Electronic resource discovery systems: from user behaviour to design

Hanna Stelmaszewska, B.L. William Wong, Simon Attfield and Raymond Chen
Interaction Design Centre
School of Engineering and Information Sciences
Middlesex University
The Burroughs, Hendon, London NW4 4BT
United Kingdom

h.stelmaszewska / w.wong / s.attfield / r.chen @ mdx.ac.uk

ABSTRACT
Information seeking is a central part of academic development for both students and researchers. However, this is often hindered by complex and highly complicated electronic resource discovery systems. One approach to improving these resources is to understand the difficulties and likely causes of problems when using current systems and how people develop their searching, retrieval and storage strategies. These might provide useful information about the requirements for future design. In this paper we present our findings from UBiRD, a project investigating user search behaviour in electronic resource discovery systems based on a qualitative study of 34 users from three UK universities. We then describe how the information gathered during the study helped inform the design of INVISQUE, a novel non-conventional interface for searching and querying on-line scholarly information. In addition, the theories and design principles used during the INVISQUE design are discussed.

Author Keywords
Information seeking, resource discovery systems, search, design principles, interactive visualization, user interface

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous, H5.2 User interfaces:

INTRODUCTION
Information seeking is an important part of the intellectual development of academics and scholars who have access to many specialized electronic resource discovery systems. Electronic resource discovery systems (ERDS) include electronic databases, e-journals portals (such as EBSCO EJS, Emerald, ProQuest), federated search engines, catalogues, e-books and various electronic newspapers subscribed to by higher education institutions. Despite availability of these resources, scholars often bypass the library-provided ERDS and use Google, Google Scholar or Wikipedia. One way to address these issues is to understand how people develop their information seeking strategies and how the current resource discovery systems support or deter them.

This paper reports on key aspects from two related projects: UBiRD (User Behaviour in Resource Discovery), and INVISQUE (INteractive Visual Search and Query Environment) both funded by JISC, the Joint Information Systems Committee in the UK. One of the challenges of the UBiRD study was to investigate how scholars use electronic resources when searching for academic material. The aim of the INVISQUE project was to propose and prototype a new innovative user interface and search concept that would address user problems identified during the UBiRD study.

The empirical findings discussed in this paper are based on a part of the UBiRD study. They illustrate how knowledge obtained from studying user search behaviour can be used as a starting point in the development of a ‘new generation’ of ERDS as exemplified by the INVISQUE project.

The remainder of the paper is structured as follows: section 2 describes the UBiRD study; section 3 presents the findings that were carried forward to inform the design of the INVISQUE system; section 4 discusses the theories and design principles applied within INVISQUE and provides a summary of the system features. Finally, the discussion and conclusions in section 5 are presented.

UBIRD STUDY: DESCRIPTION
A qualitative research approach was adopted to identify, understand and compare the information seeking behaviours of scholars searching for quality materials using different ERDS. In addition, the study focused on problems and challenges users encountered during their search sessions.

In total, 34 volunteer (16 female and 18 male, aged between 22-55 years) undergraduates (UG), postgraduates (PG), and post-doctoral researchers (Experts) in Business and

NordiCHI 2010, October 16–20, 2010, Reykjavik, Iceland.

Economics were recruited for the study (see Table 1). The participants represented students from three categories of UK university: a large research-intensive university, a smaller research-led university and a former polytechnic, which became a new teaching and research university. Participants were given pseudonyms (e.g. MP2, LP5, CP4) to maintain their anonymity. The intention in working with these groups was to study users with different levels of understanding of resource discovery systems.

		Under-grads	Post-grads	Researchers	Total
Stage 1: Focus Groups		6	3	0	9
Stage 2: Observations and in-depth interviews	Large research intensive university	5	4	3	12
	Former polytechnic	5	3	4	12
	Smaller research-led university	0	6	4	10
	Total	10	13	11	34

Table 1. Distribution of participants across universities

The study was conducted in two stages: (i) focus groups, (ii) user-observations.

Focus groups are a qualitative method, which can be used alone or with other qualitative or quantitative methods to improve the depth of understanding of the needs and requirements of users [10]. Two sessions with a total of 9 students each were conducted. We used these sessions to learn about the language and concepts used by students when searching online for scholarly material. Data gathered from the focus groups helped in the development of three task scenarios of varying levels of difficulty and ambiguity that were used in the user-observation study.

The user-observation stage consisted of a series of 2-hour sessions. Each session comprised of an observation with 'think aloud' followed by an in-depth interview. The observation with 'think aloud' was used to investigate what people do, how they do it and why, when searching for information. Participants were asked to individually carry out three information search and retrieval tasks using the ERDS. The tasks were of increasing level of ambiguity and difficulty starting with the simplest Task 1 then Task 2 and the most difficult Task 3. Each observation session lasted between 40 – 80 minutes, was screen and audio recorded and later transcribed for analysis. The tasks are briefly presented below:

Task 1: Find a range of examples from film and television programs, which illustrate product placement 'in action'.
Task 2: Find evidence of film tourism from a range of different film industries to illustrate the impact this may have had on tourism.

Task 3: Imagine that you are the brand manager for a new range of mobile phones for Nokia; you are required to produce evidence to demonstrate how you might use the film/television medium as a way of reaching your target audience.

Following each observation session we carried out *in-depth interviews* using a combination of CTA techniques such as Critical Decision Method (CDM) and the Cued Recall Method, to provide supportive evidence for our observations. The CDM is a semi-structured, open-ended and retrospective interview technique originally designed to elicit expert knowledge in order to understand the nature of expert decision making in naturalistic settings [8]. CDM uses critical and memorable incidents as a catalyst and framework for the interview ([6], [8]). For this study, we maintained the structure and approach, but adapted the probes to investigate the nature of information seeking strategies used, decisions made and problems encountered when participants were searching for information.

