

UNIVERSITY OF COPENHAGEN



Information Seeking of E-government Employees An Analytical Study

Svarre Jonassen, Tanja; Lykke Nielsen, Marianne

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Svarre Jonassen, T., & Lykke Nielsen, M. (2008). *Information Seeking of E-government Employees: An Analytical Study*. Paper presented at 5th Scandinavian Workshop on eGovernment, Copenhagen, Denmark.

Information Seeking of E-government Employees: An Analytical Study

Tanja Svarre

Royal School of Library and Information Science, Information Interaction and Information Architecture, Aalborg Branch, Fr. Bajersvej 7K, 9220 Aalborg East.
tas@db.dk

Marianne Lykke Nielsen

Royal School of Library and Information Science, Information Interaction and Information Architecture, Aalborg Branch, Fr. Bajersvej 7K, 9220 Aalborg East.
mln@db.dk

Abstract. The paper presents studies of information seeking within the domain of e-government and related domains. The purpose of investigating the studies is to find out, which methods of automatic indexing could be relevant to apply, if the overall aim is to support a more effective handling of work tasks and information needs within e-government. On the basis of the literature review, a number of different types of work tasks and information needs among employees in e-government are identified. Meeting the diversity of needs call for either different types of indexing supplementing the strengths and weaknesses of each other or an indexing method integrating extracted and assigned indexing.

Key words: automatic indexing, information seeking, information needs, work tasks, employees, and e-government.

1 Introduction

Over the years a number of studies have been performed of information seeking and information needs within e-government. The main part of the studies has focused on citizens as end users and information searchers (e.g. Craig & Schriar, 2001; Marcella, Baxter & Moore, 2003; Chau, Fang & Sheng, 2007). The present study aims at exploring the information seeking behaviour of employees within the domain of e-government. In the years to come, it is expected, that the share of complicated work tasks will increase due to the fact that simple tasks are taken care of by citizens themselves, because of the increased possibilities of self-service (Landsforeningen af

Kommunale Servicecentre, Arbejdsgruppen om IT [2005], p. 4). A derived effect of this will be changes in the types of information needs and information seeking strategies of governmental employees. Meeting these changes in administrative information systems, e.g. electronic document management systems (EDMS) calls for some changes in the way documents are indexed. One way to improve the support could be the use of automatic indexing in EDMS.

Hence, we are examining studies of people having a professional approach to information retrieval in an e-government workplace context. The study seeks to investigate the following research questions:

- 1) What characterizes information seeking and information needs among employees within e-government?
- 2) What approaches to automatic indexing are relevant when trying to meet the identified information seeking behaviour and information needs of employees in e-government?

2 Methodology and theoretical framework

The basic assumption for the study is that the choice and design of indexing methods should take its point of departure in the *content* of the retrieval system (the information objects), the *users* of the content, and the *contextual setting* (Soergel, 1985, p. 41; Nielsen, 2001, p. 780; Byström & Hansen, 2005, p. 1050). In this study, we focus on the context, i.e. the context of e-government and the derived influence on indexing. We also study the users. However, we do not study individual users and their individual characteristics, but focus on the users' roles, work tasks and function in the work setting. We study work tasks and derived information needs and their influence on the information seeking behaviour and information needs. Based on this element it is relevant to analyze studies not just within the domain of e-government, but also from related work environments. Working with cases is a central work task within e-government. Depending on the single government in question, project work also occurs to different extents (FESD-projektet, Alliancen. [2004], p. 8). Therefore it is relevant to draw on the experiences from studies of working environments, e.g. engineers, since they have work tasks in common with the ones identifiable within e-government.

The first part of the paper will go through central theories and models of information needs and information seeking, as they are understood within the domain of information science. The examined theories and models will work as a frame of reference for reviewing existing studies on information seeking and information needs (research question 1). Furthermore, we will present the basic theory on automatic indexing in order to have a baseline for the analysis regarding the second research question.

