From physical to digital: a case study of computer scientists' behaviour in physical libraries

Hanna Stelmaszewska[1] & Ann Blandford[2]

[1] Interaction Design Centre, Middlesex University

[2] UCL Interaction Centre, University College London

e-mail: H.Stelmaszewska@mdx.ac.uk

A.Blandford@ucl.ac.uk

Abstract. There has been substantial research on various aspects of people's usage of physical libraries but relatively little on their interaction with individual library artefacts; that is: books, journals, and papers. We have studied people's behaviour when working in physical libraries, focusing particularly on how they interact with these artefacts, how they evaluate them, and how they interact with librarians. This study provides a better understanding of how people interact with paper information, from which we can draw implications for some requirements on the design of digital libraries, while recognising that the term 'library' is a metaphor when applied to electronic document collections. In particular, improved communication with other library users and with librarians could facilitate more rapid access to relevant information and support services, and structuring information presentation so that users can make rapid assessments of its relevance would improve the efficiency of many information searches.

Keywords: HCI, Digital Libraries, user requirements, physical libraries.

1. Introduction

Physical libraries have been around for thousands of years, providing people with information on a wide range of subjects. A library acts as a collection of literary and artistic materials, or the place for preservation of audio recording, films, videos, and other formats of information. The emphasis of traditional libraries is often on storage and preservation and, for this, its activities include cataloguing and classification. In recent years, libraries have been transformed due to new technologies being introduced. This transformation is having a profound effect on the structure and organization of libraries and the ways they are used by people.

Users have become able to access electronic library catalogues and other on-line information resources, making information services – at least in principle – more effective and efficient. Such improvements of information technologies have led to the development of a wide range of digital libraries (DLs). Although, on the one hand, DLs are a major advance in information technology, offering a rich spectrum of information, on the other hand, they pose many problems (see, for example, [4]). Although a considerable amount of research has addressed issues connected to usability of digital libraries, there has been less directed towards understanding the nature of the detailed interaction between users and information in either traditional or digital libraries. This is particularly important for evaluating the interaction with digital libraries, where the metaphor picked up from physical libraries relating to the cataloguing, classification and

retrieval of information is often implemented in what appears to be an obvious, but may not be an appropriate, way.

As Duncker [13] points out, the term 'library' in the phrase 'digital library' is only a metaphor. A physical library is characterised by many features that do not transfer well to the digital context. For example, the library represents a space in which particular activities (reading, studying, being quiet) are expected, as well as being a repository of information; conversely, digital libraries offer capabilities that are not well matched by physical libraries, such as ease of accessing information within the ongoing work context (without the explicit change in focus that is often inherent in the physical activity of moving from an established place of work to the physical library). In addition, electronic publishing allows easier and quicker copying, correction, and updating of documents. Although people often prefer digital libraries to physical ones, as they can be accessed from anywhere with access to the computer network, the amount of information available is often overwhelming [4]. Although there are clear contrasts between physical and digital libraries, they share many purposes in common, and it appears likely that a better understanding of how people interact with information in a physical library can yield useful insights into how digital libraries might be better designed to support effective interactions.

The work reported here focuses on the details of people's activities in physical libraries, and relates those findings to the design of digital libraries. To set the scene for this work, we review past work studying people's search behaviour in physical libraries.

1.1 Studies of searching in physical libraries

Over the years, a number of studies had been conducted to monitor people's behaviour in physical libraries. Kuhlthau [20] studied activity of high school students working on an assignment, based on which she presented a six-stage model of the information search process, consisting of: initiation, selection, exploration, formulation, collection, and preservation. Another model, derived from Ellis' [14] empirical study of the information seeking activities of academic social scientists, identified six characteristic patterns of information seeking as being: starting, chaining,

browsing, differentiating, monitoring, and extracting. Ellis argues that if users' information seeking patterns are broken down into their basic activities and then facilitated by the systems, users should be able to mirror their interaction seeking pattern when interacting with the system. However, physical libraries give no explicit support for this pattern, and neither do extant commercially available digital libraries. Nevertheless, the strength of both models is that they are based on empirical research and had been tested in other studies – for example, with physicists and chemists [15] and engineers and research scientists [16].

Taking a finer-grained approach, Nichols *et al* [26] present a 'discovery model' of a person and an information object, describing the various stages of an interaction between a user and an object in order to evaluate a list of retrievals. These stages are: glimpse, consider, examine, use, and assess. Though each stage provides greater detail of the information object, users do not always proceed through all of them, often using only a sub-set of these activities to establish the relevance of information retrieved. A study by Twidale and Nichols [35] focused on a different aspect of information seeking: the collaborative activities of students and library staff and informal collaboration of workers in an office. They identified three significant aspects of successful information seeking: collaboration between students and librarians during the information seeking process, peer collaboration in learning information systems, and interactions between workers and their colleagues (when seeking help) rather than domain experts.

Crabtree *et al* [7] also studied collaborations between librarians and users in a physical environment, and identified several problems that they argued were crucial for the effective design of digital environments. The librarians studied were library assistants, rather than subject librarians, and queries that were outside the capabilities of the assistants were typically referred on to more specialist staff. They emphasised the importance of talking in the library, as users work towards identifying and articulating their needs through dialogue.