In the Cued Recall method, participants are presented with selected segments of the screen recordings in order to prompt recall of the specifics of interaction at particular points in time [9]. Cued Recall helped us to further probe aspects of the participants' interaction that we did not understand or had doubts about. The interviews focused on: (i) identifying the expertise and underlying rationale for the search behaviour demonstrated during the observation session; (ii) problems and difficulties users experienced (iii) clarifying ambiguity that occurred during the observation session; and (iv) exploring the differences in attributes between physical and electronic libraries.

The data from all 34 observations with 'think aloud' and interviews were analysed using the Emergent Themes Analysis (ETA) approach [21]. ETA is a technique for rapid yet systematic and rigorous extraction of key themes emerging from the data. The data can then be identified, indexed and collated. Starting with the observation transcripts, we identified broad themes by indexing and collating the data at the concept level. The data was then further broken down and organized within each theme allowing for the concepts and the relationship between concepts to be discovered. The same strategy was then used with the interview data. The details and supporting evidence for each theme were then organized into categories following the questions developed for the CDM interviews.

UBIRD STUDY: FINDINGS

The following sections discuss the issues that emerged from the study.

Popular resource discovery systems

The study showed that different user groups across all three institutions used a variety of resources when looking for academic material. Postgraduates and Experts used EBSCO, ProQuest and Emerald whereas Undergraduates preferred to use the library catalogue and federated search engines (these differed between institutions). In the case of resources freely available on the Internet, Google was top of the list followed by Google Scholar, YouTube and Wikipedia. The choices participants made for using particular resources were based on their prior knowledge and experience of resources, knowledge of their strengths and limitations and the belief that selected resources would provide reliable and relevant results.

The visual representation of search results

Resource discovery systems typically produced lists of search results (often many pages), which users had to scroll through, clicking on a numerous links trying to find the results they required. Even when the users got to see the article, journal or book, their 'journey' was not over. When drilling-down to access elements such as full text, table of contents, related paper, keywords and so on before making the final decision about the relevance of the documents, they spent time and faced different barriers. Often after following a promise of access to a full-text document, users were presented with a set of links to different organisations that held a copy of required document. However, these were not always available. This was time consuming, did not always provide required information and most of all irritated users.

Searching for information: using combination of searches

When looking for information the study groups very rarely used only one search strategy (e.g. Simple Search) but changed their strategy during the information seeking process depending on the results returned (i.e. refine or re-formulate a search, abandon a search or resource or change resource). Moreover, not all of these searches were used with the same level of frequency. For instance: the Link Search (follow the hyper-links within documents to find relevant information) was one of the most popular searches used. Participants followed links in order to find more information, to confirm previous findings or simply to explore other possibly related material. The 'Advanced Search' (where a number of search terms are combined with Boolean operators such as AND, OR, and NOT) was used occasionally by the user-groups. The study showed that more experienced users performed this type of search not only to limit the scope of their search but also to bring together results of a number of searches within one action (e.g. 'television' OR 'movie' OR 'film' AND 'product placement' AND 'Times'). It would appear that the reason for using the 'Advanced Search' lay in the users' knowledge about the scope, structure and kind of information these databases provide. "Normally I use simple when it doesn't come up well I may... or there is too much, too many results I will go to advanced." (LP5, UG).

Poor understanding of structure and search mechanisms work

It was often not obvious to users what information was available, contained, organised or stored in the electronic systems. "I don't always know which is the most appropriate [database]" (CP2, EX). They often did not have a good or useful 'mental model' that they could use to explain to themselves how to search the disparate data sets. They simply could not tell how big the data sets were, what they covered, and how useful they were to their information search problem. As such it was difficult to find, use, and to re-find information. "…all the resources …for data are a bit more difficult and change quite a bit …when you look for data it is not that easy so you have to learn and ask" (CP1, EX). "If I did not have luck on there I would go to Web of Science, but it's so messy. It used to be awful but you knew how to work it, you can't type 'strategic management journal' you have to type 'strat manag jrnl' but you need to know these. They have had a redesign, which did not improve its design. The new way was more awful" (CP4, EX). In order to access a particular database users had to learn procedures with limited transferability. Search engines that had fewer rules, are less complex and are hence less procedurally rigid, allowed users to find information at a semantic level whereas databases required users to know the procedures, have some basic idea of how the data was organized, indexed and which search mechanisms were employed. "[Journal Citation Index'] has the worst searching capability ever. It doesn't do the nearest match [i.e. smart text searching]. You need the exact title and it isn't easy to find that either " (CP2, UG).

Poor usability and complex user interfaces

Formulating queries to find information is highly dependent upon the functionality and user interface of a specific resource discovery system. Current systems are often built on database structures that participants found difficult and complex to use. They required users to have procedural knowledge for using a particular database and also have some basic knowledge of how the data is structured, organized and what search mechanisms were employed. "Going to the library database and then putting in some keywords, first of all there are so many options there, you know, do you want this, do you want that, I mean keep it simple" (MP12, EX).

This higher level of difficulty amongst library electronic discovery systems distracted users from focusing on the content, analysis and evaluation that would help them learn and make sense of what they have discovered. Users did not like a user interface that was too complicated as it would require investing a lot of time navigating and trying to understand how the system worked. What users preferred is a system that they can use straight away without having to spend much or any time learning how to operate it. "This is much more difficult to use [referring to Library Resources] ... Google in that sense is much easier to use" (LP1, UG). Current library systems are too complicated and users often get lost or cannot find the information they want.

