Online Information Access Systems in the Construction Industry

Sameh M. Shaaban[1] and John M. Kamara[2]

ABSTRACT | The growth and widespread use of the Internet and World Wide Web (WWW) has led to increased dependence on online information resources in the construction industry, and with it, the problem of information overload. This has also led to the development of various Online Information Access Systems (OIAS) such as search engines and domain specific web portals, to ensure the retrieval of relevant information from the WWW. This paper provides an overview of current OIAS in the construction industry. It discusses the issues relating to OIAS and uses the results from a detailed study of information seeking patterns of industry professionals to assess their effectiveness. The paper concludes with suggestions on how Online Information Access Systems can be improved using Information Visualisation (IV) techniques.

KEYWORDS | construction industry, information seeking, information visualisation, online information access systems

1 Introduction

Information has a crucial role in the delivery of construction projects. High levels of rework, faulty design, and project delays can be attributed to the lack of ready access to relevant information. The implications of not finding sought information have been considered by a number of studies. For example, Lockley et al. [1] identified high levels of rework in the design and design-development phases and argued that insufficient information leads to faulty decisions that have to be revisited. Another study by Abdel Meguid [2] on delay-related claims revealed that in nearly 60% of the cases judged in court the unavailability and lack of access to information was mentioned as a "claims' causation factor."

Vast amounts and different kinds of information are utilised on construction projects, mostly for decision-making. This information is obtained mainly from internal sources (i.e. from within participants involved in a project) and from external sources (e.g. technical information on building regulations and product information from manufacturers). External sources and varied and may have unique contents, organisation and presentation styles and language.

1.1 Traditional Information Resources

Traditional resources include the two categories of resources that provide information from published documents and information based on private or public centres and educational institutes [3]. Information from

1. Architectural Informatics Group, School of Architecture, planning and landscape, University of Newcastle-upon-Tyne NE1 TRU {j.m.kamara@nd.ac.uk}
2. RIGAE Informatics, The Old Post Office, St Nicholas Street, Newcastle-upon-Tyne, NE1 {sameh.shaaban@theNBS.com}

published documents includes a wide variety of different resources including periodicals, official advisory documents, and trade literature. Periodicals include journals, newsletters, bulletins and newspapers. They are the types of publications issued in parts. Periodicals are considered the most widely used published source of information in construction [4]. They are the main reference for current design practice, which provide primarily topical information about the profession, the economy, building design, and technical development [5]. Official Advisory Documents cover building codes, standards and regulations. They include information related to obligatory requirements established by official or professional agencies regarding the practice of the profession. Professionals consult these reference sources to check specific requirements for their design decisions. Trade literature deals with sales information that introduces products and suggests applications and technical literature that explains detailed use and performance of products, e.g. Sweet's Catalogue Files. Information based on Private or Public Centres and Educational Institutions include that from building and research centres, universities, colleges, and schools of architecture. Building Centres are market-oriented organizations providing permanent or temporary exhibitions of materials and products, and related documents and information services. Research centres and academic institutions are more research-oriented organizations, offering research facilities and specialized libraries for the use of professionals, academicians and students.

Almost all traditional information resources, with the partial exception of information from participants of the design processes, have to do with printed, written or drawn documents. These are considered to be easy to browse through page by page, can be touched, and in most cases, they do not require any additional hardware (apart from computer-based media). However, these traditional resources cannot cope with the overwhelming increase in the industry knowledge base. It is also difficult for industry users to search through this information thoroughly. In other words, users can almost never be sure whether they have obtained all the information needed, even if a considerable amount of time is invested in reviewing this material.

1.2 Online Information Resources

The limitations with traditional information sources, and the growth and widespread use of the Internet and World Wide Web (WWW) have led to the development of online information resources. These refer to construction/project related information stored electronically on distributed databases and accessible via the WWW. Electronic forms of doing business have pushed information providers to adopt new methods of dissemination (especially of information from external sources). These allow for the storage of significantly larger amounts of information than traditional sources. They also provide quicker and rigorous methods for finding relevant information through search facilities to information repositories:

The rapid development of the Internet has introduced new horizons to the application of, what used to be called, 'internal resources'. New concepts of doing business across wide area networks have been introduced to the construction industry. Concepts of "virtual enterprises" and "project extranets" have also emerged [6]. Not surprisingly, there has been a relative increase in the use of the Internet when compared to other sources of information (Table 1) [7, 8].

This increase in usage of online information resources for products information has led to a corresponding increase in users' expectations of these systems with respect to the quality and depth of provided information and the effectiveness of the search capability [7, 8]. Another consequence of the increased use of online information is that of information overload and accessibility [9, 10]. The design and capabilities of online information access systems (OIAS) is therefore crucial to the effective retrieval of online information.

Table 1: Usage of information sources for products information [7].

Source of Information	2004	2002	2000
Manufacturers literature	87%	95%	93%
In-house library	80%	92%	94%
Internet	79%	73%	49%
CD-ROMs	40%	46%	44%
Other information from manufacturers	N/A	31%	N/A
Magazines and journals	N/A	1%	7%

1.3 Objectives and scope of Paper

The objective of this paper therefore, is to provide an assessment of OIAS with respect to the way they respond to the information seeking needs of users. Following an overview of OIAS, issues related to their design, development, and accessibility to information held by OIAS (information seeking) are identified. A detailed study on the information seeking patterns of construction professionals, which was based on a particular OIAS, the Construction Information Service (CIS), is then presented. The findings from this study are used to assess the appropriateness of OIAS. The paper concludes with suggestions on how OIAS can be improved through Information Visualisation (IV) techniques.