Another aspect of people's behaviour, investigated by Vakkari and Hakala [36], is what criteria are used to determine the relevance of a document. They investigated master students' (in

Information Studies) relevance criteria when working with references and documents during the writing of a research proposal. The relevance criteria were grouped into six major categories: the information content of the document, the source of documents, the document as a physical entity, the user's situation, the user's experience and background, and information types. Vakkari and Hakala [36] found that the relevance criteria applied when assessing references and documents relate to the problem stage during information seeking. Moreover, the findings suggest that relevance criteria for evaluating documents change more than for assessing references and this depends on how much the document content matters to the overall topic.

Although these studies investigated the use of physical libraries, focusing on people's information seeking activities when interacting with an object or another person (e.g. colleague, librarian or an expert in the field), there is little literature about how people work with different library resources (books, journals, papers) and what features, within these artefacts, determine their relevance. This study explores this particular issue: how people use artefacts in a physical library, and how elements of that understanding might be transferred to, and realised within, a digital library.

1.2 Scoping studies

Different researchers have studied different kinds of users – for example, Kuhlthau [20] studied high school students, Blandford *et al* [4] computer scientists, Covi and Kling [6] academics from various disciplines, Ellis [16] research engineers and research scientists, and Adams and Blandford [1] academics from Humanities, Computing and Business. Although Ellis *et al* [15] concluded that there are no fundamental differences in information seeking behaviour between physical and social scientists, Adams and Blandford [1] found important differences across disciplines. For example, humanities required the context and full text of original documents, whereas Business and Computing were concerned more with content (regardless of structure). Studies by Stone [31] and Watson-Boone [38] established that humanities users typically need a much wider range of resources than those in other disciplines; for example, they may need to refer to material

which is much older than that used by researchers in the sciences and social sciences, and currency is less likely to be a concern for them. They may still need to use historical material in the form of manuscripts or early printed books even if digital surrogates are available [11, 37]. Although information seeking behaviour may be very similar in different fields [15], it is not possible to generalise results beyond the boundaries of any particular study. The work reported here focuses on computers scientists.

2. Method

There are a variety of techniques that could be used to investigate people's perceptions of physical libraries and their resources. Because we wanted to gather detailed information about how people work and interact with different library artefacts, two approaches to data collection were taken: one was to interview individual library users; the other was to ask users to 'think aloud' while using the library to complete a task of their own choosing. Although both techniques are time consuming, they allowed users to articulate their perceptions and expectations related to information seeking in a library, providing rich qualitative data that formed the basis of analysis for the study.

The subject group consisted of 14 computer scientists, especially recruited for this study. Six of these were PhD students (3 – 1st year, 1 – 2nd year, and 2 – 3rd year); five were MSc students; one was a BSc student; and two were part-time tutors. Nine of these people were first observed using the library and subsequently interviewed about their use and perceptions of the library. The remaining five (four MSc students and one BSc student) were only interviewed. The subjects from the 'video and interview' group were from the Department of Computer Science. The other group of subjects, who were only interviewed, were five Electrical Engineering students on project placements working with British Telecom.

The subjects working with physical libraries were asked to look for any material relevant to their study or research, and to use the library as naturally as possible, giving their comments while looking for information and using any library resources that were available as well as appropriate for their study. Because users were recruited and given the task description prior to the

arrangement to meet in the library, users had the chance to prepare as they saw fit for their library session – the consequences of this are discussed below. The physical library sessions were carried out in one of the university libraries (this study was conducted as a multi-campus university that has several libraries, each containing potentially relevant material for this user group). The physical library sessions with users' commentaries were video recorded, and the interviews were tape recorded. Both sources of data were then transcribed and analysed by applying qualitative methods to extract the high level concepts and themes that represent activities of the users and systems.

3. Results

The study identified various activities computer scientists perform when interacting with physical information, the resources they most commonly use and how they evaluate them, and the role of librarians in their information seeking. This case study helps us to understand people's interactions and their needs in the context of physical information and librarians. Some of these requirements translate into electronic contexts although, as noted above, there are also important differences.

In the following sections, we highlight the main themes that emerged from analysis of the data. Each subject in the study has been 'named' with a letter of the alphabet; these letters are used in any direct quotations from users.

3.1 Resources used

Physical libraries consist of various resources, from traditional books and manuscripts to more advanced resources, including journals, magazines, newspapers, CD ROMs with databases of abstracts, video tapes, CDs, records, audio tapes and many more. Although the choice is extensive, people tend to use only a few of those available. For all users in this study, books were the most commonly used resource. Some of the subjects also worked with journals, magazines, and newspapers. The least popular resources were video tapes, which were made available primarily as an additional teaching aid. This finding is clearly particular to the group of users studied, being both computer scientists (who are less likely to have

resources available in video format than, for example, arts and humanities users) and researchers / students (rather that teachers, who might use such resources to support their teaching).

3.2 People's activities

There were two common strategies for finding material in the traditional library; one was to conduct a quick online catalogue search, to identify suitable resources and their locations, followed by looking for a particular book or journal within the appropriate section; the other was to work directly in an appropriate section, skipping the catalogue search as all necessary information had already been gathered using the catalogue outside the library. People following the latter strategy said that they were concerned about wasting time in the library, and they wished to check that they could find the required information before coming to the library. Such behaviour was explained by User X: 'I need to be prepared before I go to the library, so it's quicker to find information from the library'.

