<http://informationr.net/tdw/publ/papers/2004ISBandDigitalInfo.html>

|  |  |
| --- | --- |
|  | Information seeking behaviour and the digital information worldProfessor T.D. WilsonInformation Research:  an international electronic journal; Sheffield, U.K. **Abstract**  *This paper explores the nature of information seeking behaviour, showing how behaviour in relation to electronic databases and the World Wide Web may be formed through previous experience of non-electronic sources, such as the telephone directory and the book index. Models of information seeking behaviour are presented and the difficulties experienced by people in discovering how to search for information are explored, with reference to research in information science.* Introduction Scholarly information behaviour today is governed to a significant extent by the existence of a wide variety of electronic information sources, from the online databases that emerged more than thirty years ago, to electronic journals and the World Wide Web. In academia, the use of libraries as places in which to search for information is being replaced, to an increasing degree, by access to these information resources from the scholar's office desk, from the desk at home, from the laptop in the conference hotel, or, with the increasing pervasiveness of wireless connectivity, from the airport lounge. With this increasing mobility comes challenges for the developers of information services to develop systems that support the user in his or her search. Models of information seeking behaviour Various models of information behaviour and information search behaviour exist, usually based upon different assumptions. For example, that of Ellis (1989, 1993) attempts to describe the characteristics of information seeking behaviour, identifying a number of activities that are not assumed to take place in any specific sequence, but which may be undertaken to varying degrees and at different times. The list of activities is:  *Starting*  *deciding to undertake the search;*  *Chaining*  *following citations backwards and forwards in time;*  *Browsing*  *casual, relatively undirected search;*  *Differentiating*  *discriminating among potentially relevant items;*  *Monitoring*  *regular scanning of sources for items of interest;*  *Extracting*  *taking information from sources;*  *Verifying*  *identifying potentially useful citations correctly;*  *Ending*  *completing the process.*  All of these activities will be well known to anyone who has ever embarked upon a search for information in relation to, for example, preparing a research proposal.  Kuhlthau (2004) on the other hand, proposes a process that links stages in the search to associated feelings:  *Initiation of the search: associated with uncertainty; Selection of potentially relevant items - optimism; Exploration for further material - confusion, frustration, doubt; Formulation of a clear strategy - clarity; Collection of further material - sense of direction, confidence; and Presentation of results - satisfaction or disappointment.*  Wilson (1981, 1997, 2002) has proposed several models - Figure 1, from 1995 is a development of a model first proposed in 1981, which drew attention to the link between cognitive and affective elements in the motivation to search for information and which suggested that 'barriers' (here referred to as intervening variables) existed between the need for information and the decision to search for it.  **figure1**  **Figure 1: A general model of information seeking behaviour**  The 1981 model was further developed through an analysis of research on information seeking behaviour in a variety of fields from consumer research to health information management.  Wilson's third model involves the idea of a problem solving process to which recurrent searches for information are connected as information problems are experienced in the process of solving the more general research problem:  **Figure 2: A problem solving model of the search process (Wilson, 2002)**  This model was tested in a study of uncertainty in information seeking (Wilson, 2002)  These models are described here since they may help the editors of science journals to envisage the circumstances in which journals may be used and to illustrate the complexity of the information seeking process. In the world of electronic information resources it is also reasonable to ask whether all appropriate steps are taken to make the search for information in electronic journals a 'comfortable' process for the information seeker. Information seeking over time An historical digression may be useful to set the present use of information resources in context. Probably the earliest form of written communication among scholars took the form of letters. Until the invention of the scientific journal and beyond, the chief means of communication among scholars in Europe was by letter and, indeed, these letters were sometimes copied and circulated to fellow scholars, or, at times summarised and built into a new letter. As might be expected the information source was conceived of as the person, rather than the means of communication. Of course, this continues today, formalised to some extent by the e-mail mailing list, but being carried on between individuals and on private discussion boards.  