2 Online Information Access for construction professionals

Online Information Access Systems (OIAS) refer to those 'systems' which are designed to improve access to relevant information from the millions of documents stored on the Internet. They include sophisticated search engines (e.g. Google) and domain specific portals (e.g. Construction Information Service – CIS). OIAS are becoming increasingly popular sources of information in the construction industry. Surveys show the high increase in the use of online resources among the industry professionals as a source of information [7, 11]. A major advantage of OIAS, as reported by user surveys, is that they provide a medium for rapid and up-to-date industry information.

2.1 Information Seeking

A key element of OIAS is the ease of access to stored information, with respect to seeking and retrieval of information. Marchionini [12] defines information seeking as: *"a process in which humans purposefully engage in order to change their state of knowledge"*. Information seeking is a fundamental human process closely related to learning and problem solving. While seeking is characterised as a more human oriented and open ended process, retrieval implies that the object must have been known at some point; most often had been previously organised for later use. Seeking connotes the process of acquiring knowledge. It is more problem oriented as the solution may or may not be found. Retrieval is applicable to database management and most applied problems, but seeking is closer to answering questions or learning and requires high level cognitive processes. Figure 1 illustrates the relationships between learning, information seeking, information retrieval, and both analytical and browsing strategies. Analytical strategies are concerned with searching mechanisms, query building, and Boolean operations [13]. Browsing strategies are more focused on navigating document structures and hierarchical classifications [14]. Much of information seeking may require identifying and retrieving previously stored information. Thus, information retrieval is one type of information seeking except when it is conducted by a machine, as machines cannot seek information but can retrieve it. Shneiderman [15] advocates that information seeking tasks are primarily decomposed into browsing or searching. These tasks are represented by interface actions, where each of these key strategies is accompanied with certain strategies and tactics [12]. Information seeking needs [16] include the following:

- The need to browse product and material choices, and see them in combinations. Where proper visual representations should be adopted.
- Whereas selecting a product represents a commitment, designers may need to keep options open and keep informed with similar products information.
- The need to retrieve information in different abstraction form, e.g. from a generic element, to individual instantiation of an element that appear in the documentation schedules.
- Information seeking is often based on precedents. Practitioners commonly develop "palettes" of favourite products that they use repeatedly on different projects. This case-based aspect of product selection introduces the need for storage and management as well as searching previously used products.

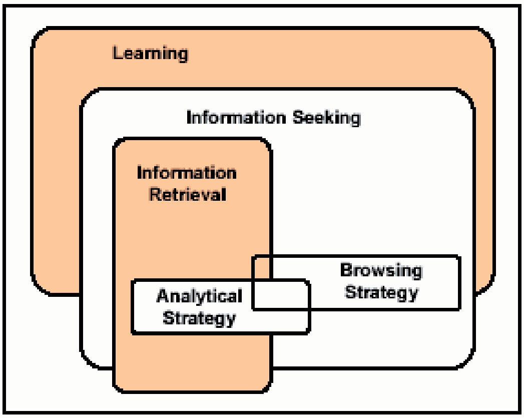

Figure 1: The relationships among learning, seeking and retrieval (after [12])

Although information seeking is crucial to the effectiveness of OIAS, surveys on ICT usage in the building industry [17, 18] reveal that there is an evident reluctance among industry users in adopting advanced ICT systems for information seeking. On the other hand, development interest and business investment are still directed towards computer aided design (CAD) solutions. This is because practitioners do not see information seeking as a productive activity, and there is therefore a lack of information seeking skills among industry professionals [19].

2.2 Factors affecting the development of OIAS

Rich and structured published information is, inevitably, the foundation for successful information delivery. It facilitates the design of efficient online information access systems. Research efforts have identified several factors that influence information delivery as well as the development of online construction information sources [9, 20-23. Some of these are described below.

Representation of information. Authoring information is a key determinant to the design of information resources often referred to as data models or underlying data structures. Construction industry research is caught in the middle between two directions of data modelling. *Core models* are often called product and process models. They model deep knowledge areas but still difficult to be applied to large data sets [24-26]. On the other hand, *classification models* model shallow knowledge areas and are mostly applied to wide data sets [27, 28]. The diversity of building standards and regulations and the variety of international constraints has diverted the aim of developing unified models that could represent the construction knowledge base. Consequently, none of these models has yet been implemented successfully for current information resources. Classification models adopt some simple conceptual models for representing entities to develop extensive categorization or classification breakdowns for the purpose of providing classification systems or enumerating all of the specific elements of construction. The diversity of building standards and regulations as well as the variety of international environmental constraints has diverted the aim of developing a unified classification model to more national based models. Examples are: the UNICLASS in UK [27], the MASTERFORMAT in USA and Canada [29], and the BSAB in Sweden [28]. The main drawback of these systems is that they are designed for shelving and filing systems of paper-based documents, with minor emphasis on the document contents itself. They model shallow knowledge areas and are applied to wide data sets.

User-centric development of information access systems. The swift development of online environments has been accompanied by a rather sceptical and slow change of information seeking culture among industry users. Where traditional information resources are deemed to be easy to use and navigate through its paper form, users have to acquire new searching skills. The domain independent developments in search engines have contributed to the users' lack of enthusiasm. It has resulted in overloading users with information that could be irrelevant to the construction domain. It also ignores users' information needs and domain implicit information searching tasks. Hence, it expects users to learn advanced query languages in order to find the bits of information they are looking for.

Infrastructure. Infrastructure is a major concern when developing on-line information resources. The speed of connection as well as allocated band width affects the performance of any information system. The emerging online technologies for improving information access systems, e.g. information visualisation or raster image detection, are striving to comply with minimum connection configuration [30].

Market pressure. This is a factor that directly influences the development of information resources in terms of providing funding and incentive for improvement. For example, it is acknowledged that users are not willing to pay for building products information and trade literature. This has led manufacturers to publish their own biased information on the internet. The opposite scenario is emerging regarding technical document information resources. The increasing demand by users of such systems in the UK has resulted in competition to provide better services, backed by scientific research, between information systems, e.g. Construction Information Services [31] and Construction Expert [32].

