lasting ~1 h. A number of observations were made during each session, the majority of which tended to be concerned with minor interface problems and system performance issues. An example is shown

[7] http://www.stcsig.org/usability/topics/articles/he-checklist.html

| No. | Review checklist | Yes | No | N/A | Comments |
|---|---|---|---|---|---|
| 2.1 | Are icons and action buttons concrete and familiar? | ✓ | | | Make sure that navigation is consistent within the application. |
| 2.2 | Are application choices ordered in the most logical way, given the user, the item names and the task variables? | ✓ | | | |
| 2.3 | If there is a natural sequence to application choices. Has it been used? | ✓ | | | |
| 2.4 | Do related and interdependent fields appear in the same window? | ✓ | | | |

**Fig. 14.3** Excerpt from system checklist

in Fig. 14.3. The findings of the inspection were collated and a measure of severity was assigned to each problem. The observations were discussed with the development team. The applications were then modified in accordance with the observations.

### *14.5.2 Cognitive Walkthrough*

The heuristic evaluation provided a simple and effective method for identifying many of the problems with the user interface. It did not, however, reveal all problems that a user may experience when using an application. The primary aim of the cognitive walkthrough was to find out whether each application functioned as expected under simulated operational conditions. A group of five users were asked to use the applications to undertake a set of short information seeking tasks. Each task was designed to make use of one or more of the key functions of the application. Users were encouraged to explain their actions and their concerns as they undertook each task, thus giving the assessors additional information about their thought processes and an early indication on the overall usability of the application. The findings of the evaluation were discussed with the development team. Each application was then modified to resolve all major usability issues that were identified during the evaluation.

### *14.5.3 Field T~ests (Squirrel and SEKTagent)*

The next part of the evaluation comprised a user test of *Squirrel* and *SEKTagent* applications against the current search engine in use in BT's digital library. Twenty users with a wide range of experience of using search tools were invited to take part in these field tests. In the case of *Squirrel*, we also assessed whether the information

search is more efficient in terms of information quality and time (information quality was assessed as a subjective measure of the relevance of the results to a user's information need). In essence, we wanted to find out whether the new technology would help people to get to the information they believed to be relevant to their search more easily and quickly.

A test environment, which comprised *Squirrel, SEKTagent*, and the current digital library search engine, was configured to give access to ~37,000 bibliographic records and 2,000 web documents. The field tests undertaken by each user comprised three distinct parts: (1) demonstration of the key search functions of *Squirrel, SEKTagent*, and the current search engine, followed by completion of a short questionnaire, (2) completion of information seeking tasks using *Squirrel* and the current search engine, and (3) completion of a Software Usability Measurement Inventory (SUMI) (Kirakowski and Corbett 1993) questionnaire for *Squirrel*.

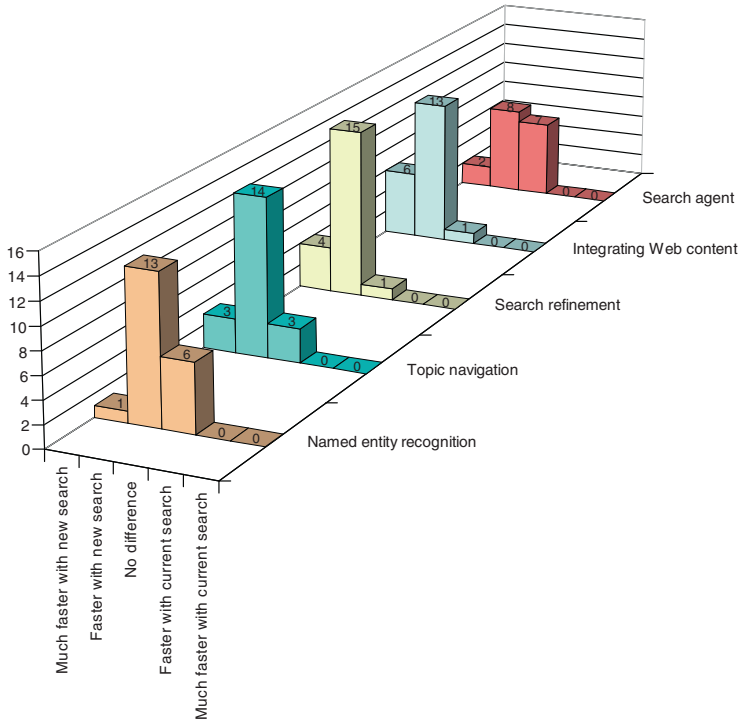
#### 14.5.3.1 Demonstration of New Search Functions