Other people preferred to search the library catalogue inside the library; as User Z commented, 'It gives me a chance to sit and read it here, reading at my desk it could be a bit distracting because there are other people there'. For some people, physical libraries are places of 'sanctuary' where they can find peace to read books and other resources such as journals, magazines, or newspapers, or they can simply work there quietly: as User J remarked, 'I often sit in them but don't really interact too much with them … I might sit in them for a peaceful environment to work in rather than go and use them as the resource'.

Different patterns of behaviour were observed, depending on the type of work people had to prepare. Those involved with projects tended to use journals as their main source of information, starting from the oldest issue and working through to the latest available one. However, paper journals are not widely used for many reasons, the main ones being the times information is accessible (constrained by the opening hours of physical libraries), the speed of information access and the limitation of available resources. For these reasons, the people studied now prefer to use online journals, which can be used to access

information much faster, and for which more resources (at least recent ones) are available. As User G stated, 'I can log on to the digital library and get it immediately.' Another subject (User J) commented: 'so if this was on the computer you could just stick on the bookmark and see it another time.' Time plays an important role, not only when accessing or receiving information, but also when seeking it, which very often determines what resources, physical or electronic, people use. For example, one of the subjects (User X) stated: 'If you do it physically, it will take a far longer time, I think, than looking for it on the net. Because on the net you get it immediately, the right article'.

The results also showed that people followed references recommended by their colleagues; this was highlighted by User J: 'Usually, I've got someone giving me references and I'll go and look at the journal, that particular journal and while I'm there I might browse through it'. This emphasizes the importance of interaction between people and information and resource recommendations, which increase the level of successful information seeking.

3.3 Working with the online catalogue

The online catalogue was used first (due to the task structure) for eliciting search terms and for other detailed information, which helped people to evaluate their search results. The online catalogue offered ten search keyword options, namely: 'title', 'subject', 'journal', 'name', 'global', 'series', 'videorecordings', 'reading list', 'ISBN', and 'ISSN'. The most commonly used search option was the title keyword, which involved people formulating a single term or a set of query terms, representing the topic. Although this might seem like an easy task, some of the subjects believed that finding the right keywords 'is always the problem'. This finding is in accordance with the work on information seeking and retrieval of Kuhlthau [20], Marchionini [22] and others. Other options applied include: subject keyword, global keyword, journal keyword, and author keyword. It was observed that when the system returned the message 'no matches', people changed the search option, hoping to obtain better results. This approach did not always work, in which case subjects reformulated their query terms. These findings are superficially inconsistent with

an earlier study conducted by Stelmaszewska and Blandford [30], which analysed users' behaviour in response to search results when using digital libraries. In that study, when users obtained 'no matches', they first changed query terms, then reformulated a query, then – if this strategy did not bring satisfying results – changed their preferences (search options), and finally abandoned the library. The difference between these two observations may be due to the different study conditions or users' different understanding of catalogues and DLs; further investigation is needed to gain a deeper understanding of this phenomenon.

Once results were obtained, subjects were engaged in an evaluation process which occurred on two levels; one comprised extracting from a list of results those titles including the relevant keywords; the second involved viewing a book's location, subject area, year of publication and status and, if all of these features satisfied people, noting the shelving number. If the required book was located in a different library, people generally discounted it and continued looking for another book that could be viewed locally. However, some subjects were prepared to travel to a different library to view a book or journal; for example, one user commented: 'I can go to Tottenham and have a little look'. Generally, they would not use the inter–library loan service due to the long delivery time. It appeared that people prefer to evaluate the material on the spot instead of waiting for it to be delivered.

3.4 Navigation through the library

Having established the details of documents, the next step was to locate relevant items within the library. This activity involved navigation inside the library. Navigation appears to be a natural behaviour that allows people to work on recognition of presented information rather then recall from memory [2]. In this study, navigation refers to ways people locate the appropriate sections of the library for target books and journals. They use landmark knowledge, i.e. "any features of the environment which are relatively stable and conspicuous" to make decisions about which path to take in order to obtain relevant information [8]. It was observed that the shelving numbers of books were recognized as landmarks by the subjects. Although the structure of the library

is such that books are shelved by categories, people ignore this information and use shelving numbers as a way to find the required material. This way of browsing differs from the one in virtual libraries proposed by Dumas and Plenacoste [12], where people can find material through a hierarchical organisation. Although a hierarchical organisation allows users to browse collections more freely, at the same time it demands substantial knowledge about the categorisation of subjects and topics as well as a greater cognitive load in processing such a hierarchical organisation. It is not appropriate for users who are unfamiliar with a particular subject categorisation, as this can increase search times and reduce search success.