Dealing with multiple paradigms and interfaces across systems

It was observed that while searching for information participants often worked with more than one resource or system at a time. Navigating from one system to another – all of which had different procedures and interfaces for searching, limiting, refining, indexing, saving, storing or exporting. This is confusing for all users. Participants have to 're-frame' their minds when switching between resources, which required patience, persistence and was time consuming. Moreover, during this process they often lost track of their progression and needed to start the process from the scratch again. Often search features appeared to be too complex and did not help inexperienced users formulate their searches and select appropriate options in order to narrow the search results and obtain relavent documents. "There are too many words. Normally I'd prefer a search box … not sure what to click on", was a comment of a Postgraduate who was puzzled by what was meant or offered by the 'Free Web Resources' page within the library resources.

'Phase shift', 'time out' and authentication issues

When searching for information users often changed a resource believing that the resource they had been using was no longer appropriate. A search may lead to a dead end where repeated searches did not reveal any useful leads, as if coming up against a brick wall. This is often when a 'phase shift' occurred, where the user switched resources and search behaviour. If they were using library subscribed resources such as one of the bibliographic databases, they may leave it and go to the Internet and use tools like Google to find their bearings, learn more about the search topic in order to find better search terms, or to re-do their search. CP4 (EX) "I'm feeling annoyed by the search I have done – and this is all I have found [when using ProQuest] … I'm going to go to Google". During the 'phase shift' process users coordinated multiple resources (moved from one resource to another and then back to the original one) to obtain material that was not available on the Internet (e.g. no access to full text). When going back to library subscribed resources users were required to log in. In addition, they needed to find their way back to the required resources, which was not always straight-forward operation. Participants did not expect the system to 'time out' without giving any warning. This caused irritation and annoyance amongst participants and created barriers to restoring coordination.

Another important issue often occurring during a 'phase shift' was the 'time out' issue and authentications. The 'time out' created problems because users invested time and effort and the investment was lost. All the searches performed were lost and there was no record of them after re-logging onto the system. This was especially irritating when participants had been working on the system for a while and had created many searches that were lost when the system 'Timed out'. "It is irritating because all the searches are lost! The library catalogue times out after like 5 minutes. So frustrating and it makes you not want to carry on with your search" (CP4, EX). Users were also irritated and confused when asked to type in their Athens username and password again to access the resources. Some participants abandoned the use of library resources altogether when prompted to authenticate, as they did not remember their log-in details. Remembering numerous log in details strain user's memory load which they often want to avoid and instead they select alternatives such as Google.

Storage and workflow

This study showed that one of the important activities people do during a search and retrieval process is the storing of information. This happens at different stages of the information seeking process with the first storage usually taking place when participants evaluate a list of results and temporarily store individual documents/material using tabs. These tabs are then re-visited for further evaluation and if information is relevant, stored permanently using different means; from notes in a Word document, saving downloaded material into a folder, bookmarking, to more sophisticated features provided by various resource discovery systems (e.g. RefWorks, Endnotes, My Research). Storing relevant information allows users to keep track of material, organise their references, but importantly, also allows them to re-visit at anytime. The notion of tracing back to documents previously found or storing information in the systems' pre-defined storage area was not always an easy task for the UBiRD users. It was observed that users were un-aware of some of these features such as 'alerting' or 'save searches'. Only one participant from the study (CP9, PG) saved his searches and then after failing to obtaining satisfactory results (he performed 10 searches in total) he went back to 'Search History' to select the search that returned the highest number of results.

It was also observed that participants often gathered information from various resources and put it together in an easily accessible place. This was either a folder, a bookmark in a browser or a number of tabs. Participants want to have access to the stored material at anytime and at a 'click away'. "I think the concept of saving to the desktop is getting more and more into the background I guess, because I tend, you're right, I tend to bookmark things more than save them because I am assuming it will be there when I click the bookmark again" (MP12, EX). Although browser's bookmarks were applied by the UBiRD users, none of them demonstrated knowledge of Web 2.0 bookmarking facilities. Instead, they made reference to not having their locally stored bookmarks available during the study, which created problems finding or retrieving found information. The existing storage spaces within various resource discovery systems were used very rarely and one can only assume that users were not aware of their existence, or not sure about what they offered. It appears that the current systems lack good ways of storing and retrieving documents allowing users to create repositories of information that can be accessed easily and be transferable across different resources.

Access to full text documents

Another important issue that emerged from the UBiRD study related to the availability of requested documents. Surprisingly, even when a document was not available, the systems suggested otherwise on many occasions. Users were annoyed when a promise of a link to a full-text article (as in those references found via federated search engines and Google Scholar) did not result in the article(s) being available and required a further step in the process of accessing materials. “Because we don’t have the full text, I’d go to SFX and follow any link it’ll give me. Although sometimes this is frustrating because even though you follow the links, we don’t have access to it. So you get there and you still can’t download it, which is just plain irritating” (CP4, EX). Users abandon searching on library-subscribed resources when this occurs too frequently and turn to freely available resources on the Internet. This kind of situation raised expectations and often upset and irritated participants as they wasted time without obtaining the required document.

INVISQUE: A BRIEF INTRODUCTION

INVISQUE (pronounced in•vi•sic) is an early stage rapid prototype intended as a concept demonstrator, and at the time of writing, had yet to be fully evaluated. It was developed to investigate how Information Foraging Theory, and other design principles such as focus+context, and Gestalt pattern perception, could be applied to create a novel interface design we call *interactive visualization* that would address the problems found in the UBiRD study. The design we proposed for INVISQUE uses animation, transparency, information layering, spatial layout and pattern creation techniques to emphasize relationships, and is orchestrated in a way that facilitates rapid and continuous iterative querying and searching while keeping visible the context of the search. This is intended to minimize problems such as ‘What Was I Looking For?’ or “WWILF-ing”, where users lose their train of thought when searching through numerous lists. The design was also intended to create opportunities for discovering relationships and unanticipated discoveries within the data [17].