The first part of the evaluation aimed to give the subjects some time to familiarise themselves with the main functions of *Squirrel, SEKTagent* and the current search engine. It also gave us the opportunity to gather some initial user feedback on the new semantic search functions in comparison with current search functions. The following functions were demonstrated: (1) named entity recognition in *Squirrel*, (2) navigation and browsing using the topic ontology in *Squirrel* against use of controlled indexing terms for navigation and browsing using the current search technology, (3) search refinement in *Squirrel* and the current search engine, (4) integration of Web content in *Squirrel*, and (5) the *SEKTagent* semantic search agent functions against use of the current information spaces implementation.

After each function was demonstrated, and after the subject had spent some time using the functions of the applications to complete a simple search task, each subject was asked to provide some qualitative feedback with respect to that task. For each function, we asked the subjects to answer to the following questions (by rating on a 4 or 5 point scale[8]):

- Does the new function offer an improvement compared to the equivalent function in the current search engine?
- Does the information seeking task become easier to complete with the new search and browse application (*Squirrel* and *SEKTagent*) in comparison with the current search engine?
- Do they expect to find information more quickly with the new search and browse application in comparison with the current search engine?
- Do they expect to find better quality information with the new search or the current search technology (where the relevance of results was to be taken as the main measure of quality)?

[8] Refer to Figure 14.4 through to Figure 14.7 for scales.



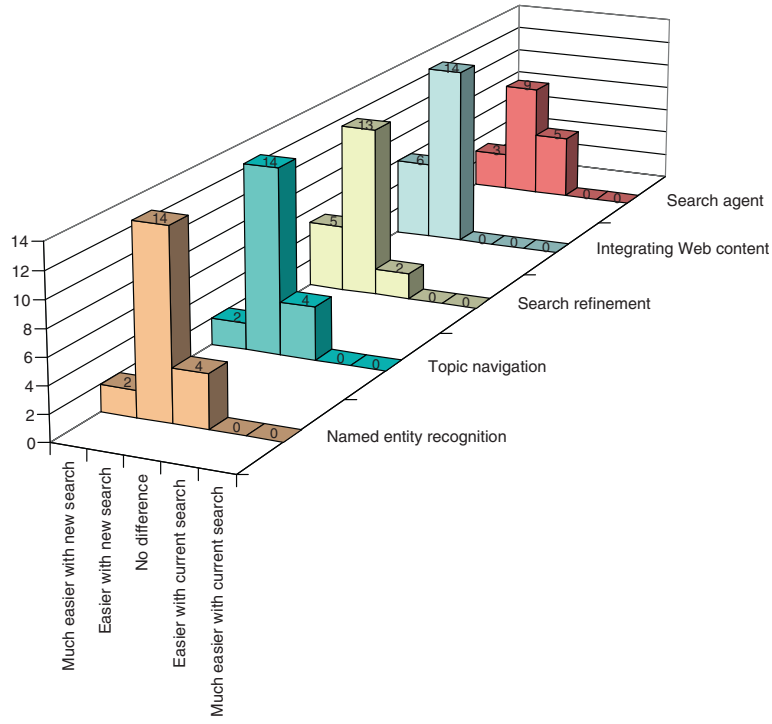

| | Much faster with new search | Faster with new search | No difference | Faster with current search | Much faster with current search |
|---|---|---|---|---|---|
| Named entity recognition | 1 | 13 | 6 | 0 | 0 |
| Topic navigation | 3 | 14 | 3 | 0 | 0 |
| Search refinement | 4 | 15 | 1 | 0 | 0 |
| Integrating Web content | 6 | 13 | 1 | 0 | 0 |
| Search agent | 2 | 8 | 7 | 0 | 0 |

**Fig. 14.4** Speed of access to information for new search functions

The results of the evaluation are summarised in Figs. 14.4–14.7. Overall, there was a very positive response to the new functions provided in *Squirrel* and *SEKTagent*. In particular, the improved functionality for searching web content stands out, being rated very positively.

