Investigating methods for understanding user requirements for information products

Mark Hepworth
Senior Lecturer, Division of Information Studies
Nanyang Technological University
Singapore
asmark@ntu.edu.sg

Introduction

This research is concerned with methods that can be used for helping to understand people's requirements for information products. Two questions are central to this research:

1. What kind of data should we try to capture about people and their interaction with information so that we can have a detailed understanding of their requirements?
2. Having determined what we need to find out about, what research techniques are most appropriate for capturing the relevant data?

To help answer these questions literature from 'user studies' and 'information retrieval' (library and information science), human computer interface design, and systems analysis and design were reviewed. This resulted in a conceptual framework that indicated the data that need to be captured. This was followed by a review of literature from the areas mentioned above as well as research methodology to identify techniques that could be applied. Different techniques were evaluated.

Together this formed a methodology that was applied to a community to see whether a useful understanding of their requirements could be derived. The community chosen, students, is one that is relatively well understood and has been the focus of research for many years. It was felt that choosing a community where data on their needs exists would help to provide an indication of the effectiveness of the methodology. At minimum it should be able to derive similar insights to those that have been built up over the years. At best it would provide a more detailed understanding of this community in terms of their interaction with information and the implications for an information product that would meet their requirements. The application of the methodology therefore enabled the research to understand the strengths and weaknesses of the methodology.

The stages of the research

The following outlines the stages the research has gone through and briefly the findings.

Stage 1. What kind of data should we try to capture about people and their interaction with information so that we can have a detailed understanding of their requirements?

To help answer this question a number of papers were reviewed. The following cites those that had particular impact on the researcher: from user studies (Taylor, 1968; Garvey, Nan Lin, and Nelson, 1971; Wilson, Streatfield and Mullins, 1979; Wilson, 1981; Streatfield 1983; Hogeweg De Hart, 1983, 1984; Dervin and Nilan, 1986; Borgman, 1996); from human computer interface design (Preece, *et al.*, 1994; Hill, 1995; Hartson and Boehm-Davis, 1995; Shackel 1997), and from systems analysis (Brown, 1994; Robinson and Prior, 1995; Mehdi Sagheb-Tehrani, 1995; Wixen and Ramsey, 1996; Underwood, 1996; Tudor and Tudor, 1997; Simonsen and Kensing 1997). This process was facilitated by a number of articles published in the 90s that helped to pull together and differentiate between various research approaches, (Hewins, 1990; Allen, 1991; Westbrook, 1993; Ellis, 1993; Wilson, 1994; Ingwersen, 1996; Ellis, 1996). There were found to be a great diversity of approaches. Few are specifically related to the specification of user requirements for information products.

However four common themes were identified. First the '*sociological*' which highlights the importance of the social context, (the roles, the tasks), of the respondent. Second the '*content*' area which includes the physical environment and tools that the respondent interacts with and are associated with their roles and tasks (such as books, the Internet, etc.). Thirdly the '*psychological*' which emphasises the cognitive and affective domain. Fourthly the '*behavioural*'. Bandura's (1986) triadic framework of,

- Behaviour e.g. browsing the shelves or entering search statements;
- Environment e.g. tasks, roles, subject matter, informationsystems, services and products
- Cognition & Personal Factors (thoughts and emotions) e.g. wanting to refine a search and feelings of confusion.

was then adopted to help conceptualise the relationship between these themes identified.

Bandura's notion of **environment** therefore incorporates both the *sociological* and the *content* dimensions. When studying the users the **environment** relates to their context both in the sociological sense of the tasks and roles they have to perform but also the physical environment i.e. the books, the articles, the OPACs, World Wide Web pages etc. that they interact with and get feedback from. Feedback that follows and to some extent stimulates **behaviour** that is based on experience, knowledge and perception. **Cognitive and personal factors** influence and result in behaviour and are related to the environment, background and tasks of the respondents. In this research it should be noted that personal factors, such as psychometric data, have not been explored but also have implications for the user's requirements. For example people may prefer either a virtual reality or three dimentional interface design, the 'rats-eye view', or the more abstract two dimensional 'birds-eye view' such as used by Windows 95 Explorer software interface (Howlett, 1996).

User requirements analysis therefore needs to capture the respondent's thoughts as well as their behaviour and the materials necessary to undertake information intensive tasks. Limiting one's study to any one dimension would result in only a partial picture of the users' requirements and hence would lead to the development of products that did not adequately support the user requirements.

The three dimensions 'environment', 'behaviour' and 'cognitive and personal factors' can therefore taken to be fundamental to understanding the potential user of information and hence user requirments analysis and provide the researcher with a framework for data gathering.

Stage 2. What research techniques are most appropriate for

capturing the relevant data?