3.5 Evaluating books

For the subjects studied, books were the most commonly used materials. It was claimed by subjects that books contain solid information. Two different approaches were applied when working with books, relating to the source of recommendations: one referred to recommendations given by a supervisor or colleague; the other referred to working without recommendations. In the case of the former, subjects did not evaluate the book, as they obeyed or trusted the source of recommendations. When they search autonomously, each of the books selected by them was evaluated. The evaluation process was initiated by viewing the title, searching for keywords related to the research topic. However, for some subjects the thickness of the book mattered as well. For example, User E said: 'if it is really thick, usually it's dealing with lots of things inside'. Conversely, User M commented: 'It doesn't have to be high level but it needs to be an introduction of HTML so I'm looking for a thin book actually'. In addition, the appearance of the book's cover was an important determinant for User K, as she remarked: 'Some books look really, really old and you think no one read them for like ten years, and then they are probably out of date, but you don't want to read it'. Having done the first step of evaluation, which involves selection of the book, the next step involved reading the summary at the back of the book to note its general content. This was followed by viewing the table of contents to learn more about it. In addition, some subjects read the introduction and preface to gain extra information about the book content, though this was not a common behaviour. To complete the evaluation process, some subjects flicked through the book, checking whether the content fulfilled their requirements. For example, User K checked that the book she chose did not contain too much coding, as she was not interested in this aspect of the topic. Another subject viewed whether enough information was included, as she noted: 'with each of the titles if this is enough or it is too much.' Some subjects also referred to a book's index, especially if they were searching for a specific term. They not only checked that the term was indexed, but they read pages including this term to evaluate how much relevant information was contained. Although subjects could view the required books in the library, some of them read the book's review on 'Amazon' prior to visiting the library, which helped them in the evaluation process. The online catalogue lacks such information, which was stated by some subjects to be very important. Reviewing books online helps them to eliminate irrelevant books and saves their time on 'empty trips' to the library.

This study shows that most subjects not only worked with a required book but they also browsed relevant sections of the library, searching for other interesting material, which increased the level of serendipity. For example, User K noted: 'I was looking for a specific title of the book so I was trying if I can find it … and then see if anything else would grab my eye that I like as well.' Most of the subjects experienced serendipity during their library session and for them it was a natural way to find material. For example, User V expressed her experience: 'Actually this one seems good. I wasn't expecting to find this'.

3.6 Working with journals

Journals were perceived as very rich sources of information. They present developments and activities occurring within different fields, helping people to keep up to date with the latest developments and use the work of others as sources of inspiration for their own work. Two approaches have been observed when people work with journals; one involves browsing the journal table of contents, looking for indicative keywords within titles and other indicators of what an article is really about; the other involves browsing individual papers within a journal trying to find links to relevant work. For the latter, the focus of

attention is not only on the title of the paper since, as one subject commented, 'by the name it's quite difficult to understand what it is about', but also on the author. Some people have preferences for particular authors and like to read their publications. There are further approaches people apply when working with journals relating to how they browse a set of journals. One referred to a chronological browsing through issues, monitoring how technology has developed over time. Another approach was to start with the latest issue then work backwards in time. This approach could provide relevant references worth following. These activities are characteristic for people working on projects where there is a demand for vast amounts of information. In other cases, people tend to apply the first approach, on just recent issues, to be updated with the latest work. These activities were supplemented by flicking through journal pages, searching for images, figures, or formulae that could help the reader in evaluating the journal. For example, one of the subjects commented: 'I think an image is a kind of things that often say a lot, maybe more figures, which are on paper'.

When flicking through or viewing journals, subjects also experienced serendipity. Sometimes the information found was not directly related to the research topic but made such an impact on the user that it diverted her from the task. For example, one user found such information and started reading it, postponing the search for a while and commenting: 'almost wish I haven't found this … because it's not what I was looking for but it is very interesting, you see. I will forget what I was looking for for a minute and have a look at this.'

3.7 Evaluating papers

Working with an individual paper is the next step when seeking information. Regardless of the selection method, people performed the same activities when evaluating it. This involved reading the abstract in the first instance to grasp an idea of the paper's content. However, the abstract alone is not always enough to make a decision about its relevance; for example, User X stated 'if I look at the abstract of the paper and if I'm not sure but I have a hint that might be something I look at the conclusion'. There are other features, including the introduction and sometimes recommendations for future work, that give extra information helping people to make decisions about the relevance of a particular paper. A references section is considered one of those. Although references are not considered to be a primary source of information, they are a crucial source that people follow in order to broaden their knowledge about the topic. One of the subjects commented: 'I pay a lot of attention to the references. Sometimes you can pick up a good source of information from it'. Other times, people viewed only a selected section within a relevant paper. For example, User M said: 'this paper is relevant but I'm looking for the methodology … so I go to the methodology description part'. In addition, people sometimes scan the paper's content, searching for pictures, figures, or formulae to make sure that it fulfils their requirements. In cases where the paper was relevant, people photocopied it for later reading.

3.8 Librarians as guides

Although, while being observed, subjects were generally autonomous library users, occasionally they reported referring to librarians for specific knowledge. They reported seeking help at the preliminary stage of their study, where the research topic as well as the library environment was new to them; in this situation, they would ask librarians for different sources of information on a specific topic and also where that information was located. However, the most common inquiry related to finding a book that was not on the shelf as specified in the catalogue. There were also other instances when people requested librarians' assistance that related to various administration issues such as: issuing a new card, proceeding with an inter-library loan request and making reservations for books. In summary, the primary role of librarians in this study was found to be guiding users on how to use specific resources of their choice, rather than providing specific support in developing people's understanding of their own needs as reported by Theng [34].

4. Discussion

There are many types of libraries, including school, academic, specialist, public and national libraries; all of them share the same property: they store and provide access to a collection of

various artefacts, which can be viewed, evaluated and used by people. These artefacts have various attributes that help people to assess their relevance. At this level of description, DLs have the same key property. However, the change of medium, from traditional to electronic, changes many details of the interaction and also the space of interaction possibilities. Some features that are central to the design of physical libraries, such as the provision of quiet study space and casual, almost incidental, low-level interactions between library users and librarians, are naturally absent from DLs, and need to be explicitly designed in if required. Conversely, DLs, particularly those accessible via the Web, are generally accessible anytime, anywhere, automatically deliver a personal copy of any item (rather than having to queue for the photocopier or borrow a book), and can offer a range of personalised services such a 'bookshelves' and alert options that are difficult to design into a traditional library. Digital library developers face a constant tension between 'evolutionary' design, which takes the familiar as a starting point and gradually evolves in response to user requirements, and 'revolutionary' design, which introduces new ways of working and creates new interaction possibilities. A risk of evolution is that exciting new possibilities are missed; conversely, a risk of revolution is that users get lost, incapacitated or alienated by their inability to achieve what used to be easy in a new environment.