### 14.5.3.2 Information Search Quality Measures

One of the main quality dimensions that users will use to assess the value of a search and browse application is the effectiveness of their search. The fundamental value to the user is the quality of the results returned by the search engine in relation to their information need, that is, the problem that motivated their search. The effectiveness of their search also depends on their ability to carry out the search task efficiently, that is, successfully and timely, and without undue cognitive effort when using the application.

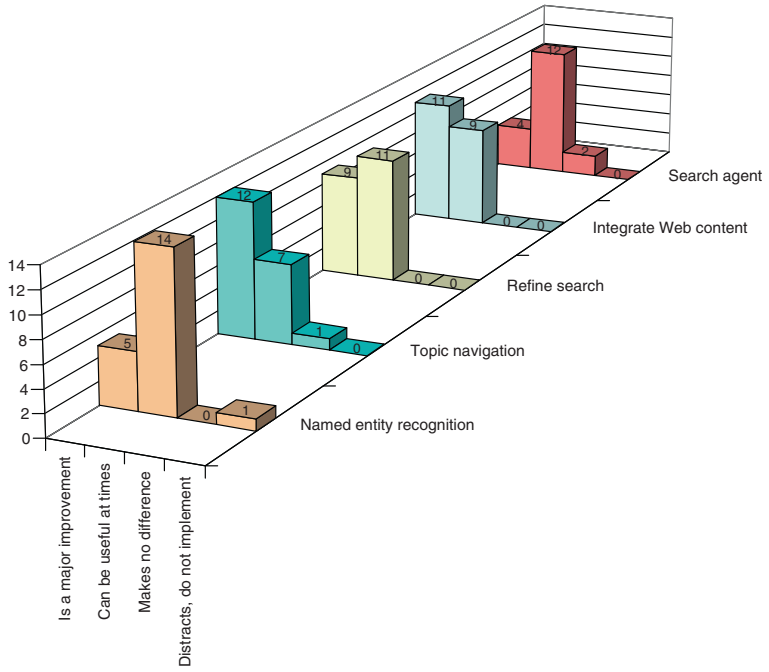



| | Much easier with new search | Easier with new search | No difference | Easier with current search | Much easier with current search |
|---|---|---|---|---|---|
| Named entity recognition | 2 | 14 | 4 | 0 | 0 |
| Topic navigation | 2 | 14 | 4 | 0 | 0 |
| Search refinement | 5 | 13 | 2 | 0 | 0 |
| Integrating Web content | 6 | 14 | 0 | 0 | 0 |
| Search agent | 3 | 9 | 5 | 0 | 0 |

**Fig. 14.5** Ease of access to relevant information

The main goal of users is to maximise the efficiency of their information seeking process, that is, maximise the quality of the information returned by the search engine, whilst minimising any personal cost to them in terms of time and the cognitive effort required to obtain a set of results that they are satisfied with. Ideally, the gain in the quality of the information returned from the search and browse application in terms of relevance to a user's information need should increase as the user's search progresses.