The second stage of the research concerned the identification of techniques to elicit the various dimensions identified above. Choice of techniques was therefore driven by the need to gather data on the three dimensions outlined above.

Choice of research techniques depends on the ontology and epistemology of the researcher. In this case the researcher was influenced by a number of approaches including ethnographic, which emphasises the importance of the respondents local context, and sense-making in that people were perceived as active participants in the process and their perceptions and actions relate to and are formed in a dynamic way as they encounter specific situations (Dervin, 1992, 1994). This implied that data should be gathered while the respondents undertake the task and that perceptions need to be captured. It was also assumed that if data on a number of people in similar situations are studied generalisations could be made. Partly due to the influence of Kuhlthau (1991) and HCI research, respondents were studied over a period of time starting from task initiation until they had gathered what they thought was relevant information. This also reflected the notion that the objective was to derive user requirements for a product that would support the entire task.

Qualitative techniques were chosen partly because they are recognised as appropriate for exploratory research where variables are not clearly defined and also because of their efficacy in highlighting themes, processes and cognition of the respondent.

To help determine the most appropriate techniques a review of appropriate literature on research methodology was conducted, (Patton, 1990; Miles and Huberman,1994; Churchill, 1995; Nicholas, 1996; Neuman, 1997; Zikmund 1997). In addition various equipment (tape, video, forms, Lotus Screencam (screen/voice capture software)) and methods including observation, interview, talk-through, task analysis, task hierarchy diagrams and the critical incident technique were reviewed.

Finally a combination of the following were chosen as most appropriate,

- task analysis to capture the respondents perception of the task and sub-tasks; as well as the task environment (services, systems etc.),
- talk-through capturing verbalised thoughts while respondents conduct tasks, providing data on the cognitive dimension
- observation to capture behavioural data (actions) as well as information about the environment with which the respondent is interacting.

To capture data, forms were designed that sensitized the researcher to collecting data on the three dimensions: cognitive and personal factors, behaviour, and the environment.

Stage one and two therefore resulted in a possible framework for understanding what data should be captured and how the data should be gathered.

Stage 3. Implementation of the methodology.

Fifty Master of Science Information Studies students were divided into six groups. Each group chose one of three research topics. Each week the students rotated between being either researchers or respondents. Peer pressure and the fact that a small proportion of marks for course assessment were awarded for 'participation' in the project helped to ensure that they were serious about the task. At the end of semester each group were also expected to derive their own solutions for an information product.

Implementation took the following course.

a. An initial task analysis interview was conducted with the respondents. The aim was to capture the users' overall perception of the task. The task being to go through the process of gathering material on the topic to the point where they would start writing an essay.

The interview used 'What' questions to identify major tasks, 'How' questions to identify sub-tasks and 'goal' questions to identify objectives and outcomes, (Sebillotte, 1988). Researchers were monitored to try to ensure that they did not depart from this format or provoke 'correct' responses. Task Hierarchy Diagrams were derived. These diagrams helped to reveal aspects of the environment that the respondents expected to interact with including the services and products they expected to use and also, to a lesser extent, the associated cognitive and behavioural tasks.

b. Once respondents started to undertake the assignment the processes, perceptions, actions and objects were recorded using a combination of talk through technique and observation. Predefined forms sensitised researchers to capturing cognitive, behavioural and environmental data.

Researchers limited themselves to only asking about the thoughts of the user, 'what are you thinking now?' and were not expected to ask probing questions, such as 'Why…' or make any suggestions that would lead the respondent in a particular direction.

- This generated eight Task Hierarchy Diagrams and the 52 talk-through and observation forms. A total of 1160 incidents were identified in the transcripts. An inductive approach was used to categorise these incidents.

An overview of the respondents' information tasks

The respondents having chosen a question to answer

a. spend time understanding what the question is about. Respondents "look at the question carefully", "think back to what they have read", think about "what is expected" and generally try to define the topic. They may wish to contact an expert for help.
  - After searching and retrieval respondents still returned to and continued the process of topic definition. In general respondent's found this aspect of the overall task very difficult.
  - As Kuhlthau (1991) has pointed out this stage was associated with confusion and trepidation.
b. The respondents then started to choose systems and services to search. This was influenced by physical location of resources, the perceived subject content, types of material available and also familiarity.
c. Search terms and combinations of terms were identified (often with great difficulty) and tried out. Depending on the response from the systems and the relevance of material retrieved respondents may narrow ("refining"), broaden or "redefine" their searches. Numerous attempts are often made using different terms. Different systems, locations, organisations may also be tried.
  - Respondents were not clear how the various systems worked. Systems such as Yahoo were chosen because it was "memorable". A great deal of frustration was associated with using these services. Respondent's were "exasperated", "befuddled", felt "irritation" and "inadequacy".
d. After viewing hits, headlines, abstracts, texts relevant information may be identified. Relevance was identified by recognising significant terms. Additional or more appropriate terms were identified and searches refined and redefined. Useful terms and bibliographic data were noted. Searches became

more precise with more boolean 'anding'. Again iteration, narrowing, broadening, redefinition, choosing alternative systems and locations took place.