2.1 Information seeking

In 1999, Wilson analyzed a number of models in information behaviour research (Wilson, 1999). On the basis of a review of central models for information seeking, Wilson outlines a nested model identifying three

research areas and the mutual relation between them; the three areas being information behaviour, information seeking behaviour, and information search behaviour. In Wilson's model, information behaviour refers to the general field of investigation. Information-seeking behaviour is subordinate to information behaviour and is concerned with "...the variety of methods people employ to discover, and gain access to information resources" (Wilson, 1999, p. 263). The third field of research is information searching, which yet again is nested within information seeking behaviour. Information searching refers to "...the interaction between information user [...] and computer-based information systems..." (Wilson, 1999, p. 263).

In this article we want to explore relevant approaches to automatic indexing when trying to meet information seeking behaviour and problems within e-government. Hence, we presuppose that a considerable part of solving the information problems within e-government will be consulting some kind of information system, e.g. an electronic document management system (EDMS). Thus, the central study object is information searching in electronic information retrieval systems (Wilson, 1999, p. 263). However this does not mean that research concerned with information seeking and information behaviour will be left out. This is caused by the fact that elements of information seeking and information behaviour can have an indirect influence on information searching. For example Wilson's (1980) study of social workers and social administrators show a quite short average time for most work sessions, including information retrieval. The limited time span is expected to have some degree of influence on the users' information searching behaviour.

2.2 Information needs

The concept of information need denotes a problematic situation which, unless the problem is very simple or routine, causes an information need (MacMullin & Taylor, 1984, p. 93). Different theories of the nature of the information need have been presented. Belkin, Oddy & Brooks (1982) and Taylor (1968) constitute central contributions within information science. Here, Ingwersen's (1986) three types of information needs will work as frame of reference for the analytical part of the paper. The three types are verificative needs, conscious topical needs and muddled topical needs. Originally the identification of the three types was based on empirical results users of libraries. When presenting the types below, we convert them into a governmental setting.

When having a verificative information need the user wants to locate or verify an item. The user possesses characteristic bibliographic data on the item wanted. Within information science bibliographic data usually refers to title, author, source and the like (Ingwersen, 1986, p. 223). Converted into the context of e-government bibliographic data could for instance refer to a citizen's CPR, full name, or address. In other words bibliographic data in the context of e-government refers to a more or less unique identification of e.g. a case or a citizen. The conscious topical information need refers to a situation, where "the user wants to clarify, review or pursue aspects of known subject matter" (Ingwersen, 1986, p. 223). Within the context of e-government an example of this could be an administrator receiving a new

case on a hitherto unknown citizen that is within a subject area known to the administrator. Another example could be initiating a development project within a known subject area. Perhaps the administrator worked on similar projects earlier. Finally, the muddled topical information need describes a user wanting to explore new concepts outside of subject matters known to the user ahead of the information need (Ingwersen, 1986, p. 223). In e-government an example of this could be a case within a subject area unknown to the administrator. Or it could be starting up a development project entirely or partly within a subject area, that are not prior known to the administrator.

2.3 Automatic indexing

Basically, there are two methods for indexing documents automatically, automatic extraction and automatic assignment of words. Indexing by extraction of words means that words from the document texts are extracted to represent the content of the document (Lancaster, 2003, p. 284). The major advantage of extracted indexing is, apart from the financial perspective, is that it enables an up-to-date indexing and searching with the vocabulary of the documents. Hence, new words and subjects occurring in the database documents will be searchable whenever the documents containing the words have been registered.

Automatic assignment of keywords, on the other hand, means that, words in documents are matched against a controlled vocabulary on basis of a set of rules (Lancaster, 2003, p. 287-288). Here the searching is based on some predefined topics and keywords, and the documents can be searched from some predefined keywords. This is a more expensive process compared to automatic extracting since the rules for the assignment of words are developed partly by humans. On the other hand automatic assignment indexing implies the advantages of controlled indexing; the possibility of coupling of keywords across a collection of documents, and representing and searching the documents from the perspective of the work domain, e.g. a specific e-government agency.