Conventional measures of information retrieval systems (van Rijsbergen 1980) tend to focus on search precision (the proportion of relevant results that are retrieved to all the results that are retrieved), search recall (the proportion of relevant results that are retrieved to all relevant results available), and the F-measure (the weighted harmonic mean of precision and recall). The relevance of a set of search results is, however, a highly subjective measure. Search results that are highly relevant to one person's query are very likely to differ in terms of relevancy to another user using the same search engine and submitting the same query. Moreover, it is anticipated



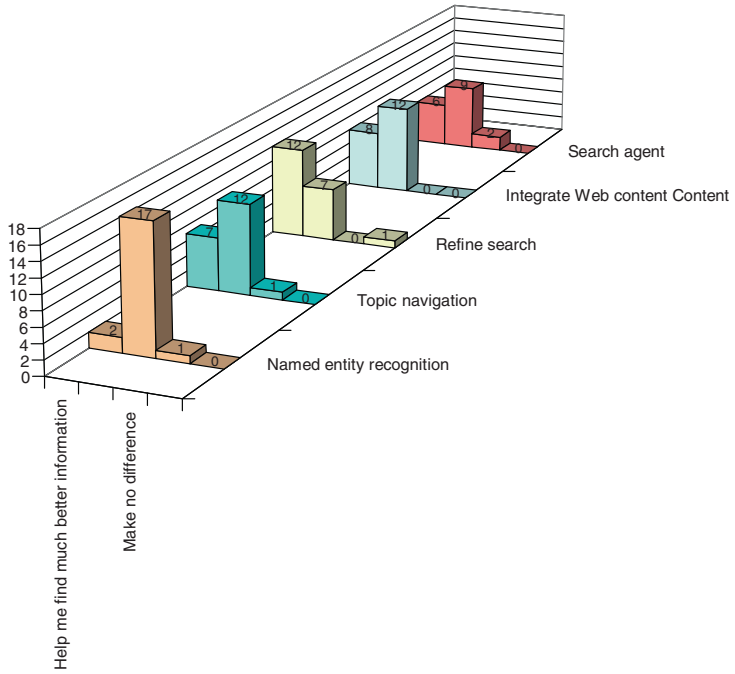

| | Is a major improvement | Can be useful at times | Makes no difference | Distracts, do not implement |
|---|---|---|---|---|
| Named entity recognition | 5 | 14 | 0 | 1 |
| Topic navigation | 12 | 7 | 1 | 0 |
| Refine search | 9 | 11 | 0 | 0 |
| Integrate Web content | 11 | 9 | 0 | 0 |
| Search agent | 4 | 12 | 2 | 0 |

**Fig. 14.6** Improvement in search engine functions

that there is a shift in a user's perception of the relevance of a search result depending on the amount of relevant information that is accessible. When there is an abundance of relevant information, a user's acceptance criterion in terms of the quality of the search results is expected to be high, whereas in the opposite case, where only a few relevant or some marginally relevant articles are found, the acceptance level of the user is likely to be reduced. We therefore decided to take a different approach to measuring the effectiveness of the *Squirrel* search and browse application, replacing the more conventional measures of precision and recall in favour of an approach that enabled us to assess the quality of information returned from the search engine from an end user's perspective, that is, we concentrated on the subjective assessment of the information quality in our analysis.

#### 14.5.3.3 Information Seeking Tasks

In order to gauge the effectiveness of *Squirrel*, we conducted a series of tests where users were asked to carry out defined tasks, with a defined objective, under controlled




| | Help me find much better information | Might help me find better information at times | Make no difference | Preclude me from finding some information |
|---|---|---|---|---|
| Named entity recognition | 2 | 17 | 1 | 0 |
| Topic navigation | 7 | 12 | 1 | 0 |
| Refine search | 12 | 7 | 0 | 1 |
| Integrate Web content Content | 8 | 12 | 0 | 0 |
| Search agent | 6 | 9 | 2 | 0 |

**Fig. 14.7** Information quality expectations

conditions. Subjective user assessments were collected from users. Each user was asked to complete two similar, but separate, information seeking tasks from a set of six. One task was completed using *Squirrel*, the other was completed using the search and browse application currently in use in BT's digital library. The order in which a subject used the applications was alternated, so that half of the tasks were performed using the keyword-based search engine first, and the other half using the semantically enabled search engine first.