- Respondents were "excited", "relieved", once relevant material was found.

e. Once material was found either electronically in full text or via locators, such as library call numbers, extensive browsing of the location and the media took place. Shelved material was located via call numbers and titles browsed for specific or significant terms.

- In both journals and books contents pages, chapter/article headings and sub-headings, indexes were scanned. Introductory sentences and paragraphs, conclusions, citations were also scanned. Other criteria such as format and appropriateness of theinformation were considered.

f. Respondents captured information but also returned to earlier processes of choosing systems, services, defining the topic, refining and redefining the search.

Conclusion

The conceptual framework (the triadic environmental, behavioral and cognitive dimensions) developed on the basis of previous user studies, information retrieval, systems analysis and HCI studies helped the researcher to determine what data should be collected. It also helped to identify appropriate research techniques such as the talk-through technique to capture cognition. Recognising that respondents are involved in a highly interactive and contextually sensitive sense-making process has also influenced the choice of a qualitative and ethnographic approach.

The techniques for data capture were able to be used by relatively inexperienced researchers i.e. the students. These techniques served to provide a rich picture of the respondents' user requirements and identified six main tasks and sixty-three sub-tasks. There was evident correspondence with previous findings such Kuhlthau's (1991); 'initiation', selection', 'exploration', 'formulation', 'collection' (but not 'presentation'), as well as the affective dimension. Eisenberg's and Brown's (1992) categories of information skill, 'task definition', 'development of information seeking strategies', 'location and access', 'information use', can also be recognised in the tasks and sub-tasks identified.

Ellis' (1993), 'starting', 'chaining', 'browsing', 'differentiating', 'extracting', 'verification' and 'ending' were also identified. Chaining however was less apparent perhaps because this tends to be associated with the later stages of the research process, which was not studied. This is also true of 'verification' and 'ending' or Ellis' 'monitoring' which were not evident in this study due to the nature of the task.

To develop an information product that meets these needs will require a great deal of work in the areas of,

- generating metadata about collections, media, and their information content and the character of 'texts'.

- Enabling subject definition and the identification of search terms will also require the development of tools that the user can interact with, rather than automatically generating terms in the background.
- Digitisation of key parts of media such as contents pages, chapter headings, sub-headings and introductions is necessary to aid relevance judgements.
- Cognitive tasks such as being able to broaden or narrow the search will need to be supported.

Some of these features are already evident in evolving information products. However no system or product currently supports the full range of requirements identified.

It should be emphasised that, although the study can be seen to have implications for an information product for students, the study did not cover the entire research and report generation process and that due to the qualitative nature of the study generalisations and solutions are specific to the community studied. However judging from the literature and current digital library and information retrieval solutions some of these requirements may be generic and extend beyond the specific community.

Individual tasks and sub-tasks need further research in terms of the respondent's perception of the task. This would involve additional research techniques. The effect of personal factors (such as psychology or knowledge) may be significant and may, for example, have impact on the 'look and feel' of the interface design as well as the undertaking of specific tasks. Different types of question, task and role will also have implications for the user requirements.
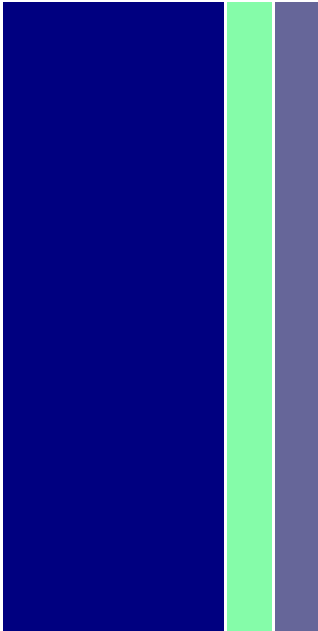
Acknowledgements