2.3 Problems of current OIAS

Table 2 shows a comparison between the three types of online construction information resources for scientific literature, product information, and technical documents. Seven (OIA system) parameters (derived from the OIAS development considerations outlined above) are used to provide a comparison between these resources. They are: the information provider, the information broker, data representation, consumer, accessibility behaviour, investment flow, and users' perceived required system enhancements. It also includes addresses of the examples mentioned in the previous discussion. The goal of this comparative analysis is to summarise the pros and cons of the current systems as well as to identify existing problems.

An analysis of the different profiles of the three types of information resources listed in Table 2 suggests the following:

1. All construction information documents are multi-authored. They neither share common format nor adhere to a unified industry-wide classification system. This phenomenon has lead to the formation of large repositories of unstructured data that information resources need to interrogate. The end results are less accurate search results and, consequently, more time and effort to find relevant information from the users' part.
2. Construction industry professionals are the main user of all information resources. However, there is no evidence of their involvement in the design of these systems. Furthermore, there are no studies that investigate their patterns of information seeking behaviours and/or usage of these systems.
3. The "Pull" approaches are the most common accessibility method applied among online resources. On the other hand "Push" approached are mostly used in product information as a mechanism where manufacturers can promote their

Table 2: Comparison between forms of online construction information resources

Parameter	Scientific literature	Products information	Technical documents
Information Provider	Academics Industry researchers	Manufacturers	Multi-disciplinary Multi-author
Data structure	Un-structured documents Un-structured data sets No unified classification system	Un-structured documents Some structured data sets (paper-based classifications)	Un-structured documents Un-structured data sets Efforts to some industry classification systems
Information broker	Publishers Digital Libraries	Information houses e-business providers	Information houses e-business providers
Information consumer	Professionals, Engineers, Researchers, In non-university environments	Professionals	Professionals Educational institutes
Accessibility behaviour	Information retrieval Mostly “Pull” Minor “Push”, e.g. email alerts	Mainly push as manufacturers promote their wares Legacy of printed catalogues “Pull” mechanisms are also offered	Information retrieval Mostly “Pull”
Investment flow	Largely paid by users Find it too expensive Increasing pressure on publishers to reduce the prices, if not to offer free access Less development budgets	Basically Manufacturers Some goes to brokers, in case of portals or product directories Manufacturers have more say over development Increasing development budgets	Paid by users Too expensive for small organizations
Required system enhancements by users	Better searching mechanism Full text searching instead of only metadata More push approaches	In depth information Information about performance, cost, and visual representation Better searching mechanism Less time consuming	Better searching mechanism

wares. However, users have reported that product catalogues contain much un-needed marketing information and they would rather have access to technical data.

4. Evidence of information assimilation problems has also been noted. Surveys reported the need for in-depth information, e.g. cost and performance, rather than merely metadata searching, i.e. searching titles and abstracts only. The need for supporting different types of information objects was also reported, e.g. visual representations.
5. There is an obvious need for enhancing searching mechanisms as more information is added to these systems daily, which reportedly caused information overload problems. Less time consuming and less demanding user interfaces improvements were clearly emphasised by users.

The above limitations suggest that current OIAS need to be improved to better incorporate the needs of users. A first step in achieving this is to develop an understanding of the information seeking behaviour of construction industry professionals and assess how these are reflected in the design of OIAS. The next section describes a study that addresses this need.

3 Information seeking patterns of construction professionals

Information seeking involves an indeterminate sequence of events [12]. Identifying patterns of usage behaviours and seeking tasks can be achieved by analysing these events in three levels of abstraction: seeking strategies, tactics and user moves – depending on the kind of study being conducted. Seeking strategies are the conscious approaches that information seekers take to a problem. They comprise of sets of tactical tasks, sometimes in a particular order, which are applied to fulfil a particular information need situation [12]. Tactics are discrete intellectual choices that users make during an information seeking episode (e.g. narrowing the search space by selecting a date range – [33]). User moves refer to discrete actions such as clicking a mouse or downloading a document, and are evidences of tactics. A focus on user moves [34], and to some extent, user tactics [35] is more suitable for exploratory ('what is?') studies. For explanatory ('why?') studies, the focus should be on the more abstract levels of seeking behaviour, i.e. seeking strategies and tactics. However, whatever the focus of the enquiry, information behaviour research strategies can either be "intrusive" or "unobtrusive" [36]. Intrusive approaches (e.g. participant observation, interviews, questionnaires and experiments) are most effective in stimulating evidences of the cognitive aspects of the information seeking process, which could include pre-search and post search activities. Limitations associated with this approach include:

- Sampling includes demographic considerations formed in subjective manners.
- Users behave differently in controlled environments.
- User feedback is hindered by users' perception of their seeking behaviours and existing technologies [37].
- Users are required to subjectively assess and comment on their tactics and strategies, which is least useful due to the unconscious nature of these seeking behaviours [12].
- Results are less reliable due to difficulties inviting large enough representative samples.

Unobtrusive approaches are those that do not intrude into the phenomenon being studied. They concentrate on evidence found by the examination of assertive traces of human behaviour. Examples of the use of unobtrusive methods include Leydesdorff's study on citation analysis [38], Julien and Duggan's content analysis of the information behaviour literature [39], and Chen and Cooper's investigation of usage patterns of online libraries [34]. A popular method used in unobtrusive studies is archival analysis [34, 38, 40]. This is based on collecting data about the user interactions with a particular system while performing information seeking processes. By using transaction log records as the data gathering method, researchers are able to examine much larger user samples where data about the users' interaction represents their real life experience. Whilst unobtrusive approaches overcome most of the problems indicated by the intrusive methods, several problems have been reported. These include:

- Difficulty to gain access to logged records of user interaction data, due to security considerations.
- Results can only be interpreted subjectively since it does not provide insight into participants' perspectives or cognitive reasoning.
- Ethical considerations, where not all gathered data could be used in order to ensure users' privacy.