With the emergence of the scientific society in the seventeenth century, we have the development of the proceedings and their circulation - initially as a record of events and as a communication to members unable to attend the meetings and then as a communication to the world of scholarship at large - which, at the time, of course, was not very large. As science (under all of its definitions) grew, so journals proliferated and the task of keeping track of developments in any single discipline, let alone across the whole field of science - hence the emergence of abstracting and indexing services, subsequently converted into online databases with the emergence of the computer age. The WWW is the present culmination of this development - but, of course, its impact is more revolutionary than evolutionary. The development and displacement of habit It seems reasonable to suggest that as the as the means for the delivery of information have changed, certain habits of searching become embedded in scholarly practice. For example, in the early years of science, the members of scientific societies would receive the proceedings of similar organizations - there would not be very many. This practice would become embedded and it continues today, when it is not unusual for someone to belong to more than one society in his or her field and to receive the associated journals. The informal network also flourished with people in the same institution exchanging information, journal copies, and, when the technology arrived, photocopied papers. With the advance in technology the offprint and the photocopy are replaced by electronic files.  Then, as libraries developed to house the increasing number of scholarly productions, people got into the habit of having a regular time to check up on the latest journal issues and to search new issues of indexing and abstracting journals as they came along. The browsing behaviour persists, although it is less well catered for in the digital sphere.  My basic proposition, therefore, is that our habits of searching have built up over time and constitute a kind of stratified silt of activities. The older we are, the more ways we will have acquired of doing things as the scholarly communication process has changed and as the technology of information delivery has changed. We will tend to fall back on the familiar when we first tackle a new technology and only slowly acquire the new habits. For the young scholar, in his/her first post-doctoral position, there is less to unlearn and the habits are those acquired over a relatively short period of time and relate to the here and now rather than the past. The disadvantage is that the younger scholar may assume that if earlier work on a problem cannot be found in the material available online, it does not exist - this ought to be a stimulus to the publishers of science journals to ensure that back files are digitized or, if this is uneconomic, that the indexes are digitized.  Of course, something also depends upon the nature of the discipline and how far it has been affected by the new developments in communication and technology. The science disciplines have recognized the possibilities faster than others, probably because, historically, they were originally the only departments in universities that even thought of using computers. On the other hand there are still departments in the humanities in some places that are not well provided with the appropriate technology and have not yet got into the habit of equipping their students with the skills of computer use.  There is also the point, of course, that the sciences differ from the humanities in the nature of their practice - for the scientist, the record of the experiment, assuming it is validated as authentic and rigorous, is acceptable, whereas the historian will need to seek out the original records of events in preference to a secondary analysis of the same documents. Familiar search tools When we come to consider the way people behave in searching the 'digisphere' we have to review that matter of habit, and look at the kinds of tools we are familiar with in the general area of information seeking. Probably the two commonest forms of information device that we come to know from fairly early in our lives are the telephone directory and the book index.  Consider, for example, the telephone directory:  *BOWERS-GREEN L,*  *100 High Road, Wortley..........(0141) 268 9148*  *BOWES C.E., 31 Rayburn Road, 9......(0141) 286 5266*  *BOWIE K, 12 Cliffton Av, 10.........(0141) 320 1200*  *L, 137 Sandybridge Rd, 6........(0141) 268 2453*  *M, 16 Worrlyn Pk., 4............(0141) 203 1280*  *BOWINS K, 3/12 Ratcliffe Bank.......(0141) 243 0971*  *BOWKER P, 91 Maple Dv, 7............(0141) 265 5308*  *BOWLER A, 61 Owen Gro, 9 ...........(0141) 246 9734*  *A, 10 Hollins Clo, 6 ...........(0141) 223 6133*  *D.R, 34 Owen Dr.,9 .............(0141) 245 7823*  *E, 10 Chestnut Av, 10 ..........(0141) 238 9120*  *E, 175 Ingram Rd, 11 ...........(0141) 230 4349*  *K, 130 Fulmere Av. 8 ...........(0141) 232 1515*  *BOWLES D, 5 Hereford Rd, 3 .........(0141) 267 4278*  **Figure 3: the telephone directory**  The key point is that the 'search terms' are very simple - persons' surnames, followed by a brief indication of the address and telephone number. Generally, however, there are no 'back indexes' that allow one to search for an address and discover who lives there, of for a telephone number to know whose it is.  The index to the classified part of the telephone directory shows a similarly simple structure:  *Ear and Body Piercing 70*  *Educational Services 70*  *Electric Motor Manufacturers and*  *Suppliers 70*  *Electric Motor Rewinds and Repairs 70*  *Electrical Appliance Rental 70*  *Electrical Appliance Stores 70*  *Electrical Component Manufacturers 70*  *Electrical and Domestic Appliance*  *Repairs 70*  *Electrical Engineers 71/*  *Electrical Instruments Makers 71*  **Figure 4: Index to classified telephone directory**  The other common tool, familiar from early studies at school onwards, is the book index:  *Pisano, G., 59*  *Polanyi, M., 4, 71*  *Price discrimination, 158*  *Productivity*  *ICT and, 27*  *knowledge management and,*  *222-223*  *paradox of, 31-35*  *R&D and, 114*  *Prusak, L., 207-208*  **Figure 5: A book index**  Again, the search elements are very simple - personal names and topics.  