INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN
ICED 05 MELBOURNE, AUGUST 15 – 18, 2005

UNDERSTANDING THE SEARCH FOR INFORMATION IN THE AEROSPACE DOMAIN

Francisco del-Rey-Chamorro, Ken Wallace

Keywords: information access, information retrieval, knowledge management

Abstract

This research was triggered by a concern expressed by some experienced engineering designers about the usability of new electronic retrieval systems. These systems are substituting the traditional paper-based manuals, but there is no clear understanding about how these new systems alter the way designers search for information. In fact, there is little understanding of how designers search for information in company manuals. In total, twelve designers from five company sites of an aerospace company were observed undertaking the design case of a flying control surface, accessing information only from the company's manuals. Six of the designers were provided with paper copies of the manuals, whereas the other six were asked to access the standards and manuals through the electronic system. The results presented in this paper provide a model of how designers search for information in company manuals that may help computer scientists to create more aligned retrieval tools to designers.

1 Introduction

Aerospace products are of a remarkable complexity. Designers require knowledge and information ranging from structures and materials to manufacturing processes. The breath and depth of this knowledge make it unfeasible for a designer or a team of designers to develop products solely based on their retained experience. Designers, therefore, need to seek information through the resources available to them, such as other designers, company manuals or previous projects. Marsh , in 1997, reported that 12% of a typical designer's time is spent seeking and retrieving information [1]. However, this percentage is likely to increase as technology evolves and the complexity of aircraft increases.

The way in which designers in the aerospace industry obtain information is changing [2]. Traditionally, young designers would join a company as apprentices and leave at the retiring age. During their careers, these designers would become extremely knowledgeable and a valuable source of information for the next generation of designers. Thus, a master-apprentice relationship would be established and knowledge retained in the company. However, career development and job opportunities encourage many to change companies and to transfer out of the design process. This represents not only an invaluable loss of experience and knowledge in itself, but also interrupts the knowledge transfer to the next generation. If experienced designers are not available in the future, designers will have to retrieve that information from what remains in the company. At present these are essentially records.

The aerospace industry has recognised the importance of information management and has invested heavily in the development of new information systems. The introduction of such systems has had an immediate effect on the way information is handled within companies. For example, the use of email and the Intranet decreased the use of traditional means of communication, such as telephone and face-to-face communication [3]. File sharing facilities have enabled the development of virtual workplaces. Electronic databases allow the creation and storage of electronic copies of technical manuals and the old paper copies to be withdrawn [4].

The investment in information technologies is difficult to justify if designers are not able to access and retrieve electronic information effectively. Designers are obliged to access and retrieve information in a completely different environment, in which previous searching practices may no longer be supported. The migration of technical manuals from paper format to electronic format can be a burden on the way designers search for information.

Thus, there is an increasing need to understand the process of retrieving and using information in the aerospace domain. Based upon an empirical study, this paper aims to provide a theoretical model of the process of accessing and retrieving information from technical documents.

2 Characteristics of the access and retrieval of information

Research undertaken in the engineering design domain provides a better understanding of the behaviour of designers and the characteristics of the access and retrieval of information. New technologies are providing a platform that enables documentation to be easily updated, but some designers feel they are not benefiting from these new technologies.

The retrieval of information is one type of design episode embedded within the design process [5]. Information retrieval is divided into two stages [6]: (1) information-seeking; and (2) information-searching. Information-seeking is referred in this paper as information access and information-searching as information retrieval. Information access is driven by users, and takes place when designers need to identify and locate promising sources of information. For example, a consumer may have just bought a new video player, but does not understand how to use the remote control. There might be three promising sources of information: the video catalogue, the manufacturer's help-desk, and the manufacturer's website. The information access would be driven by the abilities and preferences of each consumer, and would be the process of identifying and accessing one of those promising sources, e.g. finding the help-desk telephone number and getting through to a customer service representative.

The information retrieval stage is driven by the type of source, and takes place when designers need to find a piece of information within a promising source. For example, the previous consumer obtaining the information from the help-desk [7]. The information retrieval would be driven by the type of source that the consumer is accessing, e.g. automated help-desk or customer service representative, and the retrieval process would be the dialogue established between the consumer and the customer service representative.