3.1 Research Environment

The research reported in this paper adopted an unobtrusive approach, and investigated seeking strategies, tactics and user moves. The main data source was the transaction log records of Construction Information Service (CIS) (http://www.tionestop.com/ihsti/index.html) (Figure 2).

CIS is an online information access system that contains over 24,000 technical documents in its collection [31]. It is the product of a joint venture between Technical Indexes, the Royal Institute of British Architects

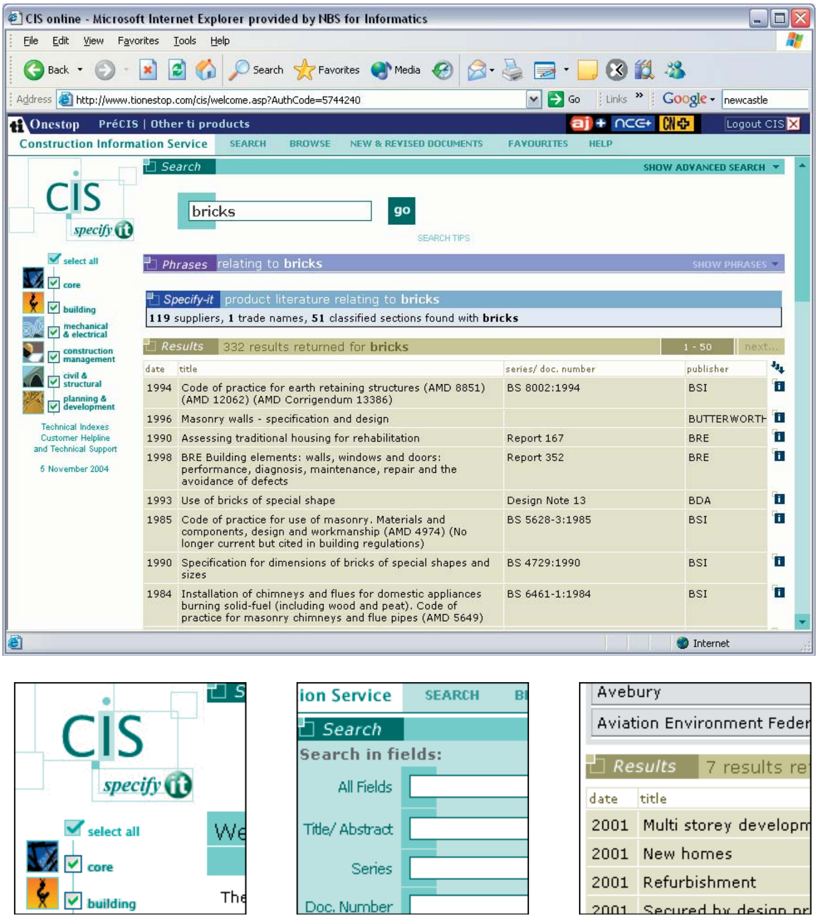

Figure 2: Images of CIS [41]

(RIBA), and National Building Specifications (NBS). CIS is a subscription-based system that provides continuously updated technical information to its users, e.g. building regulations approved documents and British Standards. Information is provided in sections so that the wide variety of users that access the service can "mix and match to create the information resource they need" [41]. Users can search for information by document number, document title, series or issuing body [41]. The reasons for choosing CIS were as follows:

- CIS is one of the key construction information sources in the UK, capturing a large percentage of users. Hence, it is appropriate to indicate common behaviours.
- It was possible to have access to CIS server log records, since one of the authors was (and still is) a member of the CIS design and development team.

3.2 Research Methodology

Two datasets were collected over two different periods with a total of 104,224 sessions over the first half of the year 2001 (the "2001 dataset") and 86,678 sessions over January and February of 2002 (the "2002 dataset"). The first dataset was used in the main analysis and the second for validation. Log data typically consists of electronic records of the user's requests and the system's response. Each log record contains a number of parameters describing the user's state, i.e. which page he/she has accessed and the page he/she is coming from including all the parameters describing the server request. Hence each line of the transaction logs represents a particular user move, e.g. searching, moving from a results' page to another or to download a document. In order to investigate user moves on CIS, all possible user states were mapped into 21 states, and these were further into eight higher level states [42], which are: 'start session', 'simple search', 'advanced search', 'explore results', 'browse index', 'get help', 'open document', and 'end session'.

The AWK programming language [43] was used to compile the session vectors or characteristics (i.e. unique identifier, count of covered identical sessions, session length in seconds, total count of user moves, etc.) from the raw log records. The AWK language transforms the raw log files into vector files, which are in the form of text files that contain comma-delimited data. These were then imported into Microsoft Access database tables to facilitate statistical analysis of the data. In addition to descriptive analysis of log data, cluster analysis was performed in order to find whether there are clusters of common information seeking patterns among the user samples. A three-phase clustering process, similar to a framework introduced by Chen and Cooper [34] was used. In the first phase, the 2001 session vectors were transformed into clusters using a 2-stage methodology. In the second phase, the 2002 session vectors were transformed into clusters using the same methodology. In the third phase, the 2002 session vectors were classified into the clusters derived from the 2001 clustering in phase 1. The purpose from phase 2 and 3 was to validate that the resultant clusters from both 2001 and 2002 datasets are not formed by chance. This will be achieved by comparing the two resultant clusters from both phases. According to Everitt [44], if the clusters found were 'real' and not merely artefacts of the particular technique being used, they should not be altered by performing the clustering on different sub-sets of the data.

3.3 Key Findings

The key findings of the research are summarised below. A more comprehensive description is provided in Shaaban et al. [45].