When we examine these familiar tools, we can readily understand why people use very few words in searching for information - several studies have drawn attention to this fact, reporting the average number of terms used in a search (without Boolean connectors) as being in the range 1.7 to 2.8. The database search model With the introduction of the online databases, most notably those developed by DIALOG, the academic user was introduced to Boolean search strategies - although, because these databases were not directly accessible by the end users, it was generally a librarian who acted as intermediary to carry out the search. Because the online databases needed to be searched by an intermediary, few information seekers acquired the habit of using Boolean formulations in their searches and, although today's Web search engines often allow for Boolean searches, few people make use of them.  Some years ago I had an opportunity to download the records of one million searches on the Excite search engine - this was rather more than I could manage easily to organize, but I did analyse 1,000 searches. The most complex strategy was:  *(bikini OR swimsuit) AND NOT (CD-ROM OR CDROM OR throbnet OR catalog OR catalogue OR video)*  In fact, only 12.8% of the searches used Boolean operators, and I was generous in my interpretation of "Boolean" - accepting a phrase as Boolean if it contained the word 'and' (as in, for example, "soccer history and the name of soccer") However, the focus on words is a little misleading. In preparation for this paper I took a look at the data again, and found an average of 2.11 words for a search statement - but only 1.13 for the number of concepts. For example, the search statement "bose lifetime 12" has three elements, but simply defines a single electronic product - so, in general people are searching for a single concept. A number of other studies support these findings, e.g., Spink, et al., 1998 and Spink & Xu, 2000.  In a forthcoming paper (Wilson, forthcoming) I explore how people talk about their information problems and the factors that motivate their seeking help in searching for information. It is clear from this work that people feel that they lack the skills to conduct a search effectively and turn to the professional information worker because they believe that he or she knows how to do it.  One enquirer, for example, identified his need for information as being concerned with the "eco-vegetation of shingle plants" - this phrase revealed no sources of information either on the Web or in the Web of Science databases. However, when these alternative search terms were used, relevant information was found:   * shingle AND plants AND water * (shingle OR pebbles) AND plants AND water * shingle AND plants AND (water OR tidal) * (shingle OR pebbles) AND plants AND (water OR tidal) * shingle beach AND plants AND nutrient  Conclusion People's information seeking behaviour is formed as they learn, throughout life - it is part of the general learning process and, in all probability, structured according to the tools they find in the course of daily learning. To a significant extent, therefore, it hss been the tools of print materials that structure behaviour, of which the book index and the telephone directory are the most obvious. Increasingly, however, a new generation is arising that is accustomed to find information through electronic systems and their behaviour may be formed by the tools they find there.  In general however, information providers must assume that information users will adopt very simple search strategies in seeking information - their systems must be structured either a) to infer complexity from the simple terms entered in a search, or b) provide an interactive process through which more complex requirements can be expressed. In various areas of information retrieval research both of these possibilities are being explored. However, past information retrieval research gives us little confidence that any major breakthrough can be expected: whatever tools are employed to determine semantic relationships in text they appear to be inadequate to the task. Note This paper is based on a presentation at a Symposium of the European Association of Science Editors, Barcelona, 7th May, 2004. References  * Ellis, D. (1989). A behavioural approach to information retrieval design. Journal of Documentation, 45(3), 171-212. * Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. Journal of Documentation, 49, 356-369. * Kuhlthau, C. C. (2004). Seeking meaning: a process approach to library and information services. (2nd ed.). Westport, CT: Libraries Unlimited. * Spink, A. and Xu, J.L. (2000). Selected results from a large study of Web searching: the Excite study. Information Research, 6(1), paper90. Retrieved 6 January, 2003 from http://informationr.net/ir/6-1/paper90.html * Spink, A., Bateman, J. and Jansen, B.J. (1998). Searching heterogeneous collections on the Web: behaviour of Excite users. Information Research, 4(2), paper 53. Retrieved 6 January, 2003 from http://informationr.net/ir/4-2/paper53.html * Wilson, T. D. (1981). On user studies and information needs. Journal of Documentation, 37(1), 3-15. * Wilson, T. D. (1997). Information behaviour: an interdisciplinary perspective. Information Processing and Management, 33(4), 551-572. * Wilson, T.D. (Forthcoming). Talking about the problem: a content analysis of pre-search interviews. [Paper to be presented at the Information Seeking in Context Conference, Dublin, 1-3 September, 2004.] * Wilson, T. D., Ford, N. J., Ellis, D., Foster, A., & Spink, A. (2002). Information Seeking and Mediated Searching. Part 2. Uncertainty and Its Correlates. Journal of the American Society for Information Science and Technology, 53(9), 704-715.  How to cite this paper: Wilson, T.D. (2004) "Information seeking behaviour and the digital information world" European Science Editing, **30**(3), |