2.1 Information access

The access of information is user driven. Factors such as company culture and personality also influence the selection of an information source [8]. For example, extrovert designers may prefer to ask colleagues for their information needs. Introvert designers, on the other hand, may prefer to access company manuals.

Previous experiences also have an important role in information access. Designers do not only use memory to retrieve pieces of design information from previous experiences, but also information about information, i.e., meta-information [9]. Some designers can remember in which manual and where in that manual they can find information on a particular design topic [4]. For example, a designer may not know what the density of titanium is, but may well know precisely where to find that information. Designers also use 'vague memory' to retrieve the approximate location of where more unstructured information may be found. For example, it may happen that a designer does not know precisely where a piece of information is but remembers that this information was used in a previous project.

In the particular case of accessing information from documents, Vakkari and Nakala [10] also add the source of documents and documents as physical entities as important factors for the access of information. When alternative electronic sources are available to users, the success of electronic sources depends on their perceived accessibility, and user's assessment of strengths and weaknesses of networked services compared to alternative media [11].

The perceived accessibility is therefore a critical factor once designers are introduced to a new electronic system. If systems are not perceived as usable, designers are reluctant to switch from traditional sources to electronic systems. This can be partially explained by lack of computing skills and the need for training [12]. Positive experiences can be reinforced if users perceive that their search skills improve, and urgent tasks can be performed more rapidly.

Previous experiences also have an important role in information access. Designers do not only use memory to retrieve pieces of design information from previous experiences, but also information about information, i.e., meta-information [9]. Some designers can remember in which manual and where in that manual they can find information on a particular design topic [4]. For example, a designer may not know what the density of titanium is, but may well know precisely where to find that information. Designers also use 'vague memory' to retrieve the approximate location of where more unstructured information may be found. For example, it may happen that a designer does not know precisely where a piece of information is but remembers that this information was used in a previous project.

2.2 Information retrieval

The retrieval of information is source dependent. How the information is searched for within a particular source depends on the type of source. For example, designers use different information retrieval strategies when they are obtaining information from a colleague and from a database.

Kuhlthau described the retrieval of information in three steps: exploration, collection, and presentation [13]. Information retrieval strategies have been researched mainly in the computer science domain. Xie identified the strategies used by students retrieving

information from a library database [14]. He identified some patterns between information retrieval strategies and the purposes of searches: identifying, learning, finding, accessing, locating, evaluating, keeping, and obtaining. This work suggested a way to provide users with a useful tool that suited their needs, but it is still at a very early stage of development. His research was based on retrieving information from library catalogues, which have a very uniform structure compared to the engineering domain. The records in the engineering domain are not as consistent as in library catalogues. Thus, his findings may not be applicable to the engineering domain.

Some researches believe that there is a need not only to understand the process of retrieving information itself, but also the context in which information needs arise. Sheldrick Ross investigated the information encountered in reading for pleasure by book readers, and identified why readers found books of interest to them [15]. Five reasons were identified for why readers retrieved information that was different from that which they were looking for: (1) Awakening new perspectives, (2) Enlargement of possibilities, (3) Models for identity, and (4) Reassurance.

3 Research methodology

An empirical study in the aerospace industry was carried out in order to understand the retrieval and use of documentation. The research methodology was based on an experiment based on a realistic design case study using the think-aloud protocol. Such experiments have been successfully used in the past to study information management in engineering design [5]. The use of a single design case allowed the control of some factors influencing the process, and allowed the researchers to focus on the main research issues.

The design case study was based on a military aircraft flight control surface (FCS) (see Figure 1) and prepared by an experienced designer within the company. Twelve experienced designers were given a scheme of the FCS and a set of suggested topics. The designers were only allowed to search for information in the company manuals. Six of them were provided a paper version of the company manuals, whereas the other six were asked to access an electronic version of the same manuals using the company's Intranet. Six sources of information were identified:

- Indexes
- Documents
- Search-engine results
- Memories
- Current project
- Previous projects.