Information seeking strategies: Analysis of the two datasets showed that 'exploring results' and 'simple searching' were the most popular activities among all users. The results also showed that 76% of search queries over the whole period of the study were a maximum of two terms long. This suggested a very strong preference for simple searching over advanced searching.

Clusters of information seeking behaviour: Four clusters of information seeking behaviour were identified from the clustering exercise. Table 3 shows description of the resulted clusters using variables derived from user moves, and Table 4 summarises the seeking behaviours of each group. The variables are grouped into four groups according to their nature. These groups are:

1. *Session characteristics:* includes the total number of sessions and their means.
2. *User activities:* includes sums of user moves accessing each part of the CIS online system.
3. *Download decision:* includes variables that indicate what triggered the user to make the download decision.
4. *Measures:* includes three session performance measures: the ratio of successful information seeking activities, ratio of system difficulty and ratio of seeking activities per download.

The clusters of information seeking behaviour (Table 3) were as follows:

G1. Exploratory information seeking behaviour with highly interactive system usage, 16.8% of total sessions with average length of 43 minutes.

G2. Knowledgeable users of the domain that are lacking experience of system usage, 68.6% of total sessions with average length of 16 minutes.

Table 3: Description of the four behaviour groups

Clusters					
Var.	**Description**	G1	G2	G3	G4
(1) Session characteristics					
Ns	Number of sessions	9277	37822	6326	1728
Np	Percentage of number of sessions	16.82	68.58	11.47	3.13
Sm	Mean of Session length in minutes	43	16	5	2
(2) User activities					
B	Simple search	12.52	26.23	7.92	9.75
C	Advanced search	5.21	6.01	28.09	2.80
D	Explore results	45.93	22.79	16.28	15.75
E	Browse index	3.81	3.63	6.42	43.30
F	Help	0.23	0.49	0.91	0.48
Gn	Open non-copyrighted document	8.40	8.63	4.36	5.55
H	Read document copyright	10.02	12.50	7.74	6.28
Gc	Open copyrighted documents	9.59	12.02	7.28	5.87
G	Open document	17.99	20.65	11.64	11.42
Ah	Users hitting home buttons	2.76	4.73	16.85	6.38
(3) Download decision					
Gs	Open document after searching	6.21	13.96	7.16	3.21
Gb	Open document after browsing	0.70	0.49	0.66	3.88
Ge	Open document after exploring results	11.52	6.68	4.27	4.75
(4) Measures					
Rs	Ratio of successful seeking (100x)	83.52	57.57	27.43	20.45
Rf	Ratio of system difficulty (100x)	0.35	0.83	1.54	0.67
Rd	Ratio of seeking activities per downloads	2.36	0.77	0.46	0.19

G3. Known-item searching accompanied by help-intensive behaviour, 11.5% of total sessions with average length of 5 minutes.

G4. Passive users who experience unsuccessful short seeking episodes, 3% of total sessions with average length of 2 minutes.

G2 represents a dominant behaviour among the other groups (Table 4). Its pattern features, in terms of making download decisions from the first page of results as well as minimum utilisation of system features, coincides with findings from similar studies by Chen and Cooper [34]. It also suggests a notable influence from the Web searching, using generic Web search engines like AltaVista and Google, on the information seeking behaviour of the construction industry users in their work-related information seeking and retrieval.

Although there are time constraints imposed on construction industry users, especially when asked to perform an information seeking task, they tend to seek information in an exploratory manner. This is evident from comparing groups G2 and G1. The more time they spend (G1) on the system the more exploratory behaviour they perform.

There is clear evidence that the industry users lack information searching skills, which is clear from their short query searches and rare advanced searching. This finding highlights the importance of devising new mechanisms that provide the industry professional with the means to recognize and improve their information skills. It also emphasises the need for enhancing systems with better user interfaces.

Table 4: Summary of the groups of seeking behaviours

Group description		Session characteristics	Most frequent pattern	Pattern features
G1	Exploratory information seeking behaviour with highly interactive system usage.	16.8% of total sessions with average length of 43 minutes.	Explore results Explore results Explore results Explore results Simple search Download document	Rarely go back to the home page. Explore the search results extensively before making a download decision. Highest successful seeking rates. Lowest difficulty reported. Significantly more time spent in the seeking process before the download decision.
G2	Knowledgeable users of the domain that are lacking experi-ence of system usage.	68.6% of total sessions with average length of 16 minutes.	Simple search Simple search Explore results Explore results Download document Download document	Make a download decision on the first page of the search results. Lack of information skills in terms of manipulating the system interface. Use minimum system features. Almost a download for each search.
G3	Known-item searching accompanied by help-intensive behaviour.	11.5% of total sessions with average length of 5 minutes.	Advanced search Advanced search Explore results Download document Start a new search	Mostly known-item-searching. Fewer download decisions. Either unsuccessful searching or the documents do not exist. High percentage of resetting the system to start a new search. The most help-intensive sessions.
G4	Passive users who experience unsuccessful short seeking episodes	3% of total sessions with average length of 2 minutes.	Browse index Browse index Browse index Browse index Explore result Download document Simple search	Short sessions. Highest browsing index activities The lowest successful seeking rate. Mostly unsuccessful seeking episodes.

By mapping the resultant groups of user behaviours to the task-action model of seeking strategies [15], higher tendencies of adopting fact finding strategies, rather than open ended browsing can be recognised. Patterns of exploration of availability behaviour were also found. No evidence of open ended browsing was found. Table 5 shows the mapping between the study's groups of information seeking behaviours and Shneiderman's [15] task-actions taxonomy of seeking strategies.