INVISQUE was developed with a combination of rich animation tools such as Adobe Flash and Adobe Flex using ActionScript and a XML (MXML) dataset as the test database. This will enable the later connection of the Rich Internet Application front-end with enterprise systems such as the library catalogue and the various publisher resource discovery systems. It will also be able to run on any web browser or desktop. In its current version (v1.0), most interactions are performed using the mouse as <clicking> or <dragging and dropping> the data from the searches. However, the current mouse-driven point-and-click interaction can be easily replaced with multi-touch and gestural interaction.

In INVISQUE, search results appear on the screen in a large windowless and borderless display space where size is limited by hardware memory constraints. The search result for each journal article appears as an 'index card' with bibliographic attributes such as title, keywords, authors, journal, and number of citations (Figure 1).

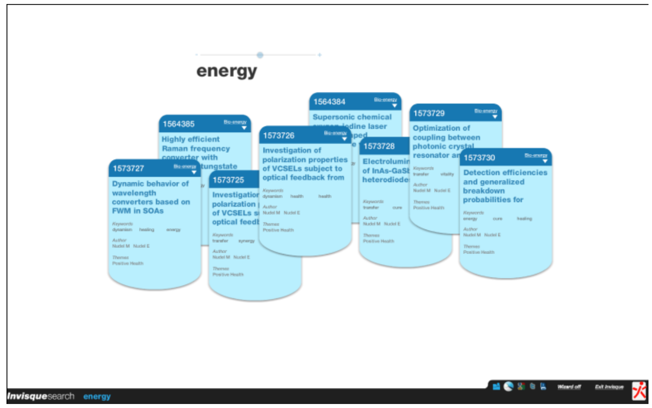

Figure 1. Search results screen showing ‘index cards’ with bibliographic information, e.g. title, author, journal, citations. Relationship: x-axis=year published, y-axis=num. of citations.

The number of index cards to be displayed can be controlled by a slider (line with dot above the search term on the display). The next basic design feature of INVISQUE addresses UBiRD problem or need for visually representing relationships within the results in a meaningful way. The ‘index cards’ are presented and organized along the x- and y-axes, where the x-axis represents time (year of publication) and the y-axis represents the number of citations (a measure of significance). In this manner, the relationship of which article is the most cited and most recent becomes immediately apparent. The axes can be readily changed to other dimensions if needed. INVISQUE is also designed so that users can interact directly with the data of the search results. By selecting, dragging, and dropping sets of ‘index cards’, the user can activate Boolean operations such as merging sets or creating intersections between sets, revealing information that is common between sets. Here, physical manipulations of result sets are interpreted by the system in terms of Boolean operators. Following this brief introduction to the basic INVISQUE design, we next describe how it addresses the user problems encountered in UBiRD.

INVISQUE APPROACH TO PRESENTED PROBLEMS

Problem 1: poor understanding of structure, search mechanisms and complex user interfaces

INVISQUE’s solution to these problems is by applying: *Simple and implicit query formulation and filtering:*

Query formulation and *progressive* modification are supported in a number of ways: (i) Figure 2 shows the simple, any-word search field interface with the option of activating more advanced Boolean search operations that do not require the user to have explicit knowledge of Boolean operators and syntax. Users can search for articles that have “all these words”, or articles with the “exact wording or phrase”, or containing “one or more of these keywords”.

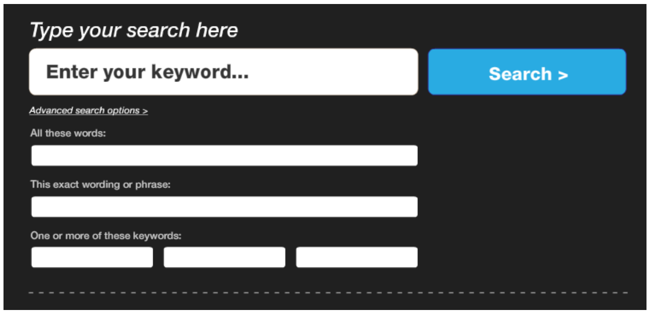

Figure 2. Simple search screen; click 'Advanced search options' to reveal more complex Boolean search fields.

(ii) INVISQUE supports the progressive modification of the search while minimizing the chances of losing track of 'where you are' by enabling the user to type a new query anywhere on the white display space with the earlier search results cluster still visible (Figure 3a).