Indexes were systematic lists of manuals, chapters or section titles providing a cross-reference to the location of each of them. Manuals gathered together information that was considered to be interrelated by the authors of the manuals.

Documents were the core information repository in both systems. The length of the documents accessed during the experiments varied between 3 and 93 pages. The documents were divided into sections and subsections and some of them had their own indexes.

Search-engine results were computer-generated lists of cross-references to chapters and sections of the manuals identified by the search engine. The main difference between indexes and search-engine results was that indexes were human generated, whereas search-engine results were computer-generated. This can lead to different retrieval performances depending on the quality of the indexes and search engines.

Memories contain information from previous experiences stored in the long-term memories of designers [5]. Sometimes the retrieval of information from manuals triggered further related information from previous experiences. The amount of information retrieved from memories was, as would be expected, dependent on the range of previous experience of the designers.

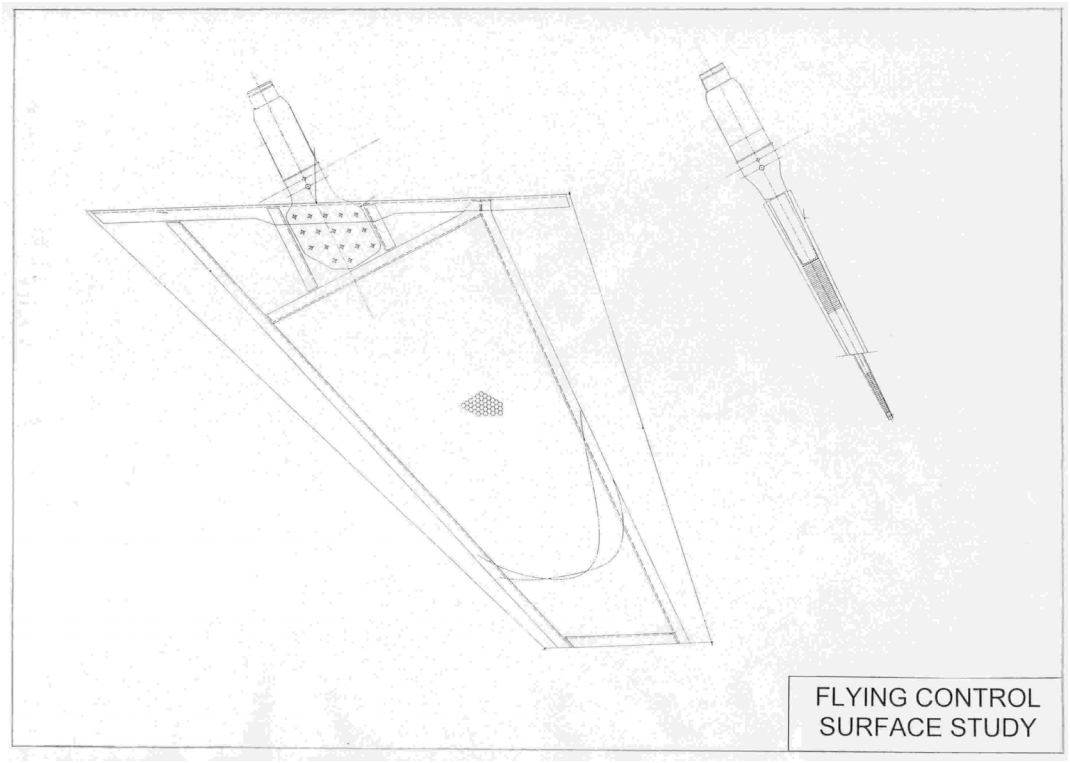

Figure 1. Flying Control Surface (FCS)

Current project documents provided some of the information associated with the current design task, e.g., dimensions, pre-selected materials, and expected performance. This source of information provided designers with background information that helped them to search for information more accurately.

Additionally some designers brought documents and drawings from previous projects that dealt with similar issues. It was observed that designers did not use parts of previous designs, but they used them to become aware of the issues and the documents relating to a subject. For example, one designer was dealing with the sealing of the honeycomb structure in the FCS. He remembered that the rudder of a previous aircraft was also made of honeycomb. He then decided to access a past drawing of the rudder and retrieve the standard numbers of documents related to the sealing of honeycomb structures from that drawing.

