Using Existing Website Ontologies to Assist Navigation and Exploration

Guy Saward, Mahmood Shah, Tracy Hall

School of Computer Science University of Hertfordshire
College Lane, Hatfield, AL10 9AB, UK
{g.r.saward, m.1.shah, t.hall}@herts.ac.uk
http://homepages.feis.herts.ac.uk/~comqgrs

Abstract. Successful navigation of on-line information is highly related to the structure of that information (variously called an ontology, taxonomy or classification). We present the *Naviguidance* method, which exploits this structure in browsing and searching so as to improve usability. We present three case studies where naviguidance has been applied to real Internet information systems.

1 Introduction

The ubiquitous problem of information overload is widely recognized and well reported, along with its extreme manifestation as "web" or "search rage" [1]. Many approaches exist to understand and solve this problem including novel visualizations, improved relevance algorithms, web log mining, and personalization. More recently the semantic web project addresses it by enriching the structure of information.

Rather than creating additional structure, in this paper we describe a method for using the structure already present in web sites to aid user navigation, exploration and learning. Called *naviguidance*, it uses an initial search to guide users to the most relevant information, similar to work by Olston and Chi [2]. In contrast to Olston and Chi we focus on the explicit browsable navigation structure (rather than individual page links) as a way of identifying and highlighting the most relevant navigation choices.

Our work focuses on new information retrieval methods and interface designs through the synthesis of existing models, and their application in a number of specific Internet Information Systems (IIS) domains. These include websites, digital libraries and intranets. Section 2 presents related work while section 3 describes the generic naviguidance method and architecture. Section 4 contains case studies that describe specific naviguidance implementations before concluding in section 5.

2 Related Work

Our work builds on widely used models of information retrieval and system acceptance. The first provides tools for investigating how users navigate complex information sources like websites, while the second provides insight into what makes users adopt or revisit particular sites. This is particularly important in the Internet's open market where many alternative providers may supply particular information, products or services.

The Role of Navigation in System Acceptance. The Technology Acceptance Model (TAM) [3, 4] highlights the importance of effective navigation. TAM predicts that users only use systems they both perceive to be easy to use (PEU) and perceive to be useful (PU). Therefore, effective navigation systems not only need to make a system easy to use, they need to demonstrate the usefulness of the system.

Navigation Issues. The problems of current information retrieval methods are well known [2] and typically center on two key issues: the difficulty in interpreting browsable options correctly, and the difficulty in specifying searches effectively.

Difficulty in browsing arises when there is no clear perceived match between a site and a task [5] and users get lost. Exploration can improve task/site match but may require too much effort. Search difficulties come from limited user input and a focus on a small subset of results. Over-precise queries return too few results [7] and limited PU, while imprecise queries generate too many options to review and low PEU.

Ontologies and Navigation. Ontologies are a key component of the semantic web. Although rich ontologies support data exchange or reasoning [8], query expansion and complex visualization, we are interested in a pragmatic interpretation of website structure or classification schemes as ontologies [9] to support simple visualization .

3 Naviguidance

The objectives of naviguidance are to assist and encourage exploration of IIS. The former should impact the Perceived Ease of Use (PEU), as defined by TAM while the latter should promote understanding of the breadth of site content, thereby increasing Perceived Usefulness (PU). The end result should be an improvement of users' attitudes to using the system and a consequent increase in the overall site usage.

Naviguidance uses the site ontology: in filtering search results to increase search effectiveness; and in directing users in browsing. This approach should aid the interpretation of browse results as well as reducing the need to specify searches so exactly. Our IIS design experience e.g. [10] leads us to advocate the following principles.

Table 1. Principles for the design of navigation methods

Browse Principles	**Search Principles**
• Give user full control of browsing	• Exploit "good enough" search results
• Show whole taxonomy when browsing	• Specify search context via categories
• Re-use existing static structure	• Specify context after search
• Explore structure from the top	• Support more general queries

Naviguidance uses an existing search engine in a site to dynamically highlight the parts of the site ontology that contain items relevant to a particular query. The method is described with reference to an abstract architecture (Fig.1).