The key challenge, then, in comparing users' experiences in traditional and digital libraries is to draw out the important lessons from observing users working in traditional libraries and consider whether these requirements apply in a digital environment and, if so, how they might be implemented. Here, we outline some of the lessons learned from this study, and propose a few possible solutions that at least merit further investigation. In doing this, we recognise that going from identifying requirements to developing solutions is largely a craft skill, and that any possible solution needs to be tested to ensure that it does actually fulfil requirements.

4.1 Reference recommendations

In digital libraries, users are overwhelmed with the quantity of information [4]; also, although one of the features of DLs (as compared to the Web more generally) is that material should be quality-assured, users may not make this distinction between general web quality and that of documents in a DL. In contrast, from a user's perspective, physical libraries offer quality assurance and more manageable collections. In the academic environment, it was found that people often follow reference recommendations from their colleagues, supervisors or simply from other sources of information such as papers, journals or books; this observation is consistent with Ellis' [14] work on a behavioural model of information seeking. We define 'reference recommendations' to be references to information that are recommended by colleagues or supervisors. They are shared between people and are treated as being reliable. Recommendations are important because they enable people to locate desired material more effectively and enable people to cope better with larger quantities of information. Working with recommendations cuts down the time spent on information searching and increases the level of success in retrieving relevant material. This study shows that people work with reference recommendations effectively and it was a common behaviour within the group studied.

Within traditional libraries, such recommendations generally take place outside the library context, and there is no formal structure within the library to support them. Arguably, DLs support such recommendations at least as well as traditional libraries, because people can easily send each other electronic notifications of good information sources, as well as using more traditional means of communicating recommendations. A few digital libraries provide an explicit facility to share information – whether by adding resources to an 'information space' that can be accessed by other (unspecified) users or by enabling users to alert (named) others to the existence of particular documents. However, in most digital libraries, users are unable to create lists of 'reference recommendations' that could be used effectively by other users searching on the same or similar topic. Such features, used in moderation so as not to overwhelm users, could improve users' efficiency, particularly within groups or organisations where many users have similar interests (e.g. students following the same course).

4.2 Searching, browsing and administration

Typical activities within digital libraries involve searching and browsing [5, 19], but in physical libraries these terms refer to substantially different activities. In a physical library, searching generally starts with using the online catalogue to find required material then moving purposefully through the library to locate the correct shelf and the target material, whereas browsing normally takes place entirely in the physical world and involves scanning bookshelves for relevant material. In this study, both searching and browsing were often supported by librarians. People who had difficulties in findings books simply asked librarians for help and obtained guidance on the spot. Digital libraries do not offer this kind of assistance. While there has been work over many years on understanding the skills of librarians [18], with a view to providing sophisticated on-line support, this has not yet yielded widespread workable solutions, and requires further work. However, one can imagine alternative solutions such as hybrid libraries where users having difficulty formulating queries on-line can seek support from traditional librarians, whether by telephone or email.

If a particular book was unavailable from the library, people would abandon the search for it and look for another one. This is a difficulty that digital libraries automatically address, in that if a document is (metaphorically speaking) in the DL catalogue, it should be available because numbers of copies are not restricted in the way they are in a physical library.

As noted above, many user queries to librarians are about administrative matters; in our earlier study [4] we found that many user difficulties in digital libraries were also about such matters – in particular, about authentication and access rights. For example, users were required to go to the local physical library to obtain a user name and password (e.g. Athens), or got confused about which digital libraries their user name and password entitled them to download documents from. This indicates that there is a need to provide 'digital librarian assistance' in some form, which would effectively replicate that from the physical environment. Again, there may be entirely electronic means of providing such guidance and services; in some cases, hybrid solutions may be preferable.

4.3 Evaluating search results

When people progressed through their online catalogue searching, they needed to filter relevant material from the results obtained. The features offered by the system allowed them to perform preliminary results evaluation based on information gathered from a title and, in more detail, from a 'title page', which gave links to relevant subject areas. An evaluation of an individual result influenced further activities of information seeking which involved either processing the particular material, if all criteria were fulfilled, or, alternatively, working with the next satisfactory result. Having this information not only makes people feel in control of their interaction but it also reduces the time they spend on evaluating results and the overall time spent on looking for relevant information. However, digital libraries lack this kind of information at this level of information seeking. Our earlier study [4] showed that, although some of the digital libraries provide keywords and index terms associated with a particular document, this information was not available to users from the list of results when the evaluation process began, but at the later stage of information seeking when working with an individual document. Yet again, it has to be noted that not all documents include this information so, when it is absent, the user is required to perform further document evaluation to discover whether the choice was correct or not. There is a requirement for users to be able to evaluate results earlier; this means making key, user-relevant, information about documents and their contents available at the earliest possible moment in the interaction. This relates closely to which components of a document users refer to first to assess their relevance or utility, as discussed below.