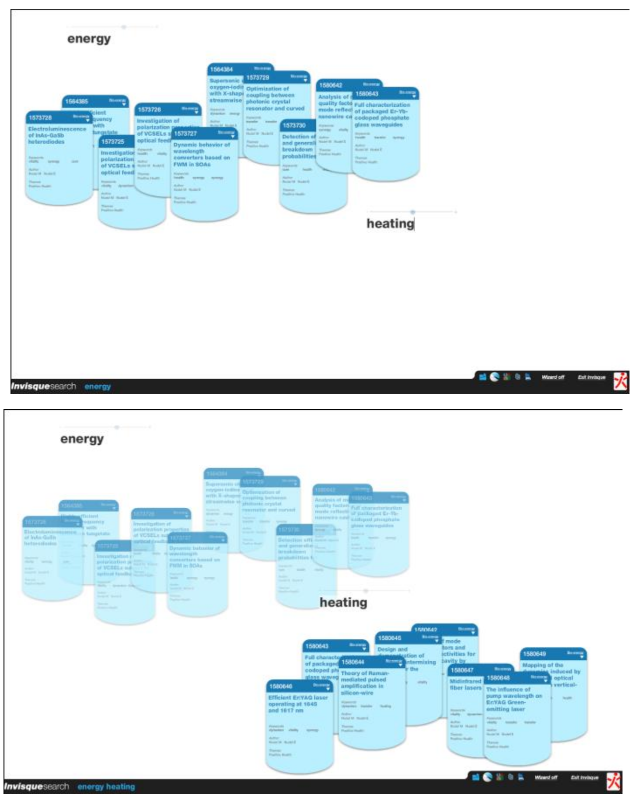

Figure 3a and 3b. Progressive modification of search strategy while keeping context of search visible (focus+context).

Instead of exiting to a separate search screen, we apply the design concept of focus+context. This starts from three premises: first, the user needs both overview (context) and detail information (focus) simultaneously. Second, information needed in the overview may be different from that needed in detail. Third, these two types of information can be combined within a single (dynamic) display, much as in human vision [3]. In addition, the application of the Gestalt Laws of Pattern Perception can be seen (see [19] for a fuller discussion). For instance, by applying the 'figure and ground law', the new results ('heating') appear brighter and in the foreground in what we call the 'primary layer'; while the previous search results ('energy') which are still visible, appear faded in the background, appearing to occur in a secondary layer (Figure 3b). In addition, by applying the spatial proximity law, where objects or events that are near to one another (in space) are perceived as belonging together as a unit, it is quite clear that there are two sets of results.

(iii) One aspect of query modification is filtering and merging of results. In INVISQUE a user can drag and drop multiple results sets to invoke Boolean operations such as create a super-set or to create an intersection. Again, this does not require knowledge of a particular syntax (Figure 4a, b).

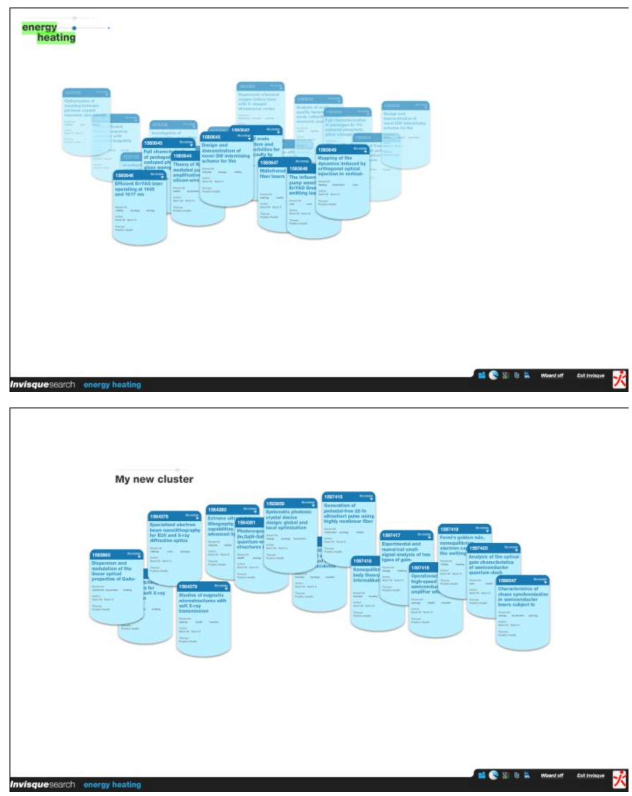

Figure 4a and 4b. By dragging and dropping the 'heating' cluster over the 'energy' cluster, we can create an intersection of the two sets 'My new cluster'.

(iv) Any of the bibliographic attributes on the index cards can also be used to progressively modify searches. For example, by clicking a keyword on one of the 'index cards', all cards across different clusters with the same or related keywords are highlighted and brought to the foreground, quickly revealing further possible relationships (Figure 5).

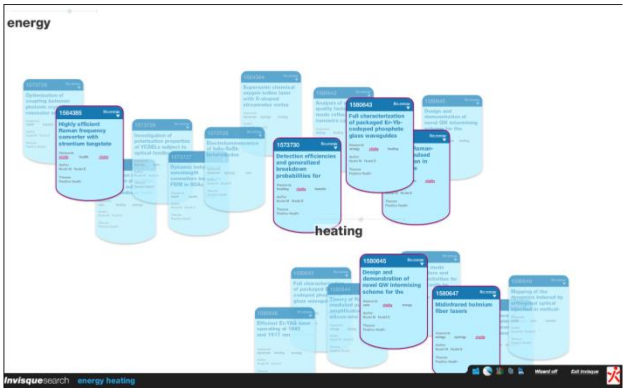

Figure 5. Common theme: common keyword and edges of card highlighted in red, and in foreground.

Attention is a process of selecting things to concentrate on, at a point in time' from a range of possibilities available [14], and is an aspect of cognition that is particularly important in the design. The highlighting combined with the figure and ground effect, directs the user's attention to the new relationship displayed–cards with the same or related keywords. This eliminates the need for linking lines, thereby reducing clutter while still being able to direct the user's attention. This allows the user to quickly identify the index cards that share a common attribute, and then construct and modify queries in non-complex, natural ways, with no required knowledge of search syntax and little explicit querying.