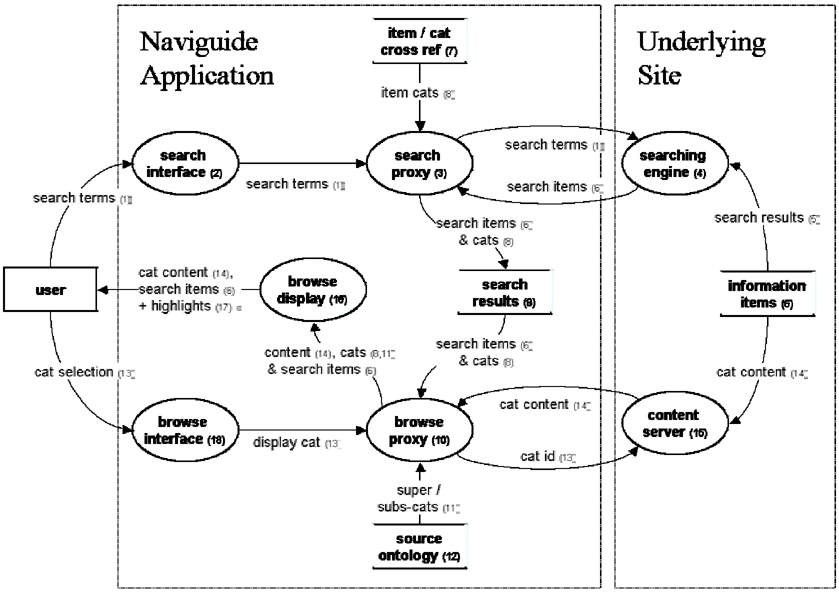

Fig. 1. A logical view of the naviguidance architecture

Naviguidance consists of the following steps (numbers relate to the items in Fig. X):

1. The user executes a search (1) using the naviguide search interface (2)
2. A proxy search service (3) passes the query to the site search engine (4).
3. Search results (5) are selected from available items (6) and returned to the proxy.
4. The proxy cross-references the results with the site ontology (7) and the results with their associated categories (8) are stored (9).
5. A browse proxy (10) extracts sub/super-category data (11) from the taxonomy (12) .
6. In parallel, the browse proxy uses the current location in the site (13) to retrieve the basic content (14) to display from the site's content server (15) .
7. A display process (16) determines which search results to show and categories to highlight (17) in the browsable content displayed to the user.
8. Browsing a new category in display (18) updates it content via steps 5 to 7.

As the user browses through the site ontology, the browse proxy in conjunction with the display process fetches the content to display and uses the site ontology to keep relevant categories highlighted. In addition, the search results will be continually filtered to show the search results related to the current category. In this way, the naviguide process continually highlights the parts of the site where search results are located, thereby guiding users in their browsing to the most relevant parts of the site. Crucially, highlighting is done within the current site design, rather than as a separate element [11], using an existing structure rather than a dynamic clustering.

4 Case Studies

Naviguidance has been implemented for a variety of systems including the Open Directory Project, a grocery e-commerce site and telephone directory. Three applications are described here to show naviguidance can be implemented using a variety of technologies and architectures for both the naviguidance application and underlying system. In addition, it shows naviguidance applied to a variety of different ontologies that are displayed in a number of different ways. The first two studies use an Apache / Tomcat / Java proxy server architecture. The third uses client-side Javascript.

Case Study 1: A Global Airline (GA). This study used GA's Lotus Notes intranet. GA wished to assess the effectiveness of previous design changes to the site structure, and to assess the feasibility enhancing the existing search engine. This gave us the opportunity to build an add-on naviguidance system and to evaluate it alongside the current GA system [12]. The system included basic category highlighting and filtering of search results as described above. In addition item counts were shown for each category and users could also choose to exclude certain categories of results.

The site ontology is displayed as a collection of links in a navigation bar (see Fig.2) with a single level shown at a time. The current location in the ontology is shown as a breadcrumb trail. Recommendations for browsing the ontology are generated using a Lotus Notes search and highlighted in a contrasting text style and arrowhead images, although the user (as with any naviguidance system) is free to browse whichever category they choose given all the available options are shown at all times.

Case Study 2: A Multiple Retailer (MR). This study used the e-commerce site of a Multiple Retailer (MR) developed using Java, JSP and Oracle to sell over 30,000 health, beauty and related products. A review of recent changes to the site gave us the opportunity to evaluate the need for, and effectiveness of naviguidance [13].

The architecture for MR followed the GA study but the presentation of the site ontology was different. The MR site used up to three nested navigation bars on the left of the screen to provide backtracking to previously seen category, a pop-up look

Fig. 2. Navigation bar for GA system showing highlights for "maternity pay" query

Fig. 3. Navigation bar for MR system showing highlights for "pushchair" query

ahead menu to select lower levels in the ontology, and a bread crumb trail showing the current location. Fig. 3 shows a two level display for the "Baby" category.

The naviguidance for MR was a simplified version of the GA system and did not include category item counts or category exclusion filters. Highlighting of recommended browse categories in both the navigation bars and pop-up menus was done with a contrasting text style and a prefixed ">" character in the category name.

Case Study 3: A Naviguidable OPAC (NaviCat). This study uses the online public access catalogue (OPAC) of a library system. The OPAC provides access to items including books, journals and electronic journals. The study is part of an investigation of the factors driving the use of competing information sources such as subject specific portals and Internet search engines. In particular we wish to evaluate if