Table 5: Mapping of behavioural groups into strategic task actions

Task-action (Shneiderman [15])	Usage behaviour groups
Specific fact finding (known-item search)	G2, G3
Extended fact finding	G1
Open ended browsing	---
Exploration of availability	G3, G4

4 Comparing user behaviour with OIAS features

The study on information seeking behaviour shows that groups G1 and G2 represent the majority of usage patterns, with 80.4% of the total number of sessions (16.82% and 68.58% respectively). Hence, evidences of system usage on the two groups are indicative of general usability of the OIAS. By studying the usage patterns of both groups, represented in Tables 3 and 4, several observations can be made:

- Whilst the system provided users with three groups of information access features that enabled them to find information, i.e. simple searching, advanced searching, and browsing the documents index, users preferred using only the simple searching feature. Advanced searching features, which include, among other features, specific field searching and publishing date range, were seldom used. Similarly, browsing the documents index, including browsing publisher names, document series, and subject classifications, were rarely accessed.
- Short, one or two term, queries were the most commonly used search formulation. Whilst the system allowed for Boolean searching, which was clearly described and explained on the help pages, they were not of particular interest among users.
- The most popular set of features used, were the exploring results group. This included searching within results, navigating to the next page of results, sorting results, displaying detailed result's information, and displaying a publisher's details. A significant increase of usage of exploring results was indicated in situations where users had more time for their information seeking, as opposed to no increase or even reduction of usage of other features.
- The help feature, which would provide assistance to users on how to use the system features, were seldom used in both behaviour groups.
- These observations suggest that the OIAS, investigated in this study, was not being used effectively or to its maximum potentials by the industry users. This is to further reinforce the point that current online systems are not used effectively. Whether the reasons behind this lack of system utilisation were lack of information seeking skills, user interface design, or unfamiliarity of the system, users did not even put an effort to find out how to use the system more efficiently. This was evident from the lack of using the help features.

4.1 Suggestions for Improving OIAS through Information Visualisation

It is observed that information seeking behaviour of construction professionals identified in the study are broadly in agreement with generic observations of OIAS mentioned earlier. This further suggests that improvements to OIAS are necessary, if users are to derive optimum benefit for such systems. One way in which this could be done is through the use of Information Visualisation (IV) techniques.

IV refers to the use of computer-supported, interactive, visual representations of abstract data to amplify cognition [46]. The objectives of IV are: (a) to enable the understanding of large amounts of data; (b) to allow the perception of emergent properties; (c) to enable identifying problems with the data and with its method of collection; (d) to facilitate the understanding of the relationships within the data features; and (d) to facilitate hypothesis information. The challenges for the design of successful IV tools include the following:

- ***Adopting task-based design*** to develop realistic tools, providing users with new features that truly facilitate real difficult information tasks and be a major factor for their acceptance and utilisation. However, rather than designing a full system replacement for information seeking, task-based visual components would be more effective.
- ***Providing undemanding information workspace,*** in terms of minimising the need of complex user abilities and skills. Building visualisations that depend on stable user abilities, such as spatial abilities, helps to gain better user performance

as well as reducing the influence of individual differences. However, sensitive selection of a spatial display metaphor is crucial to provide intuitive interface for users to understand and interact.

- ***Minimising the learning cost,*** in terms of learning time and effort to accomplish information seeking tasks. Systems, which are harder to get familiar with, tend to perform less successfully and require more training. Educating users on understanding advanced visualisation types or advanced searching algorithms is a crucial factor that influences the development of future IV tools.
- ***Selecting appropriate display type.*** This is mostly influenced by the complexity of the task rather than the composition of the display itself. It is therefore necessary to build task taxonomy for information seeking.
- Carefully utilising ***the screen real-estate*** in order to balance the number of features included, as well as the amount of data that can be visualised. On the other hand, increasing the screen size beyond the display size, where users have to scroll around in order to see the hidden parts, is not advisable. The clarity and structure of the IV user interface is an important factor for its success.

Given the IV principles outlined above and the findings from the study of information seeking behaviours of construction professionals, the following suggestions need to be considered for the improvement of construction domain OIAS such as CIS:

- The provision of alternative information seeking techniques as opposed to the unpopular advanced searching tools.
- Enhancement of results exploration facilities within a system. This will assist users in exploring the contents of the library that is relevant to their search queries.
- The provision of undemanding information work-spaces that require less information searching skills, and which suits the abilities of construction users.
- The incorporation of an intuitive visual interface, which would gain industry professionals' interest.

The above recommendations have been incorporated in the development of a prototype visual interface (VisExplorer) [47]. The VisExplorer user interface (Figure 3) provides a brief overview of its key features. It is divided into seven main sections: search panel (Fig 3a), documents explorer window (Fig 3 b), document details window (Fig 3c), download button (Fig 3d), axis selection panel (Fig 3e), list of subjects found (Fig 3f) and list of publishers found (Fig 3g).

The *search panel* includes the search input box, the submit search button and the number of maximum results returned selector. Here, users can enter their search terms, submit the search as well as setting the number of search results that they want returned per page.