References

- Allen, B. L. (1991). Cognitive Research in Information Science: Implications for Design. In Williams, M. (ed.) *Annual Review of Information Science and Technology* (ARIST), Medford, NJ: Learned Information, 26, 3-37
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs. NJ: Prentice Hall
- Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science,* 47(7), 493-503
- Brown, D. (1994). STUDIO: *Structured User-interface Design for Interaction Optimisation*. London: Prentice Hall
- Churchill, G. A. (1995). *Marketing research: methodological foundations*. Fort Worth, TX: The Dryden Press.
- Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. In J.D. Glazier, R.R. Powell (eds*.), Qualitative research in information management*, 6-84,. Englewood , CO: Libraries Unlimited.
- Dervin, B. (1994). Information – Democracy: An Examination of Underlying Assumptions. *Journal of the American Society for Information Science*. 45(6), 369-385
- Dervin, B. and Nilan, M. S. (1986). Information needs and uses. In Williams, M. E. (ed.) *Annual Review of Information Science and Technology*, Medford, N.J.: Knowledge Industry Publications, Inc., 21, 3-33
- Eisenberg, M. B. & Brown, M. K. (1992). Current themes regarding library and information skills: research supporting and research lacking. *School Library Media Quarterly*, 20(2), Winter, 103-110.
- Ellis, D. (1993). Modelling the information seeking patterns of academic researchers: a grounded theory approach. *Library Quarterly*, 63(1), 469-486
- Ellis, D. (1996). The Dilemma of Measurement in Information Retrieval Research. *Journal of the American Society for Information Science*, 47(1), 123-136
- Garvey, W. D. Nan Lin and Nelson C. E. (1971). A comparison of scientific communication behaviour of social and physical scientists. *International Social*

Science Journal, 23(2), 256-272

- Hartson, H. R. and Boehm-Davis, D. (1993). User interface development process and methodologies. *Behaviour & Information Technology*, 12(2), 98-114
- Hewins, E. T. (1990). Information Need and Use Studies, vol. 25 of *Annual Review of Information Science and Technology*, ed. Williams, M. American Society for Information Science: Elsevier Science Publications, 142-172
- Hill, S. (1995). *A practical introduction to the Human-Computer Interface*. London: DP Publications
- Hogeweg-De Haart, H. P. (1983). Characteristics of Social Science Information: A Selective Review of the Literature. Part I. *Social Science Information Studies*, 3, 147-164
- Hogeweg-De Haart, H. P. (1984). Characteristics of Social Science Information: A Selective Review of the Literature. Part II. *Social Science Information Studies*, 4, 15-30
- Howelett, V. (1996). *Visual interface design for Windows: effective user interface for Windows 95, Windows NT and Windows 3.1*. New York: John Wiley
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of cognitive IR theory. *Journal of Documentation*, 52(1), 3-50
- Kuhlthau, C. C. (1991). Inside the search process: Information Seeking from the User's Perspective. *Journal of American Society of Information Science*, 42(5), 361-371
- Mehdi Sagheb-Tehrani (1995). Knoweldge acquisition process some issues for further research. In Aamodt, A. and Komorowski, J. (eds). Proceedings of the *Scandinavian Conference on Artificial Intelligence*. Trondheim, Norway, May 29-31. Amsterdam: IOS Press. 448-452
- Miles, B. M. Huberman, A. M. (1994*). Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, CA: Sage Publications.
- Neuman, W. L. (1997). *Social research methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon
- Nicholas, D. (1996). *Assessing information needs: tools and techniques*. London: Aslib
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage Publications
- Preece, J. Roger, Y. Sharp, H. Beyon, D. Holland, S. & Carey, I. (1994). *Human Computer Interaction*. Wokingham, UK: Addison Wesley Publishing
- Robinson, B. Prior, M. (1995). *Systems Analysis Techniques*. London: International Thompson Computer Press.
- Sebillotte, S. (1988). Hierarchial planning, a method for task analysis: the example of office task analysis. *Behaviour and Information Technology*, 7(3), 275-293
- Shackel, B. (1997). Human-computer interaction-whence and whither? *Journal of the American Society for Information Science,* 48(11), 970-986
- Simonsen, J. Kensing, F. (1997). Using ethnography in contextual design. *Communications of the ACM*, 40(7), 82-83, 84, 86, 88
- Streatfield, D. (1983) Moving towards the information user: some research and implications. *Social Science Information Studies,* 3, 223-241
- Tudor, D. J. Tudor, I. J. (1997). *A comparison of structured methods*. Houndmills, UK: Macmillan Press.
- Underwood, P. G. (1996). *Soft systems analysis and the management of libraries, information services and resource centres*. London: Library Association
- Westbrook, L. T. (1993). User Needs: A synthesis and analysis of current theories for the practitioner. *RQ*, 32, 541-549
- Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, 37(1), 3-15
- Wilson, T. D. (1994). Information Needs and Uses. In Vickery, B. (ed*). Fifty*

years of information progress: A Journal of Documentation Review. London: Aslib 15-51

- Wilson, T. D. Streatfield, D. R. & Mullins, C. (1979). Information Needs in Local Authority Social Services Departments: A Second Report on Project INISS. *Journal of Documentation,* 35(2), 120-136
- Wixon, D. Ramey, J. (1996). *Field methods casebook for software design*. New York: John Wiley & Sons.
- Zikmund, W. G. (1997). *Business research methods*. Fort Worth, TX: The Dryden Press.