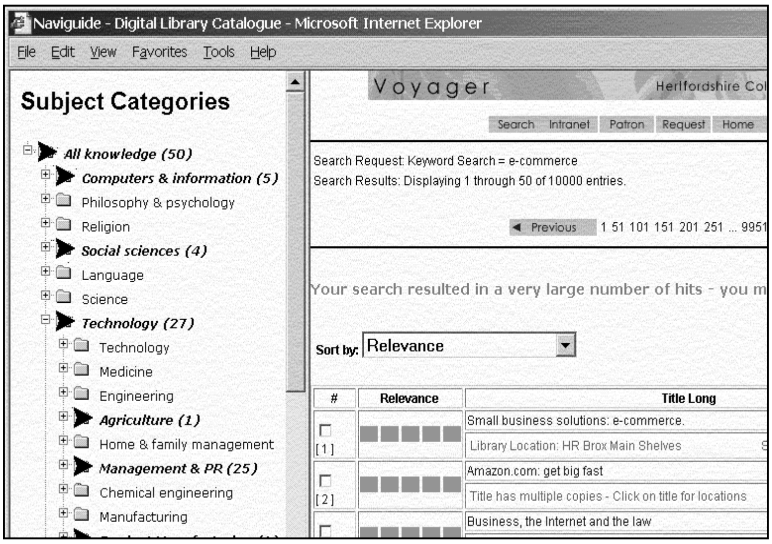

Fig. 4. The NaviCat interface showing site ontology (on left) and standard OPAC search results (on right). Highlighted ontology categories show the classification or location of search results

improved subject based searching and browsing can promote the use of high quality academic sources. NaviCat has just been developed and is about to be evaluated.

NaviCat uses a frameset (see Fig. 4) to display the standard OPAC content alongside the site ontology, which uses Dewey classification. The ontology is shown as an expandable directory tree that can be expanded or collapsed as needed. This allows users to explicitly browse the implicit ontology. Search results and classes are taken from the OPAC page and used to select the categories to highlight. The highlighting uses the same style as the GA system. Clicking on a specific category removes unrelated search results from the OPAC display. For example, clicking on "Management & PR" in Fig. 4 removes the third search result (related to law) and 24 other results not related to e-commerce (the query used) and management (the category selected).

5 Findings

Our findings relate to the ability to implement a naviguidance system for existing IIS, and to the effectiveness of naviguidance in improving the use of IIS. The case studies show the naviguidance can be built on top of a variety of existing systems that have a search engine, a browsable structure, and a means of cross referencing the two.

Naviguidance's impact on system use was assessed through individual user analysis (with 54 subjects) of the GA and MR systems. These show a positive user attitude towards naviguidance. After improving search relevance, this was rated the most useful development ahead of: limiting search by topic; increasing results detail; and increasing the number of search results shown. This implies that users want fewer, more relevant search results and that naviguidance is useful way of achieving this.

Although 60% of users thought naviguidance would help find information more quickly, this is not proved. No *significant* difference was found using task measures such as speed, completion rate or ease. However, ease of use was significantly correlated with knowledge or expectation of where items were located in the ontology.

Furthermore users who were more successful in browsing the site were more successful in acquiring this knowledge. The result is a virtuous circle where more browsing leads to more *successful* browsing. By encouraging users to engage with a site's ontology, naviguidance does appear to have a role in successfully promoting the use of particular sites. A long-term trial of NaviCat provides the opportunity to test this.

References

[1] D. Sullivan, "WebTop Search Rage Study," The Search Engine Report, February 5, 2001.

[2] C. Olston and E. H. Chi, "ScentTrails: Integrating Browsing and Searching on the Web," ACM Transactions on Computer-Human Interaction, vol. 10, pp. 177–197, 2003.

[3] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology.," MIS Quarterly, vol. 13, pp. 318-341, 1989.

[4] H. v. d. Heijden, "e-Tam: A revision of the Technology Acceptance Model to explain revisits," Vrije Universiteit, Amsterdam Research Memorandum 2000-29, 2000 2000.

[5] M. A. Katz and M. D. Byrne, "Effects of Scent and Breadth on Use of Site-Specific Search on E-Commerce Web Sites," ACM Trans. on CHI, vol. 10, pp. 198–220, 2003.

[6] B. J. Rhodes and P. Maes, "Just-in-time information retrieval agents," IBM Systems Journal, vol. 39, 2000.

[7] H. Stelmaszewska and A. Blandford, "Patterns of interactions: user behaviour in response to search results.," presented at JCDL Workshop on Usability, 2002.

[8] C. Brewster and K. O'Hara, "Knowledge representation with ontologies: the present and future," IEEE Intelligent Systems, vol. 19, pp. 72 - 81, 2004.

[9] D. Fensel, D. L. McGuiness, E. Schulten, W. K. Ng;, G. P. Lim;, and G. Yan;, "Ontologies and electronic commerce," IEEE Intelligent Systems, vol. 16, pp. 8 - 14, 2001.

[10] G. Saward, "Navigating Heterogeneous Knowledge Sources in a Customer Service Environment," in Internet-Based Organizational Memory and KM, Idea Group, 2000.

[11] B. Richard, P. Tchounikine, and P. Jacoboni, "An architecture to support navigation and propose tips within a dedicated Web site," 1st Int'l Atlantic Web Intelligence Conf, 2003.

[12] G. Saward, T. Barker, and T. Hall, "Changing Designs for Internet Information Sys-tem Navigation, Usability and Acceptance," Uni. of Hertfordshire Tech Report, 2004.

[13] G. Saward, T. Hall, and T. Barker, "Information Scent as a Measure of Usability," presented at 10th International Software Metrics Symposium, Chicago, 2004.