The search of information took place in the form of search episodes [5]. Episodes are chunks of transcripts. Search episodes present the process of acquiring information to satisfy

information needs. Each search episode is divided into segments. Segments are individual pieces of information retrieved by designers. For example:

> *This index is telling me now that the information on aluminium materials is contained in the XXX document.*

Segments were distinguished in the transcripts by the change in the content of the information provided [16][17].

4 The search for information in the engineering domain

The search for information is commonly known as the process that starts with the identification of the information need and ends with the use of the information (see Figure 2). The analysis of the experiments has shown that this process can be divided into three stages:

1. Define information need
2. Access information
3. Retrieve information

Designers define information needs when their background knowledge is insufficient to continue their designs further. Information needs therefore identify the pieces of information that designers require to progress their designs. These pieces of information are determined by the previous knowledge of the topic possessed by designers.

Once designers have identified an information need, they access information using the most promising source of information. The selection of the source is based on their previous experience of accessing sources of information. The following example shows an example of one designer who was familiar with the manuals and was trying to find a type of aluminium that met the design requirements.

> *I usually work with these types of materials. I know for a fact that they are in XXX85 manual. What I will do is I will go to that manual and I will scan through the types of aluminium available looking for one the meet these requirements.*

However, a less experienced designer may select a different strategy and source to satisfy the same information need.

I am new to this domain. I do not usually work at this level of detail. I am trying to find a suitable aluminium alloy. I do not know where I can find this type of information. I think what I will do is to go to the alphabetical index and navigate through the relevant documents.

Designers retrieve information from documents. Documents contain large amounts of information, but designers are generally only interested in specific pieces of information within those documents. It was observed that designers undertook different retrieval strategies depending on their level of knowledge of the topic. Consider, for example, a designer exploring a new topic when he or she has very little prior knowledge of the topic. The designer will aim to discover relevant pieces of information. However, a designer who is trying to find a specific piece of information, such as the maximum stress of an aluminium alloy, has generally a better idea of what he or she is looking for. The designer is therefore

able to verbalise the information need and is able to search for related keywords in the manual.