Price, Nielsen, Delcambre & Vedsted (2007) has developed and evaluated an indexing method that builds on automatic indexing. Semantic component indexing provides a supplementary, enriched description of document contents by marking up segments of text in a document (i.e., semantic component instances) with labels (semantic component names). Domain-specific documents tend to contain characteristic types of information (semantic components). With semantic components a searcher can search for query terms in specific semantic components, or specify a preference for documents containing particular semantic components. Hereby, the searcher can combine the advances of document-oriented automatic indexing and domain-oriented controlled indexing that focus at specific topics or components of the documents.

3 Studies on information seeking and searching within e-government

In this review we will focus on studies of information behaviour in a governmental setting or in settings with similar characteristics as the domain of e-government. These are Wilson (e.g. 1980; Wilson & Streatfield, 1977), Ingwersen (1994), Byström (e.g. 1997; 1999), Ellis & Haugan (1997), Toms, Freund & Waterhouse (2005), and Marcella, Baxter, Davies & Toornstra (2007).

3.1 Project INISS

Wilson's participation project INISS (information needs and information services in local authority social services departments) had the purpose of examining information needs and information behaviour among social workers and social administrators (Wilson, 1980, p. 1999). The results were supposed to be used for improving and developing information system organization and information service delivery (Wilson, 1980, p. 199). The project was carried out in a selected set of British local authorities departments representing both urban and rural departments. Furthermore the test persons reflected different categories of employees (Wilson, 1980, p. 203).

22 subjects were observed using structured observation, providing 6.000 records of communication events (Wilson & Streatfield, 1977, p. 282). The study is primarily a study of information behaviour. Hence, it is concerned with multiple aspects of the work situation of the subjects being studied. Still there are elements of information seeking in the results, e.g. when referring to the role of current awareness bulletins (Wilson & Streatfield, 1977, p. 285).

The study shows, that 74% of all sessions last 5 minutes or less (Wilson & Streatfield, 1977, p. 284). This indicates the necessity of easy and quick searches, which may be obtained by assigned, controlled indexing. In addition, the social service staff members stress the importance of clearly and succinctly presented texts, preferable in a format, which makes the identification of key elements easy accessible (Wilson, 1980, p. 211). The finding suggests that it might be fruitful to apply the semantic component indexing model. By the use of the indexing model the document will be structured into components and make it possible to identify specific text components or key elements. As regards information needs, the study indicates, that information needs among the participants are more complex than just verificative needs. The results indicate, that topical needs, whether conscious or muddled are present among the participants. However, this does not exclude the presence of verificative information needs in the social services departments. Put another way, this indicates that the work tasks of the employees generates both verificative and topical information needs. In the personal files observed in the study, a number of different information types are to be found; e.g. committee papers, pamphlets, reports, and statistics (Wilson, 1980, p. 211). It means that the length of the single units of information may be varying. When choosing automatic indexing methods, normalization for the length of documents is expected to influence the

performance of the indexing method (Anderson & Pérez-Carballo, 2001, p. 259-260).

3.2 System development in the Danish Folketing

In 1989 and onwards Ingwersen, working as a consultant on a project regarding the information systems in the Danish Folketing, analyzed the information structure and the working processes of people employed in the Folketing (Ingwersen, 1994). A substantial part of the study consists of a user and domain analysis (1994, p. 11). The empirical results are based on interviews with 32 respondents on the basis of a structured questionnaire (Ingwersen, 1994, p. 11).

Essential findings in relation to the purpose of the present paper are that some kind of controlled vocabulary is important, since legal documents are an important document type, often searched by the popular name that does not appear in the document text and cannot be extracted by automatic indexing (Ingwersen, 1994, p. 12). Furthermore, in favour of assigned indexing, are the complex dimensions of dependence between documents, which might even not keep up their original main subject (Ingwersen, 1994, p. 7).