The user was presented with rating scales and asked to assess the subjective value of the results returned from the application as they completed their search. Users provided feedback on two key dimensions: (1) the perceived information quality (PIQ), that is, the quality of results returned from the search engine in relation to the task as perceived by the user, and (2) the perceived progress of search (PPS), that is, the progress of their search in relation to the task they were asked to undertake. PIQ was measured according to a 7-point scale, where *1 = unsatisfactory, 2 = poor, 3 = less than expected, 4 = as expected, 5 = above average, 6 = good*

and *7* = *excellent*. PPS was measured according to a 15-point scale, where points 1–5 represented the beginning of the users search, points 6–10 represented revisions of their search, and 11–15 represented the refinement of their search as the task neared completion. Assessments of information quality were given by users at points in time that they determined themselves, for example, the moment before they submit a modified query rather than at a set time interval, so as to minimise interruptions to their search process.

Users typically provided four or five PIQ and PPS measures per task (average of 4.6). The average time for a user to undertake a search task was ~15 min (the minimum duration was 7 min; the maximum duration was 27 min). There was no significant difference in time between the two search tasks of users.

The average PIQ using the existing library system was 3.99 against an average PIQ of 4.47 using *Squirrel*. The PIQ scores given at various stages in the users' searches are shown in Fig. 14.8. The sign tests applied to the data shows the difference to be significant ($p < 0.01$). Note: in places where datapoints coincide (Fig. 14.8), one datapoint is shown inside the other.

#### 14.5.3.4 Usability of Squirrel

Finally, the Software Usability Measurement Inventory (SUMI),[9] which gives a more detailed view of the subjective assessment of the usability of *Squirrel*, was

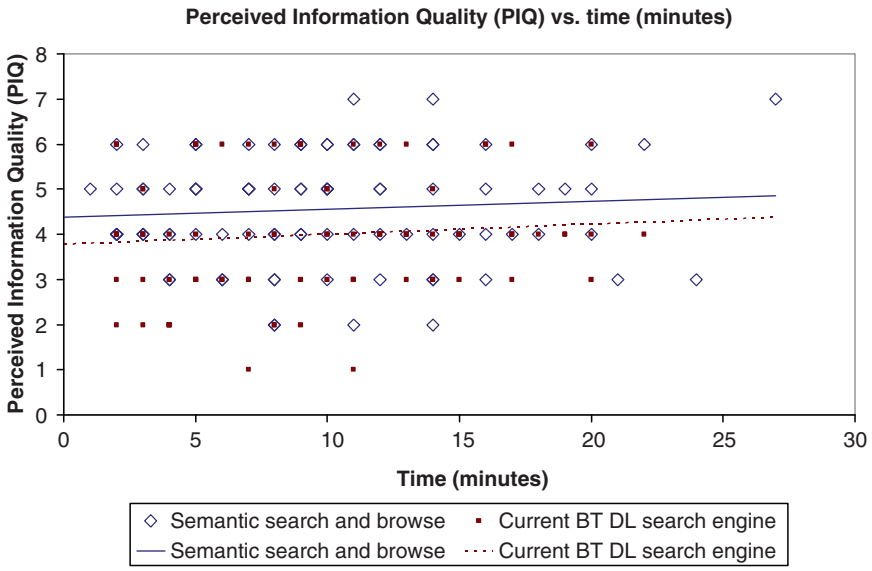



**Fig. 14.8** Perceived information quality (PIQ) against time

[9] http://sumi.ucc.ie

administered to each subject. SUMI enables the following five separate aspects of user satisfaction to be measured: (1) efficiency, that is, the user's feeling that the software enables them to perform their tasks in a quick, effective and economical manner (Table 14.1), (2) affect, which refers to the positive user feeling of the user being mentally stimulated and pleased as a result of interacting with the software (Table 14.2), (3) helpfulness, which refers to the user's perceptions that the software communicates in a helpful way and assists in the resolution of operational problems, (4) control, which refers to the feeling that the software responds in an expected and consistent way to input and commands (Table 14.3), and (5) learnability, which refers to the feeling that it is relatively straightforward to become familiar with the software. A *global* metric describes the user's view of the overall usability of the application.