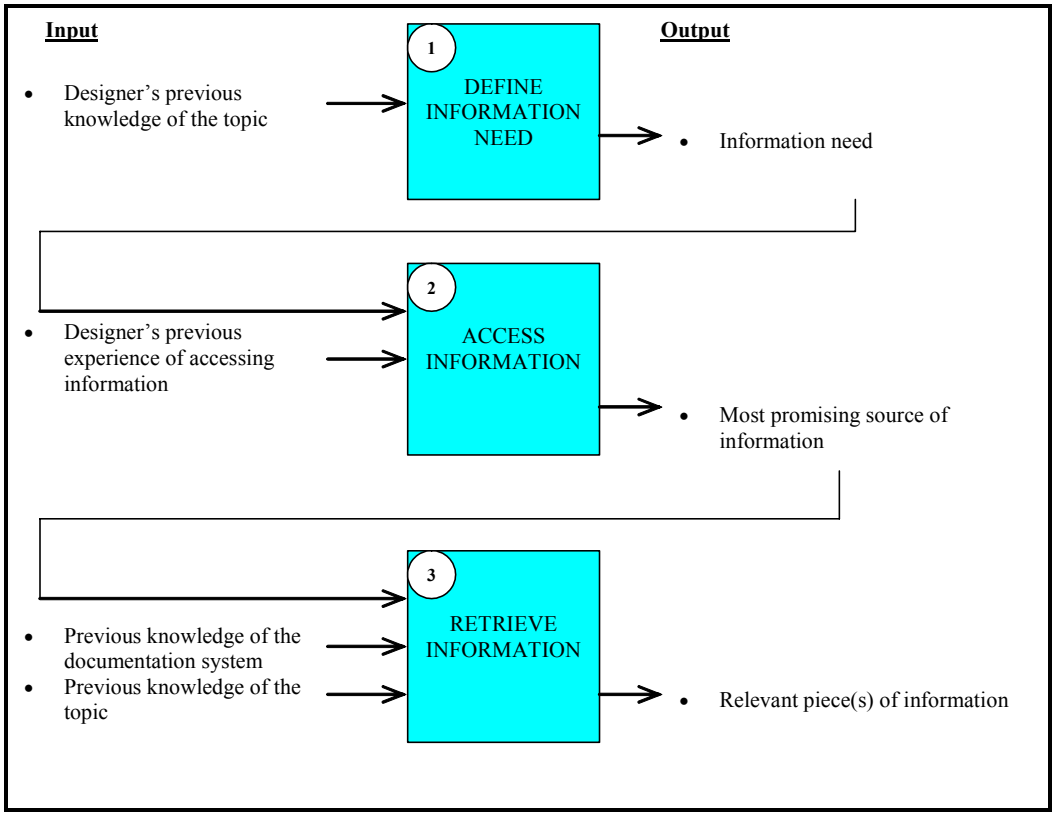

Figure 2. The search for information

4.1 Define information need

Information needs are pieces of information that designers require to progress their designs. They can be classified according to the previous knowledge that designers had on the topic. Three levels of detail were identified:

- Exploring a topic
- Understanding issues
- Defining product specifications.

Designers needed to explore a topic when they had very little previous background knowledge of that topic. The main characteristic of this type of information need was that the designers did not know what sort of information would be relevant before starting the search. The following example shows a designer, with little previous experience on honeycomb construction, expressing his need for exploring the topic of honeycomb construction.

> *We are going to work on this honeycomb construction now. I have never worked with honeycombs before. I do not know what is relevant and what is not. I need to learn a bit more about what is involved in this type of construction.*

Once designers had acquired some knowledge on a topic, their information needs changed to understanding issues. At this level, the designers had a better idea of what they were looking for. They would recognise what they were looking for if they saw it, but they were still likely to find difficulty in verbalising their needs beforehand. Continuing with the previous example, the same designer might now want to learn about issues involved in honeycombs such as crack growth.

> *I identified that crack propagation is a problem in honeycombs. I would like to know more about what are the risks of the crack propagation in my design case and how these risks could be mitigated.*

At the lowest level of detail, designers needed to define product specifications. In this type of information need the designers knew what they were looking for and they were able to verbalise it. These information needs were related to the definition of manufacturing specifications of components. An example at this level of detail is the definition of the type of aluminium alloy for the honeycomb.

> *I am going to select now the type of aluminium alloy that meets the requirements of this honeycomb construction.*

4.2 Access information

Once the information need had been defined, designers needed to access the most promising sources of information within the manuals. An analysis of the data showed that the accessing strategies were dependent on the retrieval system used by the designers. Four types of accessing strategy were identified:

- Navigating through indexes
- Scanning documents
- Linking to other sources of information
- Keyword searching (only for the electronic system).

Navigating consists of identifying relevant documents from a human-generated index. It was a common and successful accessing strategy when designers were unfamiliar with the manuals. This strategy enabled the designers to obtain an insight into the content of each document by identifying related keywords in their titles.

Scanning consists of flicking through documents looking for related keywords, figures and formulas. This strategy was especially successful for defining product specifications, when designers were often able to verbalise their needs. However, it performed very poorly when the designers did not have search criteria, as was often the case when, for example, exploring a topic.

Linking consists of triggering information from sources other than the manuals. Due to the design of the experiments, the only additional sources of information available to the designers were previous experiences stored in their memories, current project information, and past project information. One example of linking was when one designer realised that the honeycomb construction for the flying control surface was similar to the one used in the rudder of a previous aircraft. This designer preferred to retrieve the drawings of the rudder and get information about sealing from those drawings rather than from the manuals.

Keyword searching is the type of search performed by the search engine embedded in the electronic system. The performance of keyword searching depends on how sophisticated the search engine is. The search engine embedded in the electronic system only allowed Boolean searches of keywords contained in the titles of documents.

4.3 Retrieve information

The information retrieving strategies can be classified depending on each designer's knowledge on the required information. Four types of retrieving strategies were observed:

- Discovering
- Recognising
- Finding
- Triggering.