4.4 Artefacts' components

When people are working with individual documents, specific document components are very important in the decision making processes. Components such as the abstract, conclusions, introduction and references are most commonly consulted, which is partially consistent with Bishop's [3] studies. She identified six basic processes when working with a journal article: (1) read the abstract and introduction, (2) skim article headings, (3) look for and at bulleted lists, summary statements, definitions, and pictures, (4)

zero in on any particular sections that seem especially relevant (like methods, findings, directions for further research), (5) read conclusions, and (6) skim references. Physical documents generally satisfy people's demands well, because they can be rapidly flicked through, and typography, page layout and position within document provide rich cues about where each kind of information can be found. Digital documents often do not meet users' needs in this respect. Although some DLs give users access to keywords, indexes, and abstracts to help them evaluate the relevance of information, other components such as conclusions, introduction and references are not available without a complete download. In the absence of these components, within a list of results, people are forced to download a full-text document in order to evaluate the material, which can be costly in terms of the time spent downloading it and then finding appropriate components to read. In conclusion, the overall time spent on evaluating an individual document can be long compared to the same activities performed with paper material – too long for some users – which might discourage them from using digital libraries in future. This study shows that people were consistent in their way of interacting with individual documents. However, two emerging patterns of people's behaviour when evaluating the relevance of a document were observed: one involved reading an abstract, followed by introduction, conclusion, and references in this order and another one was reading an abstract, conclusion and references, omitting the introduction. Many existing studies have investigated how people read paper and online documents, focusing on speed [9, 24], comprehension [25], or activities accompanying reading such as making annotations and navigating between documents [27]. Although these studies highlight important issues and give indications of how to design better reading technologies, none of them has explored which components of a document are crucial for successful document evaluation, as investigated in this study.

One suggestion to improve the efficiency of digital library use could be to change the ways a document can be accessed and provide alternative examination methods - making the relevant artefacts such as abstract, conclusion, introduction and references available at the user's 'finger tip'. This would allow users to view a particular section of the document before committing themselves to downloading and evaluating the whole article. The effectiveness and usability of such a solution needs further investigation.

4.5 Serendipity

Another issue that arose with physical libraries is that of serendipity. Information seeking for both physical and digital libraries is based on users' full or partial knowledge about what information they seek. There is evidence [17, 32] that people also acquire information through serendipitous interaction. Roberts [28] defines serendipity as an unexpected, accidental discovery of interesting information, while Ross [29] presents it as 'finding without seeking'. Serendipity occurs frequently in the physical environment and has been associated by Marchionini and Shneiderman [21] with a browsing strategy when seeking information. Serendipitous interaction in physical libraries has been supported through generations simply by the juxtaposition of books on a bookshelf. Blandford *et al* [4] found that serendipitous discoveries were the main source of really positive user experiences when working with DLs. The question, then, is how to actively promote serendipity in interactions between users and DLs without making it harder for users to achieve their explicit information seeking goals. One possible approach is to think of DLs as stimulating curiosity, encouraging users to explore them in the quest of knowledge, so that accidental or serendipitous discoveries become common.

One way of facilitating serendipitous discoveries would be to alert users to similar or related articles to those being searched for, increasing their chances to immerse themselves in discovery. This is not a new idea: various search engines (e.g. Google) and some digital libraries (e.g. ACM) already employ such a mechanism, but it has not yet been widely adopted by DL developers. Recommender systems, as implemented by some online bookstores, can also raise users' awareness of related materials. If such a feature is implemented in an appropriate way, it could also support serendipity. The effects of these features on user experience with DLs needs to be studied – not by giving users predefined tasks, but by studying user experience in context, preferably over an extended period of time.

4.6 Familiarity

Finally, an important activity found in this study was becoming familiar with the environment. People need to acquire knowledge about the library content, especially in cases when they are new to the environment or to the topic they are researching, and about where the required information can be found. In physical libraries, the construction of the environment, including the layout of furniture, stacks, etc. and the physical appearance of different areas of the library give strong (if culturally dependent) cues about the high level organisation of the library. For example, a quick look will make it clear to most users which area contains journals and how they are organised, or where the reference books are located. More detailed familiarisation is also supported by librarians, whose specialised knowledge was tested by some subjects seeking information. Typically, for the studied group, interaction between librarians and subjects occurred within the library. In digital libraries, users are left alone to explore the features, content and structures, and their successful acquisition of documents depends on how well collections are designed. Virtual libraries may try to create a familiar environment analogous to the one found in physical libraries [12] to make people 'feel at home' and reduce the level of 'lostness' in a digital world [33]. We can identify a significant role for librarians as being 'information therapists' [34] who help people to make better use of library resources and cut the time required for such activities which has not been yet reproduced in DLs. This highlights users' need for extra help and support in the form of appropriate feedback, as identified by Stelmaszewska and Blandford [30]. One approach we have recently investigated, with preliminary results that look promising but are not yet well tested, has been to introduce two types of 'tips' to a digital library (the NZDL [39]). The first type of tip gives general but detailed information about how to perform a quick search or choose appropriate terms and explains different types of search; each explanation is supported by examples. The second type of tip is context-dependent, delivered as part of the system response to a user's query, including an explanation about what happened during the search, and suggesting ways to improve the search results.