Problem 2: multiple paradigms and interfaces across systems

INVISQUE's solution to these problems is by applying: *An engaging and familiar metaphor.*

INVISQUE is designed around a metaphor of physical index cards on a two-dimensional infinite 'canvas' workspace. This is a departure from traditional list-style information retrieval interfaces designed to provide an engaging interface. This might reduce problems of frustration and 'phase shifting', and to promote a more familiar and less complex interaction through the use of a familiar metaphor.

Problem 3: 'phase shift', the 'time out' and authentication

INVISQUE's solution to this problem is by applying: *Seamless access to Internet resources.*

In INVISQUE, live Web and social network search systems are integrated with academic search systems. Users can switch between INVISQUE and other Internet search systems when searching to get new ideas without any interruption. This solves the issue of our findings when users have to navigate between multiple systems and encounter the problems of routes becoming difficult to follow where they can get lost, or have no access. In this demo Google and Twitter were used (Figure 6).

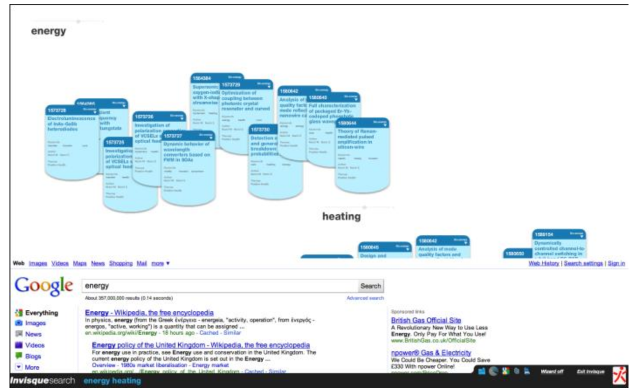

Figure 6. The Google layer seamlessly sliding over the clusters.

The maintenance of system state across sessions.
INVISQUE maintains the physical arrangements of search objects on the canvas between sessions. Users returning to the system find their past searches, documents and their own organisation of documents as they were left. This deals with the 'time out' issue where users loose all their searches and data when the system logs out automatically. They use a USB 'memory stick' where their sessions as well as different authentication details are saved transferred across when the users access the resource again. It is a solution that favours the user more in a security/usability trade-off. However, at the time of writing, this function has not yet been fully implemented and tested. When implemented, this would allow search activities to be coordinated across sessions such that the users can start again from when they left off and minimise the effects of interruption by providing strong visual cues of previous workflow context. In addition, it will minimize the load placed on memory by having to remember authentication details for a range of resources. However, securing the memory stick access is an important issue that needs to be taken into consideration.

Problem 4: storage and workflow

INVISQUE's solution to these problems is by applying: *Manipulation of search results to support visual triage and workflow.*

Sensemaking typically involves the ongoing discovery of concepts present in an information space and the development of concepts and classifications relevant to a task at hand. INVISQUE allows search results to be manipulated (freely moved) into user-determined groupings as a natural extension of the spatial metaphor to support the information triage process. The "Wizard" supports workflow by enabling the user to create sets of interest, and is currently represented by three 'hot spots'. By dragging and dropping cards on to a "Wizard" hot spot, we activate one of three specified functions: to discard, to save, and to keep aside. By dragging one or a set of cards to the "Not interested in a particular document?" hot spot, we discard the cards. This removes it from the display, and also instructs the search algorithm to lower the search weightings for documents that have those characteristics. This allows the user to filter the content by (implied) usefulness. The system would record what the user has done in this filtering process. Subsequent searches would apply filters based on what the user has done before to determine search results, filtering out the results that the user is not interested in.

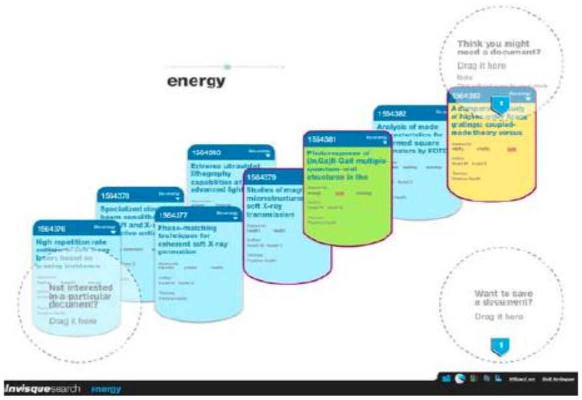

Figure 7. 'Wizard' function: Drag and drop for functions such as save, delete, set aside; and collate them into themes.

Similarly, by dragging and dropping cards on to the "Want to save a document? Drag it here" area, the information on the cards would be saved, and the colour of the card would change to green to indicate that it has been saved. It will also adjust the filters so that future searches would look for more articles in that area. Dragging and dropping a card on the "Think you might need a document?" hot spot changes the colour of the card to yellow, and indicates that it has been set aside for possible later use so that they do not get lost in the process (Figure 7). This will not affect the filtering mechanism. The sets can be collated into themes and for further action encoded either spatially or using colour as a natural extension of the interface metaphor, thus allowing users to keep track of material in a flexible, extensible and explicit way. This will not present any problems for a small number of documents. However, when a number of documents increase, the canvas might become over-crowded and impede access to the 'hot spot' areas. One way to deal with this could be to move the results around and resize them allowing easily access at all times. This, however, needs to be tested with users, which will be a part of future work.

Problem 5: access to full text documents

INVISQUE's solution to this problem is by applying: *Drill-down titles to full-text to table of contents: Supporting detail and serendipity*.