The SUMI questionnaire comprises fifty statements, each labelled with boxes *Agree, Undecided*, and *Disagree*. Users were asked to complete SUMI (for *Squirrel*) after they had completed the information seeking tasks (users are asked to answer all questions). Of the twenty SUMI questionnaires, one questionnaire was clearly incomplete and was excluded from the analysis. Three other questionnaires had

**Table 14.1** SUMI item 1 (efficiency): This software responds too slowly to inputs

| | Agree | Undecided | Disagree | |
|---|---|---|---|---|
| Profile | 13 | 1 | 5 | |
| Expected | 3.61 | 2.9 | 12.49 | |
| Chi Square | 24.44 | 1.25 | 4.49 | 30.18[a] |

[a]Indicates 99.99% confidence

**Table 14.2** SUMI item 32 (affect): There have been times in using this software when I have felt quite tense

| | Agree | Undecided | Disagree | |
|---|---|---|---|---|
| Profile | 2 | 2 | 15 | |
| Expected | 7.19 | 2.89 | 8.93 | |
| Chi Square | 3.74 | 0.27 | 4.13 | 8.15[a] |

[a]Indicates 95% confidence

**Table 14.3** SUMI item 29 (control): The speed of this software is fast enough

| | Agree | Undecided | Disagree | |
|---|---|---|---|---|
| Profile | 6 | 1 | 12 | |
| Expected | 10.65 | 3.17 | 5.1 | |
| Chi Square | 2.03 | 1.48 | 8.99 | 12.51[a] |

[a]Indicates 99% confidence

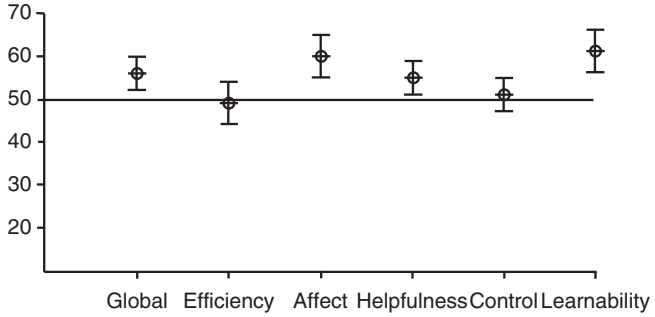



**Fig. 14.9** Results of the SUMI analysis

between one and three statements unanswered (in the analysis, the responses to these questions was considered as being *Undecided*). An analysis of the responses given to the questionnaires was completed using the SUMI scoring program SUMISCO. The results of the analysis are shown in Fig. 14.9.

The vertical axis of Fig. 14.9 shows the normalized scale of the SUMI test, which is constructed to have a mean of 50, a total range of 100. The following tables show statements where the subjects made significant assertions about *Squirrel* (for brevity, only the statements where a significant statement was made are shown). Each table provides a description of the statement (e.g., *This software responds too slowly to inputs*), a *profile* of the user responses (the number of people who *agreed* or *disagreed* with the statement or who were *undecided*), the *expected* response (based on data from a standardisation database), and a *Chi Squared* measure (a measure of goodness of fit between the observed and the expected values). The standardisation database, which was produced from the responses of more than 1,000 users, is used to generate an expected pattern of response for each SUMI item. The expected pattern of response is then compared with the actual, obtained pattern. In summary:

- more users than expected agree that *Squirrel* responds too slowly to inputs, (99.99% confidence),
- more users than expected disagree that the speed of *Squirrel* is fast enough (99% confidence).
- more users than expected disagree that there have been times in using *Squirrel* when they have felt quite tense (95% confidence).