If a user is to become familiar with a system, that means it must be easy to learn – an issue addressed directly within the HCI literature. Dix *et al* [10] propose five aspects of system design that improve learnability. The first is predictability – that the user should be able to anticipate what next state is reached by performing any particular action from the current state; in the case of DLs, total predictability is clearly unattainable (even undesirable), because the user is searching for novel material. However, an element of predictability – e.g. that the user can be reasonably confident that they are 'heading in the right direction' while browsing – will help users' familiarisation.

A second aspect is what they term synthesisability – that the current state should be understandable on the basis of the history of the interaction. In our earlier study [4], we identified many situations where the user clearly had no idea how they had arrived at the current position. For example, one user manually stemmed terms in a digital library that was designed to accept full words only, and consequently received 'no matches'. Synthesisability depends on high quality feedback – particularly about the detailed effects of actions.

The third aspect discussed by Dix *et al* is what they term familiarity: that the user should be able to relate their experience of working with the system to their real world knowledge. As discussed above, DLs exploit the real-world library metaphor; indeed, the theme of this paper is understanding how well real world familiarity can be exploited effectively in the design of DLs. Therefore, we will not consider this aspect further here, for fear of the paper becoming recursive.

The fourth aspect they discuss is generalisability – that the user should be able to generalise experience from one system to another. Most DLs are accessible via web browsers, and therefore inherit many generalised features, such as means of navigation, from the browser environment. However, in many other respects they perform poorly on this criterion. As users move from one DL to another, they have to learn new features, new structures, and new ways of interacting, as well as needing to (necessarily!) familiarise themselves with new content. As developers are exploring new possibilities and novel features, this lack of generalisability may be unavoidable, but as DLs become widely used standard systems, there is a real need for standardisation – not just of

underlying protocols, but also of the features and interaction styles that users experience.

The final aspect of learnability discussed by Dix *et al* is consistency – that similar functions should be implemented in similar ways throughout the system and that similar feedback should be given in similar situations.

We have discussed familiarity at length because this aspect of DLs can draw on more general HCI understanding in a way that most of the other particular features discussed cannot, since they are specific to the libraries domain and users' experiences of information seeking.

5. Conclusions

What emerges from this study is an identification of the activities people perform in physical libraries, how they interact with different resources, what resources they use, what components of individual resources are relevant, and how people evaluate them. As highlighted, physical and digital libraries have different phenomenologies: the fundamental nature of the user experience is necessarily different because the physical properties of shelves, paper, books and other tangible media are different from those of keyboards, screens, and other devices that support interaction with digital information. Similarly, the interactions between library users and librarians are experienced in different ways in the different types of library. Thus, the library metaphor is limited in terms of the extent to which user experiences can transfer from the physical to the digital environment. However, we have tried to highlight deeper user requirements and how they are realised within physical environments.

We have investigated some roles of librarians in a traditional library and discussed how these roles might be realised within a digital environment. This study has also highlighted key features of computer scientists' generation and utilisation of information while information seeking. Although these findings cannot be generalised to all types of users, they identify some key requirements on the design of DLs. As discussed earlier, although digital libraries provide access to large document collections, users experience problems with finding particular information in them. DL developers should not be trying to create DLs that mimic physical libraries, but to develop new designs that empower users, both in terms of transferring their existing skills into the new environment and also to develop extended system capabilities, such as "the ability to support rich forms of searching and link following from within documents; to annotate documents in a fluid manner and to manage those annotations; to share documents – or perhaps annotated portions of documents – as a means of collaborating with peers and experts" [23].

Another requirement is better support for serendipitous information seeking which, although it demands a relatively unorthodox and ill-defined approach to information seeking, could attract more users to DLs.

Finally, it is our hope that insights from this study can be used by designers to create digital library environments that people feel more comfortable and confident in using.

6. Acknowledgements

7. References

1. Adams, A. and Blandford, A. (2002) Digital libraries in academia: challenges and changes. In E-P Lim, S. Foo, C. Khoo, H. Chen, E. Fox, S. Urs and T. Costantino (Eds.) *Digital Libraries: People, Knowledge, and Technology*. 392 - 403. Springer LNCS 2555.

2. Anderson, J. R. (1990). *Cognitive psychology and its implications*. San Francisco: W.H. Freeman.

3. Bishop, A. P. (1998). Digital Libraries and Knowledge Disaggregation: The Use of Journal Article Components. *Proceedings of ACM DL'98,* 29 – 39.

4. Blandford, A., Stelmaszewska, H. and Bryan-Kinns, N. (2001) Use of multiple digital libraries: a case study. In *Proc. JCDL 2001*. 179 – 188. ACM Press.

5. Borgman, C. L. (2000) *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked world*. The MIT Press.

6. Covi, L., Kling R. (1997) Organisational dimensions of effective digital library use: Closed rational and open natural system model. In Kiesler, S (Ed) *Culture of the Internet*. Lawrence Erlbaum Associates, New Jersey. 343 - 360.

7. Crabtree, A., Twidale, M., O'Brien, J and Nichols, M. (1997) Talking in the library: Implications for the design of digital libraries. *Proceedings of DL'97*, ACM Press. 221 - 228.

8. Dillon A., McKnight, C., and Richardson, J. (1993). Space – the final chapter or why physical representations are not semantic intentions. In C. McKnight, A. Dillon, and J. Richardson (Eds.) (1993) *Hypertext: A psychological perspective.* 169 – 191. New York: Ellis Horwood.

9. Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, Vol. 35 (10), 1297 - 1326.