Discovering was the retrieving strategy used by designers when they did not have tacit search criteria for the required information. The designers simply read through documents to find explicit statements related to their information needs. For example, one designer was searching for issues related to the design of honeycombs. Stress corrosion is a relevant issue, but he only discovered this when he came across a statement in a document explicitly indicating its importance.

Recognising was the strategy used by the designers when they had tacit search criteria, but were not able to verbalise them. This retrieving strategy consisted of scanning through documents and focusing on those promising pieces of information that matched the tacit criteria. In these cases, if the designers were presented with relevant information, they were able to recognise it. For example, one designer was looking for the principles of cracking in metallic structures under stress and corrosion. This designer was not able to verbalise what he understood by 'principles', but he had a feeling for what he was looking for. He went into a document and concentrated on recognising the information that could match these tacit criteria. The information finally retrieved was related to threshold stress, residual and assembly stress, and protective treatments.

Finding was the strategy used by the designers when they could verbalise their search criteria. This retrieving strategy was based on scanning for a precise set of keywords and synonyms. The search engine embedded in the electronic system automated this strategy. For example, one designer was looking for blind bolts for the attachment of the spigot to the FCS. The context in which these bolts were going to be used was known. The designer was therefore able to verbalise the requirements of the blind bolts and establish precise keywords.

Triggering was the strategy used by the designers when they could remember information related to the information need. For example, one designer was navigating through an index looking for issues on honeycomb construction. He then read the heading of a section on draining holes. 'Draining holes' triggered from his background knowledge that corrosion is an issue in honeycomb construction.

The information retrieved during the experiments was classified according to the following six types:

- Product specifications

- Issues
- Characteristics
- Requirements
- Constraints
- Best practices.

Product specifications were the instructions for manufacturing. The designers searched for product specifications once they had already determined their requirements. Examples are types of material or types of fastener.

Issues were topics that the designers needed to explore in order to develop their design requirements. For example, one designer became aware from the documents that lightning strike was an issue for the design of the attachment between the leading edge and the FCS.

I am looking at what sorts of things are involved in this attachment. I am looking through the index of this company manual and there is a section on lightning strike. That springs to mind that we should design the attachment in such a way that current is able to pass from the leading edge to the spar.

Characteristics referred to the attributes of product specifications. The designers needed to retrieve the characteristics of product specifications to ensure that the specifications were able to meet the requirements. For example, one designer had to select a type of bolt (product specification) for the joint between the leading edge and the front spar. However, this joint had an assembly constraint: once the leading edge was mounted on top of the spar, there was no access to the inside of the joint. The requirement for the type of bolt was that it could be tightened from the outside only, i.e. they must be blind bolts. Thus when the designer looked for bolts in the manuals, one characteristic he was looking for was 'blind' bolt.

Requirements were the characteristics demanded or imposed on product specifications in order to address successfully the design issues. An example of a requirement is the threshold working stress of a material.

Constraints were restrictive conditions for the product specifications. The difference between requirements and constraints was how they influenced the retrieval of information. Designers aimed at finding product specifications that met product requirements, and then ensured that they met the constraints.

Best practices were the recommended procedures to undertake a task or standard values used in a company for a particular requirement. For example, an empirical formula to calculate the pitching and edge distance of bolts in a joint or a particular safety factor were some of the best practices retrieved.

5 Conclusion

This research was prompted by the view expressed by some experienced designers that new electronic systems were hindering the retrieval of information compared to the traditional paper systems. The concept of efficiency and ease of use in retrieval systems are in themselves relatively abstract and one of the challenges faced was to understand the process

of accessing and retrieving information – its nature and its taxonomy. This understanding may provide the corner stone to build a representation of the efficiency and ease of use of any retrieval system, regardless its nature. This representation will enable researchers to assess the performance of the paper and electronic systems and to determine whether the designers' perceptions matched reality.

The search for information from manuals in the engineering domain is a purposeful activity that aims to provide designers with answers to their information needs. It takes place in two stages: (1) the access of the relevant document; and (2) the retrieval of information from the document. The access of information is driven by the retrieval system used to access the manuals, and it is characterised by the access strategy and the source of information. The retrieval of information is driven by the users' information needs, and is characterised by the retrieval strategy and the type of retrieved content-information.