In contrast to the academic and Internet systems, in INVISQUE there is a visual interaction for the drill down function from the 'index card' to the abstract, to the table of content and to the full-text document of the desired article, which can be accessed instantly. In this case, the user does not have to open a lot of pages or go to another screen when they want to view other content. They call up a menu, from which they can select 'show full text document'. The pdf of the full-text document is presented over the clusters and the 'index card'. This supports necessary access to detail while still displaying the context of the overall search, minimizing 'what was I looking for' and 'loosing track' problems. In addition, after having seen the full-text the user may be curious about what other articles may have appeared in that specific issues of the journal. Clicking on the 'table of contents' button brings the user to the Table of Contents for that specific issue of the journal, enabling a seamless review of other papers that they may not have been specifically looking for, thereby fostering a degree of serendipity in the search process. See Figure 8.

Figure 8. 'Drill-down' function

Problem 6: supporting combination of searches

INVISQUE's solution to this problem is by applying: *Easy and implicit query formulation and filtering* (see section 'Dealing with poor understanding of the concept, structure, the way the searching mechanisms work and complex user interfaces' for details and Figure 1).

In the case of 'Link Search', INVISQUE allows users to follow different hyper-links that are available within individual 'index cards'. The users can access and view a number of different documents without 'loosing the track' of where they are as the remaining results are constantly visible in the back layer allowing a fast and easy access to any other documents that the user wishes to see (see Figure 4a&b). The 'Advanced Search' (Boolean AND) can easily be performed by merging two or more individual search results displayed on the 'borderless space' using direct manipulation.

INVISQUE AND INFORMATION FORAGING

Information Foraging Theory is a useful tool to describe information retrieval behaviour ([11], [12]). The theory refers to activities associated with assessing, seeking, and handling information sources. Information Foraging Theory helps to design interfaces that effectively support the key concepts: "(i) information: the item of information that is sought or found and the value it has in fulfilling the information need; (ii) information patches: the temporal and spatial nature in which information is clustered; (iii) information scents: the determination of information value based on navigation cues and metadata; (iv) information diet: the decision making to pursue one information source over another" [18].

An extension to Pirolli & Card's information foraging theory [11] is the model of a sense-making loop for intelligence analysis [12]. The authors suggest that the overall process is organized into two major loops of activities: (i) a foraging loop that involves process aimed at seeking information, searching and filtering it, as well as reading and extracting information possibly into same schema; (ii) a sense making loop that involves iterative development of a mental model (conceptualization) from the schema that best fits the evidence. Pirolli and Card's model [12] offers a novel and useful perspective on designing systems for information retrieval. It encourages the designer to think about the structure of the interface, how to support different searching and browsing strategies appropriate for the context of work and how to effectively use metadata cues to enhance item selection and patch navigation.

INVISQUE, guided by this concept of information patches and scents, has created a new way to initiating searches that maintains the context by keeping the context of previous searches visible. Users can create a new search by activating the search mode and then simply keying in new search terms near the results of an earlier search. In addition, the user may also type in a new search term anywhere within the borderless search space. Moreover, it displays search results by they spatial and temporal value within one display (x represents the number of citations whereas y axis represents the time line), which facilitates the information patches concept (see Figure 1). Information diet has been supported by providing users with an immediate access to the full text of a document with all necessary information to make decisions about the relevancy of information. Information scent is supported by providing rich metadata for each document allowing users to learn about a particular document before they invest more time in exploring it in detail (Figure 1 and 8).

DISCUSSION AND CONCLUSIONS

This paper presented the way in which knowledge about user's behaviour and problems encountered in the UBiRD project when searching for information was utilized in the design of a new generation interactive system. With the system, interaction does not require a high level of procedural knowledge of the system or advanced information literacy skills. The new ways of searching, retrieving, organizing and storing information presented in INVISQUE system is a step forward to a new era of ERDS. The new innovative interface concept employed in INVISQUE illustrates how next-generation systems would support semantic analysis and access of large data sets. The following sections discuss the unique features that have been utilized in INVISQUE.

Display of multi-dimensional information and dynamic manipulation of results

Over the years different concepts of multi-dimensional information visualization and direct manipulation of data have been implemented. Ahlberg and Shneiderman [1 & 2] used 2-dimensional scattergrams with each additional dimension controlled by a slider to display information. The HomeFinder [20] used dynamic queries and sliders so that users can control the visualization of multi-dimensional data. Other systems ([13], [15]) developed novel visualizations of time lines using the perspective wall. Others applied visual information foraging to present thematic results on a spatial-semantic interface ([4], [5]). More recently Stasko et al., [16] developed a system (Jigsaw) that provides multiple coordinated views of document entities emphasizing visual connections between entities across different documents. What is unique in INVISQUE system is not only the way that results are organized and displayed according to the x and y dimensions, which represent the time and the number of citations accordingly (these dimensions can be changed to other dimensions if required e.g. authors, titles, journals, conferences or concept) but also the way users can manipulate their search results on the 'borderless space'. Users can merge individual searches and create new clusters or move documents to pre-defined areas that will activate specific direct manipulation functions. In this way the user is free to move, re-organise grouping, and thereby modify not just the visual relationships, or creating new clusters, as it is in the current system. In future releases of the system, it will instruct the system to adjust, for example, the weighting of semantic distances. This would be the basic building block for future direct manipulation data analysis techniques.

'Borderless space'

The 'borderless space' gives users unlimited area/space to perform multiple searches in parallel, the results of which can be viewed and manipulated without having to move to a different page, tab or a window. This would help users to keep track of previous searches and their results as well as provide a space to create and work with different clusters simultaneously.