10. Dix, A., Finlay, J., Abowd, G. & Beale, R. (1998) *Human Computer Interaction.* Prentice Hall International. 2nd Edition.

11. Duff, W. M. and Johnson, C. A. (2002) 'Accidentally found on purpose: Information-seeking behavior of historians in archives', *Library Quarterly*, 72: 472-496.

12. Dumas, C. and Plenacoste, P. (2002) Libraries: comparisons between the real and the virtual in 3D, 2D zoomable and 2D arborescent. In A. Blandford and G. Buchanan (Eds.) *Proc. JCDL Workshop on Usability*. 11-14.

13. Duncker, E. (2002) Cross-Cultural Usability of the Library Metaphor. In *Proceedings of the ACM-IEEE Joint Conference on Digital Libraries*. New York, NY: ACM. 223-230.

14. Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, 45 (3) 171 - 212.

15. Ellis, D., Cox. D. and Hall. K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, 49 (4) 356 - 369.

16. Ellis, D. and Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, 53 (4). 384 - 403.

17. Ford, N (1999) Information retrieval and creativity: towards support for the original thinker. *Journal of Documentation*, 55 (5). 528 - 542.

18. Galliers, J. R. (1991) Cooperative interaction as strategic belief revision. In Deen, S. M., Ed. *Cooperating Knowledge Based Systems,* Berlin: Springer-Verlag.

19. Jones, S., McInnes, S. and Staveley, M. (1999). A Graphical User Interface for Boolean Query Specification. *International Journal on Digital Libraries*, Vol. 2 Issue 2/3, 207 - 223.

20. Kuhlthau, C. (1988) Longitudinal case studies of the information search process of users in libraries. *Library and Information Science Research*. 10.3. 257-304.

21. Marchionini, G. and Shneiderman, B. (1988) Finding facts versus browsing knowledge in hypertext systems. *IEEE Computing*, 70 – 79.

22. Marchionini, G. (1995) *Information seeking in electronic environments*. Cambridge University Press.

23. Marshall, C. C, Golovchinsky, G. and Price M. N. (2001). Digital Libraries and Mobility. *Communications of the ACM*. 44.5. 55-56.

24. Muter, P and Maurutto, P. (1991). Reading and skimming from computer screens: the paperless office revisited. *Behaviour and Information Technology*, 10 (4). 257 - 266.

25. Muter, P., Latermouille, S.A., Treunit, W.C., and Beam, P. (1982). Extended reading of continuous text on television screens. *Human Factors*, 24, 501 - 508.

26. Nicols, D. M., Twidale, M. B., and Paice, C. D. *Recommendation and Usage in the Digital Library*. http://www.comp.lancs.ac.uk/computing/research/cseg/97_rep.html.

27. O'Hara, K and Sellen, A. (1997) A comparison of Reading Paper and On-line Documents. *Proceedings of CHI'97*, ACM Press. 335 - 342.

28. Roberts, R. M. (1989). *Serendipity: Accidental Discoveries in Science.* John Wiley and Sons, Inc.

29. Ross, C. S. (1999). Finding without seeking: the information encounter in the context of reading for pleasure. *Information Processing and Management,* 35. 783 – 799.

30. Stelmaszewska, H. and Blandford, A. (2002) Patterns of interactions: user behaviour in response to search results. In A. Blandford and G. Buchanan (Eds.) *Proc. JCDL Workshop on Usability*. 29 - 32.

31 Stone, S. (1982) Humanities Scholars-Information needs and uses. *Journal of Documentation*. 38 (4). 292-313.

32 Toms, EG (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human Computer Studies*, 52. 423 - 452.

33. Theng, Y. L., Mohd-Nasir, N. and Thimbleby, H. (2000) Purpose and Usability of digital libraries. *Proceedings of the 5th ACM Conference on Digital Libraries*. 238 – 239.

34. Theng, Y. L. (2002) Information therapy in Digital Libraries. In E-P Lim, S. Foo, C. Khoo, H. Chen, E. Fox, S. Urs and T. Costantino (Eds.) *Digital Libraries: People, Knowledge, and Technology*. 452 – 464. Springer LNCS 2555.

35. Twidale, M. B. and Nichols, D. M. (1998) Using studies of collaborative activity in physical environments to inform the design of digital libraries. Presented at the *CSCW'98 Workshop Collaborative and co-operative information seeking in digital information environments*. (Also as Technical Report CSEG/11/98, Computing Department, Lancaster University).

36. Vakkari, P. and Hakala, N. (2000) Changes in Relevance Criteria and Problem Stages in Task Performance. *Journal of Documentation*, Vol. 56, No. 5, 540 – 562.

37. Warwick C and Carty, C. (2001) Only Connect, a Study of the Problems caused by platform specificity and researcher isolation in humanities computing. In Arved Hubler, Peter Linde and John W.T. Smith, (eds.) Electronic Publishing 01, 2001 in the digital Publishing Odyssey. *Proceedings of the 5th International ICCC/IFIP Conference on Electronic Publishing*. 36-47.

38. Watson-Boone, R. (1994) The Information Needs and Habits of Humanities Scholars, *Reference Quarterly*, 34(2), 203-16.

39. Witten, I. H., Nevill-Manning, C.G., McNab, R. & Cunningham, S.J. (1998) A public digital library based on full-text retrieval: collections and experience, *Communications of the ACM*, 41(4). 71–75.