Different types of information needs can be deduced from the results. The users search on the basis of known descriptive data from documents, which corresponds with verificative information needs. From time to time, the verificative searches are based on knowledge on the subject of the document alone (Ingwersen, 1994, p. 13). Further, there is a need for high precision in the results when searching for information (Ingwersen, 1994, p. 13). Thus, the participants prefer a list of just one or a few documents, when searching for information. This calls for assigned indexing with a level of specificity (Lancaster, 2003, p. 203-204) preventing, that too many documents are assigned with the same keywords.

3.3 Task complexity, information types and information sources in municipal governments

As a part of her ph.d dissertation, Byström (1999) conducted a study of two Finnish local (municipal) governments. Using diary, interview, organizational document review and observation (Byström, 1997, p. 132) 54 (80 if the cases from the pilot are included) cases handled by 19 officials are analyzed. Data on the cases were collected from the moment they arrived at the registrar's office (Byström, 1999, p. 67-68).

In the study Byström focuses on information seeking, and she is not specifically concerned with the actual information searching. Among other things, Byström (1999, p. 85) analyzes the relation between the complexity of work tasks and the subject expertise of the participants in the study. As expected with a group of fairly experienced participants, the subject expertise is in a lot of cases rather large. Working with a new case but within a known subject area would cause conscious topical needs. In the cases where the

participant is working on a case with an extremely low expertise on the subject, muddled topical information needs could be expected to occur.

Characterising the identified information needs according to Ingwersen's categorization (1986, p. 223), all three types of information needs are represented in the case study. It seems, though, that work experience has an influence on the information need; with more experience and knowledge about the work tasks, the nature of the need changes (Byström, 1997, p. 137). The muddled topical need is not the predominant type in the study, but it occurs, especially among inexperienced users, and therefore needs to be taken into account when indexing the documents. In sum, the study indicates, that documents ought to be indexed using both extracted and assigned indexing.

3.4 Information seeking in an industrial environment

Within the Norwegian company Statoil, Ellis & Haugan (1997) conducted a study of the role of information seeking in the Research Centre. Using semi-structured interviews 2 areas of interest were explored; seeking behaviour and project specific requirements for information. 23 respondents participated in the study (Ellis & Haugan, 1997, p. 385-386). As indicated in the title, the study primarily focuses on information as defined by Wilson (1999). However, since the role of information in the progress of projects, the study also contains elements of information behaviour.

The work tasks of the examined group are characterized by project work (Ellis & Haugan, 1997, p. 389). Three types of projects are identified; incremental, radical, and fundamental (Ellis & Haugan, 1997, p. 392). It appears that working with the different types of projects generates different types of information needs. The aim of incremental projects is to achieve minor advances in technology. The projects are typically built on established knowledge within the domain. The information needs expected to arise due to incremental projects are therefore verificative and conscious topical needs. Because of the previous knowledge of the subject, muddled topical needs will not usually be dominant.

Radical projects are "based on a foundation of existing scientific and engineering knowledge, but one which is considered insufficient to obtain the desired results" (Ellis & Haugan, 1997, p. 393). Since building on, but also supplementing, existing knowledge radical projects can generate all three kinds of information needs. Within a known subject area, verificative needs hardly can be excluded. Conscious topical needs and muddled topical needs may occur depending on the specificity of the desired results in relation to the previous knowledge. Fundamental projects enhance the company's understanding of a subject area. Since topics unfamiliar to the researchers are approached, this type of projects is expected to generate the largest amount of muddled topical information needs. Therefore automatic indexing should reflect the existence of all three project types and hence also all three types of information needs. One way of doing this could be using a combination of automatic and assigned indexing.

3.5 Information behaviour of software engineers

Using a combination of four methods; focus group, semi structured interviews, observation, and finally analysis of documents and digital information (phase I) Freund, Toms & Waterhouse (2005) examined the information behaviour of software services consultants. In the second study (phase II) reported in the article, 14 software services consultants interviewed on the basis of a semi-structured interview. The purpose of the study was to investigate information behaviour on the basis of a work task framework.