Organization and storage of documents: the Wizard

Jones et al., [7] discussed how people use different ways to gather or 'keep' their information such as sending emails to oneself, to others, printing, saving documents as files, passing URL's into documents, putting documents into a personal website, creating bookmarks, writing paper notes, creating toolbar links, and use the note facility in Microsoft Outlook. The problem with these methods is that they require using different systems outside the resource discovery system and they are time consuming. There are also other means of storing information, which are supported by different resource discovery tools (e.g. bibliographic management software such as RefWorks and EndNote). However, these are not easy to find or intuitive as many users seemed unaware of their availability. The INVISQUE system offers the 'Wizard' function, which is designed to be easy and intuitive, and allows users to organize, store and retrieve documents and create repositories of information that can be accessed and be transferred across different resources.

Coordinate multiple resources

The integration of live Web and social networks (e.g. Google and Twitter) offers seamless access to external resources to support users when searching for new ideas without interruptions and it lessen/minimizes the chance of 'losing the track'.

Portability of searches and authentication details cross platforms

INVISQUE uses 'memory stick' function to automatically save a user's search sessions as well as different authentication details, and transfers these across when the users access the resource again. These allow search activities to be coordinated across sessions such that the user can start again from when they left off. It also minimises the effects of interruption by providing strong visual cues of previous workflow context.

Limitations and future work

While INVISQUE offers new ways of searching, retrieving, organizing and storing information the current version presents some limitations. One of the issues that need to be addressed is scalability of the displayed results. At present the system works based on the small-scale mock data, which displays small number of results. The system needs to be tested with real data in the further versions INVISQUE in order to see how the system will behave and if the issues will continue. The system has not been rigorously evaluated by the users, which at the current state of art would jeopardize their experience and understanding of the ways INVISQUE operates as some of the features are not fully functional. In addition, the results obtained from such evaluation would not necessarily provide a true value for already stated reasons. The next step will involve using 'design briefing' evaluations that will ensure that important design issues and the functionality of the INVISQUE are considered. A further design aim is to involve users in the system's evaluation where they would review its strengths and weaknesses. However, this can happen only when the system's features are fully implemented.

REFERENCES

1. Ahlberg, C., Shneirderman, B. (1994a). "Visual information seeking: Tight coupling of dynamic query filters with startified displays." Pp. 313-321 in *Human Factors in Computing Systems*: ACM ,New York.
2. Ahlberg, C., Shneirderman, B. (1994b). AlphaSlider: A compact and rapid selector. Paper presented at the Human Factors in Computing Systems.
3. Card, S.K., Mackinlay, J.D., and Shneiderman, B. (Eds.) Readings in Information Visualization: Using Vision to Think, pp. 1-34, Morgan Kaufmann Publishers, San Francisco, Califomia, 1999.
4. Chen, C. (1999). Information Visualisation and Virtual Environments. London: Springer-Verlag London.
5. Chen, C., Cribbin, T., Kulijis, J., Macredie, R. (2002). Footprints of information foragers: Behaviour semantics of visual exploration. International Journal of Human-Computer Studies, 57(2 (August 2002)), 139-163.
6. Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). Use of the Critical Decision Method to elicit expert knowledge: A case study in the methodology of Cognitive Task Analysis. Human Factors, 40(2), pp.254-276.
7. Jones, W., Dumais, S. & Bruce, H. 2002. Once found, what then?: A study of 'keeping' behaviours in the personal use of web information. Proceeding of ASIST 2002, Philadelphia, Pennsylvania.
8. Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man and Cybernetics, 19(3), pp. 462-472.
9. Omodei, M. M., McLennan, J. P., & Wearing, A. J. (2005). How expertise is applied in real-world dynamic environments: Head mounted video and cued recall as a methodology for studying routines of decision making. In T. Betsch & S. Haberstrohe (Eds.), *The routines of decision making* (pp. 271-288). Mahwah, NJ: Lawrence Erlbaum.
10. Payne, J. (1999). Researching Health Needs. A Community Based Approach. Sage Publications, London.
11. Pirolli, P. & Card, S. K. (1999). Information foraging. Psychological Review, 106, pp. 643-675.
12. Pirolli, P. (2007). Information foraging theory: adaptive interaction with information. Oxford; New York: Oxford University Press.
13. Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). *LifeLine: Visualising personal histories.* Paper presented at the Human Factors in Computing Systems.
14. Preece, J., Roger, Y., & Sharp, H. (2002). *Interaction Design: beyond human-computer interaction*: John Wiley & Sons, Inc.
15. Robertson, G., Card, S., Mackinlay, J.D. (1993). Information visualisation using 3-D interactive animation. *Communications of the ACM, 36*, 56-71.
16. Stasko, J., Görg, C., & Liu, Z. (2008). Jigsaw: supporting investigative analysis through interactive visualisation. Information Visualisation, 8, 118-132.
17. Thomas, J. J. (2007). Visual Analytics: why now? Information Visualization, 6(1), 104-106.
18. Trepess, D. (2006). *Information Foraging Theory*. Retrieved 7 May 2010 from Interaction-Design.org. Accessed at http://www.interaction-design.org/encyclopedia/information_foraging_theory.html
19. Ware, C. (2004). *Information Visualisation: Perception for design.* San Francisco, CA: Morgan Kaufman Publishers, an imprint of Elsevier, 2004.
20. Williamson, C., Shneiderman, B. (1992). The Dynamic HomeFinder: Evaluating dynamic queries in a real-estate information exploration system. Paper presented at the ACM SIGIRi.
21. Wong, B.L.W. & Blandford, A. (2002). Analysing Ambulance Dispatcher Decision Making: Trialing Emergent Themes Analysis. In HF2002, Human Factors Conference "Design for the whole person - integrating physical, cognitive and social as- pects". Melbourne, Australia: A joint conference of the Ergonomics Society of Australia (ESA) and the Computer Human Interaction Special Interest Group (CHISIG).