The users who completed the SUMI assessment were satisfied with the usability of *Squirrel* (Global usability was rated 4 scale points higher than the average). Furthermore, users liked using *Squirrel* (Affect was rated 8 scale points higher than the average), but do not seem to find the presentation of the user interface particularly attractive (a further analysis needs to be undertaken to identify the particular aspects of the presentation that the users were not satisfied with). *Squirrel* was considered easy to use (Learnability was rated 11 scale points higher than the average and Helpfulness was rated 4 scale points higher than the average). Control, the feeling that the software

is responding in an expected and consistent way, was considered average. The majority of users considered *Squirrel* to be no more efficient than other comparable software systems (Efficiency was rated 4 scale points below the average). Users found the response to inputs, that is, the time to display a set of results, too slow.

## 14.6. Conclusions

The overall aim of the BT case study was to investigate how the semantic technologies researched and developed within the SEKT project could enhance the functionality of BT's digital library and address some of the challenges outlined at the beginning of this chapter. SEKT technology was used to integrate web content into our digital library prototype system.

The systematic inspection and the *cognitive walkthrough* analysis were found to be good tests of usability, enabling us to identify and correct a number of fundamental usability deficiencies prior to the field test phase of the evaluation.

The response of users to the new functions of the *Squirrel* and *SEKTagent* applications, when compared with the current technology in use in BT's digital library, was positive. The findings of a task-based evaluation of *Squirrel* revealed some promising gains in the average perceived information quality (PIQ) of the search results when compared with the current search engine.

A SUMI assessment of *Squirrel* showed that users rate the application positively and believe that it has attractive properties, but were concerned about its performance and speed.

Overall, the BT case study, in particular the positive results of the user evaluation, has given us confidence that the semantic technologies researched and developed in the SEKT project can be used to good effect in the next generation of semantically enabled search and browse and search agent applications. Some of these functions are expected to be integrated into BT's digital library in the near future.

**Acknowledgement** This work was supported by the IST Programme of the European Community under SEKT (IST-IP-2003-506826).

## References

Alsmeyer D, Owston F (1998) Collaboration in Information Space. In: Proceedings of Online Information 98. Learned Information Europe Ltd., 31–37.

Chen H (1999) Semantic Research for Digital Libraries, D-Lib Magazine 5 (10), http://www.dlib.org/dlib/october99/chen/10chen.html.

Davies J, Studer R, Warren P (eds) (2006) Semantic Web Technologies: Trends and Research in Ontology-based Systems. Wiley, New York ISBN: 0-470-02596-4.

Duke A, Glover T, Stincic S (2006) Search and Browse Facility – Final Version. SEKT project, deliverable D5.2.2.

Kirakowski J, Corbett M (1993) SUMI: The Software Usability Measurement Inventory, British Journal of Educational Technology 24 (3): 210–212.

Kiryakov A, Popov B, Ognyanoff D, Manov D, Kirilov A, Goranov M (2003) Semantic Annotation, Indexing, and Retrieval, In: 2nd International Semantic Web Conference (ISWC2003), 20–23 October 2003, FL, USA. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence) vol. 2870, Springer, Berlin Heidelberg, 484–499.

Lynch C, Garcia-Molina H (1995) Interoperability, Scaling and the Digital Libraries Research Agenda, A report on the May 18–19th 1995 IITA Digital Libraries Workshop, http://dbpubs.stanford.edu:8091/diglib/pub/reports/iitadlw/main.html.

Nielsen J, Molich R (1990) Heuristic evaluation of user interfaces. In: Chew C Whiteside J (eds) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People (Seattle, Washington, United States, April 01–05, 1990). J. CHI '90. ACM Press, New York, 249–256.

NSF (2003) Knowledge Lost in Information, Report of the NSF Workshop on Research Directions in Digital Libraries, June 15–17, 2003. http://www.sis.pitt.edu/~dlwkshop/report.pdf.

van Rijsbergen CJ (1980) Information Retrieval. Butterworths, London.

Wharton C, Bradford J, Jeffries R, Franzke M (1992) Applying cognitive walkthroughs to more complex user interfaces: Experiences, issues, and recommendations. In: Bauersfeld P, Bennett J, Lynch G (eds) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Monterey, CA, United States, 03–07 May, 1992). CHI '92. ACM Press, New York, 381–388.