Essential results from the study of phase I is the dependency of information needs to the type of work task at hand. Work tasks can range from short term to long term commitments and the development of information needs is highly dependent of this work task context. Phase II of the study shows that the length of a work task influences on the choice of document types to resolve the derived information need. As the time available for resolving a work task increases, so does the length of chosen documents and the diversity of sources used for gathering information. Having little time for resolving a work task requires an assigned indexing, where the user can obtain search advantages from the structure of the controlled vocabulary to identify central documents quickly. However, the more time available to resolve a work task and the derived information need, the more use a user can gain from searches based on different types of indexing.

3.6 Information behaviour of the users of the European Parliamentary Documentation Centre

In 2004 Marcella et al. (2007) examined users of the European Parliamentary Documentation Centre (PDC) regarding information needs and information seeking behaviour. The main purpose of the study is to make recommendations for service development in the PDC on the basis of the study of users of the PDC (Marcella et al., 2007, p. 921). 72 semi structured interviews were conducted between different types of administrative staff. In order to assure experienced test persons, the data were collected prior to the 2004 election for the European Parliament (Marcella et al., 2007, p. 922-923).

The study explores elements of information behaviour, information seeking behaviour and information searching. For the examined group it is essential that the time available is limited. This is expressed in different ways; the participants use other people to perform their searches and time is a constraint (Marcella et al., 2007, p. 928). Like Wilson's study, this indicates that assigned indexing must be applied in order to make searches more effective. On the other hand the study shows that the topics of information needs are dynamic and continually develop (Marcella et al., 2007, p. 926; p. 931), which implies that extracted indexing that reflect up-to-date topics and vocabulary is necessary as well.

Concerning the types of information needs present among the participants of the study, there are examples of the conscious topical information need, when the subjects mention that even though they are working with well defined subject areas, there is still a need to explore relating peripheral areas

(Marcella et al., 2007, p. 925). More complex information needs, e.g. finding supranational comparisons of statistics (Marcella et al., 2007, p. 925) may still be characterized as conscious topical needs, since the subjects know the subject area ahead of the information need. Here the process of information searching is complicated by the diverse information structure across member states. The muddled topical information need is also identifiable in different ways in the results. One way is the expression of anxiety, when the participants are searching for new or highly specific subjects (Marcella et al., 2007, p. 928). Similarly, a lot of the participants from the Parliament do not know what they need to know (Marcella, 2007, p. 931). Finally, the verificative information need is represented, when the users enters full citations and/or titles in order to locate specific needed documents (Marcella et al., 2007, p. 928).

4 Discussion and conclusions

The above analysis illustrates that the organizational context of a user, in this case the context of e-government, plays a role for the users' information seeking behaviour and the development of information needs. Further, the study indicates that different types of information needs demand different kinds of indexing and searching strategies. On the basis of the analyzed studies a combination of extracted and assigned indexing is desirable. One method of combining the advantages of the two main approaches to automatic indexing could be using semantic components (Price et al., 2007). This solution will also meet the identified need of well-structured documents that allows for easy identification of key topics and key elements of the texts. However, applying semantic components within the domain of e-government presuppose, that work tasks and/or information needs can be related to document components. From the study of Freund, Toms & Waterhouse (2005) we have seen the importance of types of work tasks on the development of information needs. The relation between work tasks and document components can be made among the work domain of physicians, which was the original domain for developing the method. Whether establishing the same type of relations within e-government is possible remains to be solved. Further, it needs to be investigated, whether it is possible to make an identification of document genres within the domain of e-government, which is also a prerequisite for applying semantic components as a method for automatic indexing.