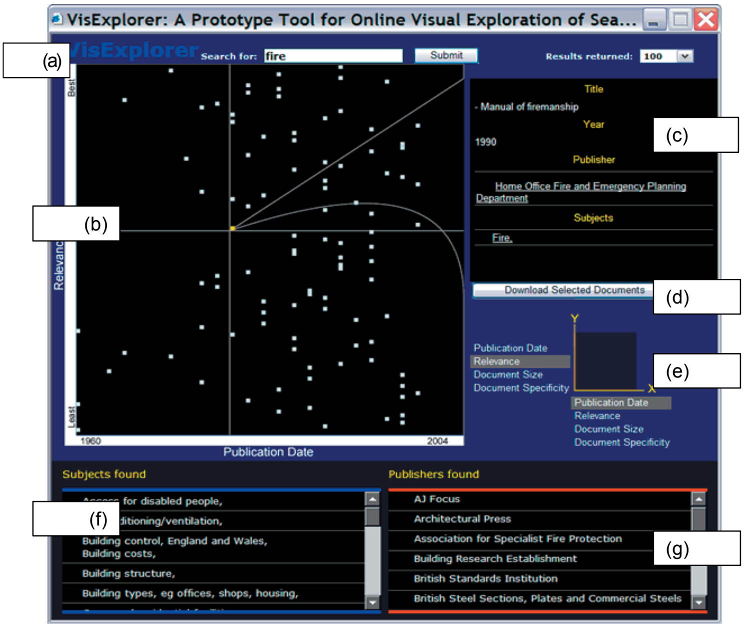

Figure 3: VisExplorer User Interface

The tool implements a one entry point to the search engine through one search box. As it allows users to start their information seeking session by adopting only simple searching, it corresponds to the findings of the study's user behavioural analysis. The *documents explorer window* is the main results area. It adopts a scatter plot visual representation. Each document returned in the results is represented by a point on the scatter plot. A square white box was used as its visual mark. The default axis variables are publication date for the X axis and Relevance for the Y axis. Range boundaries for each variable are stated on the axes bar, e.g. Least and Best for the Relevance axis. The *document details window* provides users with a short summary of a highlighted or selected document on the document explorer window. The summary includes the document's colloquial title, year of publication, publisher name and subjects under which the document is classified. Users can highlight or select a document from the explorer window by hovering over its point or actively doing a mouse click on it respectively. The *download button* provides the document download functionality. As the button caption indicates, the user can download one or a set of documents depending on the number of documents selected on the explorer window. The *axis selection panel* depicts a two-axes graph layout. It enables the user to select alternative variables to apply to any of the scatter plot's axes, i.e. X and Y axes. The user can choose one of four variables, which are: Publishing date, Relevance, Document size and Document specificity. Selecting a different variable causes the document collection on the explorer window to relocate to the respective axis values. The two panels, which list the *subjects* and *publisher names* found in the retrieved results, are located in the bottom part of the user interface. Ordered alphabetically, they provide an overview of the search results. Subjects list provides an over view of relevant subjects to the search terms, retrieved from the results' classified subjects. Publisher names list depicts a breakdown of the publishers of these current results. A full description of the development and use of VisExplorer, which is beyond the scope of this paper, is provided elsewhere in [47].

Conclusions

This paper set out to provide an overview and an assessment of Online Information Access Systems (OIAS) with respect to the needs of users. It was established that a key factor to the effective use of such systems is the information seeking skills of users, which affect the way they browse and search for information. A detailed study on the information seeking behaviour of users of Construction Information Service (CIS), using server log records, was described. The findings from this study revealed that most users performed simple searches and rarely used the advanced search features of CIS. The study also revealed that the exploration of search results is the dominant information seeking activity. It further revealed that industry users lack information searching skills as evidenced from their short searches and rare use of advanced search facilities.

It should be noted however, that since this study was based on one OIAS (CIS), the wider application of the findings to other OIAS and types of users need to done with caution. It is further acknowledged that the fast development of the Internet, and increased dependency on it for information sources, might affect the accuracy of the findings after a few years. Further and continuing research is clearly required, but the indications are that current OIAS need improvement to better serve the needs of users. Ways in which this could be achieved include: the enhancement of results exploration facilities within a system, through Information Visualisation (IV) techniques, and the incorporation of features for finding similar documents. There is, of course, the dilemma with OIAS, given that such systems have to be developed first and extensively used before your can effectively study seeking patterns [48]. This findings and recommendations highlighted in this paper therefore serve as an important contribution to the continued refinement of Online Information Access Systems in the construction industry.

REFERENCES

[1] Lockley, S.R., Wiltshire, J., & Dudek, S. (1987). Information Transfer and Knowledge Based Systems in Architectural Design, *ARECDAO, 1st international symposium of computer aided design in architecture and civil engineering.* Barcelona, Spain.
[2] Abdel Meguid, T. (1997). Managed Claims Procurement Strategy (MCPS): a Comparative Study of the Performance of Alternate Building Procurement Strategies, *Faculté de l'aménagement* (p. 181). Montréal: Université de Montréal.
[3] Adderley, P. and Nixon, M. (eds.) (1994) CIIG Manual: a guide to good practice in construction industry information, *Construction Industry Information Group (CIIG)*
[4] Goodey, J., & Matthew, K. (1971). Architects and information. *York. Institute of Advanced Architectural Studies.*
[5] Marvin, H. (1985). Information and experience in architectural design, Research paper. 23. XY/N-1: Building Research Establishment
[6] CISEMIC (2002). eSM@ART and CISEMIC, *The European Conference on Information and Communication Technology Advances and Innovation in the Knowledge Society.* Salford, UK
[7] RIBA-Enterprises (2004). The NOP Information Sources Survey 2004. London: RIBA Enterprises
[8] Barbour-Index (2004). Construction expert, http://www.barbour-index.co.uk/content/aboutservice/ce.asp.
[9] Amor, R., Langham, M., Fortmann, J., & Bloomfield, D. (1996). The UK industry knowledge base feasibility study. *CIB W78 and TG10 workshop: Construction on the Information Highway.*
[10] Shaaban, S., Lockley, S., & Elkadi, H. (2001). Information Visualisation for the Architectural Practice. In E. Banissi,K. Khosrowshahi,M. Sarfraz, & A. Ursyn (Eds.), *The 5th International conference of Information Visualisation (IV'01)* pp. 43-50). London: IEEE.
[11] Barbour (2000). Influencing Product Decisions: specification and beyond. London: Barbour Index plc
[12] Marchionini, G. (1995). *Information seeking in electronic environments*: Cambridge university press
[13] Yates, B., & Neto, R. (1999). *Modern Information Retrieval* New York: ACM press
[14] Hearst, M.A. (1999). User Interfaces and Visualization. In Ricardo Baeza - Yates, & Berthier Ribeiro - Neto (Eds.), *Modern Information Retrieval* (pp. 257-323). New York: ACM press
[15] Shneiderman, B. (1998). *Designing the User Interface, Strategies for Effective Human-Computer Interaction*: Addison Wesley Longman,Inc.
[16] Coyne, R., Lee, J., Duncan, D., & Ofluoglu, S. (1998). Applying Web-Based Product Libraries. *Proc. EuropIA'98: Cyberdesign, ed. Cherif Branki, K. Zreik, Paris, EuropIA productions, Paris,* 105-117
[17] Doherty, J. (1997). A Survey of Computer Use in the New Zealand Building and Construction. *ITcon,* 2
[18] Hansen, K., Gann, D., & Groak, S. (1998). Information technology decision support and business process change in the USA. *Engineering Construction and Architectural Management,* 5(2), 115-126
[19] Davidson, C. (1998). The Information Problem in Building - Causes and Effects, first Edition, Montréal, May: http://www.cibat.qc.ca/if/va/ifpubdex.htm.
[20] Kelly, J., Aouad, G., Rezqui, Y., & Crofts, J. (1997). Information systems developments in the UK construction industry. *Automation in Construction,* 6, 17-22
[21] Augenbroe, G. (1998). Building product Information Technology, Executive White Paper: Construction Research Center, Georgia Institute of Technology.
[22] Lockley, L., Watson, R., & Shaaban, S. (2002). Managing E-Commerce in Construction - Revolution or E-Business as Usual? *Engineering Construction and Architectural Management,* 9(3)
[23] Björk, C., & Turk, Z. (2000). A survey of the impact of the internet on scientific publishing in construction it and construction management. *ITcon, 5,* 73-86.
[24] Froese, T. (1996). Models of Construction Process Information. *Journal of Computing in Civil Engineering, American Society of Civil Engineers,* 10(3), 183-193.
[25] Tolman, F. (1999). Product modelling standards for the building and construction industry: past, present and future. *Automation in Construction,* 8, 227-235
[26] Eastmen, C. (1999). *Building product models: Computer environments supporting design and construction:* CRC Press LLC
[27] Crawford, M., Cann, J., & O'Leary, R. (1997). *Uniclass: Unified Classification for the Construction Industry:* Royal Institute of British Architects