The conclusions that emerged from the results of this study provide some valuable insights. However, care must be taken not to over generalise these conclusions. The scope of the study was limited by its size and context, with the main characteristics being:

- The observations took place in the aerospace domain
- The designers were working individually and could not consult colleagues
- The search engine used during the experiments was very basic and a more advanced one would probably have altered the results
- Think-aloud protocol may affect the way designers undertake activities. e.g., reducing instinctive actions.

References

[1] Marsh, J. R. "The capture and utilisation of design experience in engineering design". PhD thesis. Department of Engineering. Cambridge, University of Cambridge, (1997).

[2] Wallace, K. M., Clegg, C., and Keane, A. "Visions for engineering design: a multi-disciplinary perspective". ICED 01, Glasgow, Scotland, UK, (2001), pp. 107-114.

[3] Aurisicchio, M. "Characterising information acquisition in the engineering domain", PhD thesis. Department of Engineering, University of Cambridge, UK, (2005).

[4] del-Rey-Chamorro, F. M., and Wallace, K. M. "A study of information retrieval in the aerospace domain", ICED 03, Stockholm, Sweden, (2003), pp. 271-272.

[5] Ullman, D., Dietterich, T., and Stauffer, L. "A model of the mechanical design process based on empirical data", AI EDAM, vol. 2 (1), (1988), pp. 33-52.

[6] Wilson, T. D. "Models in information behaviour research. Journal of documentation", vol. 55 (3), (1999), pp. 249-270.

[7] Sheng, F., Fan, X., Thomas, G., and Ng, P. "A knowledge-based approach to effective document retrieval", Journal of Systems Integration, vol. 10 (4), (2001), pp. 411-436.

[8] Talja, S. & Keso, H. & Pietiläinen, T. "The production of context in information seeking research: A metatheoretical view", Information Processing and Management 35, (1999), pp. 751-763.

[9] Court, A. W., Culley, S. J., and McMahon, C. A. “Information access diagrams: a technique for analyzing the usage of design information”, Journal of Engineering Design, vol. 7 (1), (1996), pp. 55-75.

[10] Vakkari, P. & Hakala, N. “Changes in Relevance Criteria and problem Stages in Task Performance”, Journal of Documentation 56 (5), (2000), 540-562.

[11] Savolainen. S. “The role of the internet in information seeking. Putting the networked services in context. Information Processing and Management”, vol. 35 (6), (1999), pp. 765-782.

[12] Horton, R. “Factors contributing to the use of a company intranet. PhD thesis. Institute of work Psychology”, Department of Psychology, University of Sheffield, UK, (2002).

[13] Kuhltahu, C. C. “Principle of uncertainty for information seeking. Journal of Documentation”, vol. 49 (4), (1993), pp. 339-355.

[14] Xie, H. “Patterns between interactive intentions and information-seeking strategies”, Information processing and management, vol. 38 (1), (2002), pp. 55-77.

[15] Sheldrick Ross, C.. “Finding without seeking: the information encounter in the context of reading for pleasure”, Information Process Management (6), (1999), pp. 783-799.

[16] Goldschmidt, G. “The dialectics of sketching”, Creativity Research Journal, 4(2), (1991), pp. 123-143.

[17] Suwa, M., Purcell, T., and Gero, J. “Macroscopic analysis of design processes based on a scheme for a coding designer's cognitive actions”, Design Studies, 19 4, (1998), pp. 455-483.

Francisco del Rey
KBR E&C Division
Europe and Africa Headquarter
Hill Park Court, Springfield drive
Leatherhead, Surrey
KT22 7NL
United Kingdom
Tel: +44 (0)7917053001, Fax +44 (0)1372 860000
E-mail: francisco.del-rey@halliburton.com

Ken Wallace
Engineering Design Centre
Department of Engineering
University of Cambridge
Trumpington Street
Cambridge
CB2 1PZ
United Kingdom
Tel: Int +44 1223 330271, Fax: Int +44 1223 332662
E-mail: kmw@eng.cam.ac.uk