5 References

- Anderson, J.D. & Pérez-Carballo, J. (2001). The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part II: Machine indexing, and the allocation of human versus machine effort. *Information Processing & Management*, 37(2), 255-277.
- Belkin, N.J., Oddy, R. & Brooks, H. (1982). ASK for information retrieval. *Journal of Documentation*, 38(2), 61-71.
- Byström, K. (1997). Municipal administrators at work – information needs and seeking (IN&S) in relation to task complexity: a case-study amongst municipal officials. In: Vakkari, P., Savolainen, R. & Dervin, B. (Eds.), *Information Seeking in Context: Proceedings of an international Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14-16 August, Tampere, Finland* (125-146). London: Taylor Graham.
- Byström, K. (1999). *Task Complexity, Information Types and Information Sources: Examination of Relationships*. Tampere, FI: University of Tampere. Unpublished doctoral dissertation.
- Byström, K. & Hansen, P. (2005). Conceptual framework for tasks in information studies. *Journal of the American Society for Information Science and Technology*, 56(10), 1050-1061.
- Chau, M., Fang, X. & Sheng, O.R.L. (2007). What are people searching on government web sites? A study of search activity on the Utah.gov web site. *Communications of the ACM*, 50(4), 87-92.
- Craig, A. & Schriar, S. (2001). The *Find-It! Illinois* controlled vocabulary: improving access to government information through the Jessica subject tree. *Government Information Quarterly*, 18, 195-208.
- 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.
- FESD-projektet, Alliancen. [2004]. Bilag 5: Arbejdsgangsanalyser og Best Practices. [Homepage for Digital Forvaltning]. Available on 17. january 2008 on:
http://modernisering.dk/fileadmin/user_upload/documents/Projekter/FESD/Leverancekontrakterne/CSC/B05_CSC_30052004.pdf
- Freund, L., Toms, E.G. & Waterhouse, J. (2005). Modeling the information behaviour of software engineers using a work-task framework. *Proceedings of the American Society for Information Science and Technology*, 42(1).
- Ingwersen, P. (1986). Cognitive analysis and the role of the intermediary in information retrieval. In: Davies, R. (Ed.), *Intelligent Information Systems: Progress and Prospects*, (206-237). Chichester: Horwood.
- Ingwersen, P. (1994). Systemudvikling i et in-house miljø: Folketingets emneordssystem som case-studie. *Biblioteksarbejde*, 41, 5-23.
- Lancaster, F.W. (2003). *Indexing and Abstracting in Theory and Practice* (3. Ed.). London: Facet Publishing.
- Landsforeningen af Kommunale Servicecentre, Arbejdsgruppen om IT. [2005]. LKS: Projekt Borgerbetjening 2007: Rapport fra arbejdsgruppen om IT. Available on 9. january 2008 on:
www.kl.dk/_bin/c9f4430c-73b5-4bc4-8d1d-e0dda317a58b.doc
- MacMullin, S.E. & Taylor, R.S. (1984). Problem dimensions and information traits. *The Information Society*, 3(1), 91-111.

- Marcella, R., Baxter, G., Davies, S. & Toornstra, D. (2007). The information needs and information-seeking behaviour of the users of the European Parliamentary Documentation Centre. *Journal of Documentation*, 63(6), 920-934.
- Marcella, R., Baxter, G. & Moore, N. (2003). The effectiveness of parliamentary information services in the United Kingdom. *Government Information Quarterly*, 20, 29-46.
- Nielsen, M.L. (2001). A framework for work task based thesaurus design. *Journal of Documentation*, 57(6), 774-797.
- Price, S.L., Nielsen, M.L., Delcambre, L.M.L. & Vedsted, P. (2007). Semantic components enhance retrieval of domain-specific documents. *ACM Sixteenth Conference on Information and Knowledge Management (CIKM)*.
- Soergel, D. (1985). *Organizing information: principles of data base and retrieval systems*. San Diego, CA: Academic Press.
- Taylor, R.S. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries*, 29, 178-194.
- Wilson, T.D. (1980). Information system design implications of research into the information behaviour of social workers and social administrators. In: Harbo, O. & Kajberg, L., (Eds.), *Theory and application of information research: Proceedings of the Second International Research Forum on Information Science, 3-6 August, 1977, Royal School of Librarianship, Copenhagen*, (198-213). London, UK: Mansell.
- Wilson, T.D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249-270.
- Wilson, T.D. & Streatfield, D.R. (1977). Information needs in local authority social services departments: an interim report on project INISS. *Journal of Documentation*, 33(4), 277-293.