[28] Ekholm, A., Häggström, L., & Karlsson, H. (2000). *BSAB 96, the Swedish construction industry classification system:* Swedish Building Centre
[29] CSI (1995). *MasterFormat 1995 Edition:* The Construction Specifications Institute, (http://www.csinet.org)
[30] Barbour (2002). Exploring the Web as an Information Tool: Barbour Index plc.
[31] McKechnie, J., Shaaban, S., & Lockley, L. (2002). Computer Assisted Processing of Large Unstructured Document Sets: A Case Study in the Construction Industry. In ACM (Ed.), *The ACM Symposium on Document Engineering in conjunction with the Tenth International Conference on Information and Knowledge Management (ACM CIKM).* Atlanta
[32] Barbour (1999). The Sourcing and Exchange of Information: across the building team: Barbour Index plc
[33] Ellis, D., & Haugan, M. (1997). Modelling the Information Seeking Patterns of Engineers and Research Scientists in an Industrial Environment. *Journal of Documentation,* 53(4), 384-403.
[34] Chen, H.-M., & Cooper, M.D. (2001). Using clustering techniques to detect usage patterns in a Web-based information system. *Journal of the American Society for Information Science and Technology,* 52(11), 888-904
[35] Sonnenwald, D., & Wildemuth, B. (2001). Investigating Information Seeking Behaviour Using the Concept of Information Horizons: Technical Reports -- School of Information and Library Science, University of North Carolina - Chapel Hill
[36] Case, D.O. (2002). *Looking for Information: A survey of research on information seeking needs and behaviour* USA: Elsevier Science
[37] Höök, K. (2000). Steps to take Before Intelligent User Interfaces Become Real. *Journal of Interaction with Computers,* 12(4), 409-426
[38] Leydesdorff, L. (1998). Theories of citation? *Scientometrics,* 43, 5-25
[39] Julien, H., & L., D. (2000). A longitudinal analysis of the information needs and uses literature. *Library & Information Science Research,* 22, 291-309
[40] Cothey, V. (2002). A longitudinal study of world wide web users' information-searching behaviour. *Journal of the American Society for Information Science and Technology,* 53(2), 67-78.
[41] RIBA (2004) RIBA.IHSti Construction Information Service (http://www..thenbs.com/products/cis.asp).
[42] Chen, H.-M., & Cooper, M.D. (2002). Stochastic modelling of usage patterns in a web-based information system. *Journal of the American Society for Information Science and Technology,* 53(7), 536-548.
[43] Aho, A. V., Kerngham, B. W. and Weinberger, P. J. (1988), The AWK Programming Language, Addison-Wesley (ISBN 0-201-07981-X) (http://cm.bell-labs.com/cm/cs/awkbook/).
[44] Everitt, B.S. (1993). *Cluster Analysis, Third Ed.:* Edward Arnold, a division of Hodder & Stoughton Ltd.
[45] Shaaban, S., McKechnie, J. and Lockley, S. (2003), "Modelling information seeking behaviour of AEC processionals on online technical information resources," *Journal of Information Technology in Construction (ITCon),* Vol. 8, pp. 265-281
[46] Card, S., Mackinlay, J., & Shneiderman, B. (1999). Readings in Information Visualisation: Using Vision to Think San Francisco: Morgan Kaufmann Inc
[47] Shaaban, S. (2004), "A user-centric approach for the development and evaluation of a visual interface to online information in the construction industry", PhD Thesis, University of Newcastle upon Tyne, UK
[48] Amor, R., Betts, M., Coetzee, G. and Sexton, M. (2002), "Information technology for Construction: Recent work and future directions," *ITcon,* Vol. 7. pp. 245-258