Relevance as process: judgements in the context of scholarly research

Theresa Dirndorfer Anderson
University of Technology, Sydney
P.O. Box 123, Broadway, NSW 2007 Australia

Abstract

Introduction. This paper discusses how exploring the research process in-depth and over time contributes to a fuller understanding of interactions with various representations of information.

Method. A longitudinal ethnographic study explored decisions made by two informants involved in scholarly research. Relevance assessment and information seeking were observed as part of informants' own ongoing research projects. Fieldwork used methods of discovery that allowed informants to shape the exploration of the practices surrounding the evolving understandings of their topics.

Analysis. Inductive analysis was carried out on the qualitative data collected over a two-year period of judgements observed on a document-by-document basis. The paper introduces broad categories that point to the variability and richness of the ways that informants used representations of information resources to make relevance judgements.

Results. Relevance judgements appear to be drivers of the search and research processes informants moved through during the observations. Focusing on research goals rather than on retrieval tasks brings us to a fuller understanding of the relationship between ultimate research goals and the articulation of those goals in interactions with information systems.

Conclusion. Relevance assessment is a process that unfolds in the doing of a search, the making of judgements and the using of texts and representations of information.

Introduction
The concept of relevance is at the centre of information retrieval and figures prominently in the evaluation of information systems. Whether as the articulation of an idea, interaction with other people or in our engagement with systems, relevance is at the heart of the human communication of meaning. Both these aspects of relevance come together in information seeking situations where searchers use networked systems to pursue information relevant to the task that prompted that use. While we acknowledge that systems and people work in different ways, there has been little detailed exploration of the human behaviour associated with these interactions. A fuller understanding of the human processes associated with assessing the relevance of informative artefacts, such as documents, citations or other representations emerges through investigating the way relevance is judged and communicated in authentic, work-based situations. This paper presents findings from a longitudinal, ethnographic study that explored the decisions made by researchers involved in scholarly research. Exploration of these human experiences extends earlier empirical research about the dynamic, multidimensional nature of human relevance judgements. It demonstrates that these judgements are not singular actions but are instead embedded in very diverse and complex search and research practices.

The study explored the way relevance is experienced by searchers using networked information systems and how they used representations of information (e.g., bibliographic citations, abstracts, documents). Relevance judgements were observed during the course of a search or evaluation session as part of a broader process of understanding and seeking meaning. One key feature of this exploration is the contribution to our understanding about relevance assessment as a process. In-depth study of searchers' relevance assessments affords us new insights into the richness of the various ways that relevance (and judgement thereof) is experienced during the course of search and research practices. From the perspective of searchers engaged in task-based information seeking, relevance is more than the selection or rejection of information. It is a multi-level phenomenon communicated through the absence as well as the presence of connections that researchers recognise at the time. It is also conveyed through the creation of boundaries that evolve during search and research processes. The findings discussed here support assertions that placing more emphasis on research goals than on retrieval tasks brings us to a fuller understanding of the relationship between an individual's ultimate research goals and the articulation of those goals in interactions with information
retrieval systems.

**Conceptual background: relevance and process**

In line with much information behaviour research, the study views information seeking predominantly as a communicative process and information retrieval is an interactive process of *meaning-making* over time (Burnett & McKinley 1998; Hert 1997; Talja 2002). Systems use is related to information seeking stages (Kuhlthau 1991; Hert 1997: 20; Robins 2000). People use a variety of information sources and strategies - often within a single search session-- to learn about a subject or resolve an information problem (e.g., Byström & Järvelin 1995; Spink *et al.* 1999; Vakkari & Hakala 2000; Xie 2000).

Hert (1992: 73-4) describes retrieval as a transformative process, where the searcher shapes search results. A variety of experiences shape the processes of learning and interpretation associated with searching and relevance assessment (Hert 1997; Kuhlthau 1999; Park 1992; Vakkari 1999). Contextual variables of the search process can differ from user to user. The judgement of relevance is linked to stages of the search process, successive searches, and decision processes associated with work goals (Spink *et al.* 1998; Vakkari & Pennanen 2001). Vakkari and Hakala (2000) and Vakkari (2000) describe how different types of information are sought at different stages of problem solving. The interplay of these elements of the search process helps explains the variability of relevance assessment discussed (among others) by Schamber (1991), Barry (1994), and Maglaughlin & Sonnenwald (2002).

Particularly significant for this current study is the recognition that neither the system nor the user can judge relevance in advance. People interact with many different layers of representation during information retrieval (Hjørland 2002: 260-1; Saracevic 1996a; 1996b: 210-4). Furthermore, the judgement of relevance or usefulness is not a single event simply based on text content. Relevance drives human communication, but judging relevance is far from straightforward. Inferences, interpretation and the context of communication are central to this view of relevance (Greisdorf 2000; Harter 1992; Saracevic 1996b; Schutz 1970; Sperber & Wilson 1986). While process-oriented research provides evidence of the multidimensional character of relevance and a range of criteria that may be applied in relevance assessment (e.g., Barry 1994; Bateman 1998; Park 1992; Tang & Solomon 1998; Wang & Soergel 1998), we are still unable to ascertain which criteria become most important for users and in which situations this
significance might occur. Topical relevance, however, remains one of the easiest dimensions for both searchers and systems to judge (Froehlich 1994). A number of researchers have found that the most likely criteria contributing to relevance judgements were those associated with content matters (Barry & Schamber 1998; Bateman 1998: 31; Maglaughlin & Sonnenwald 2002; Vakkari & Hakala 2000: 553).

In summary, even a cursory look at the literature shows that relevance is embedded in human cognition and situated practices of communication and information seeking. The act of retrieving something from an information retrieval system is not a single interaction, but a complex process of interaction of representations communicated between a searcher and a system. Emerging from this view of information and users is a view of relevance—or more precisely the user's judgement of relevance—as dynamic and multidimensional. A searcher's interaction with representations of texts centres on communication, language and meaning.

The study: background and method

The research examined relevance assessment within the context of an ethnographic exploration of the research practices of two academics, looking at the ways they experience the concept of relevance while using networked information resources. It sought to portray relevance as a complex interwoven activity involving a researcher in authentic search situations, framing relevance assessment, information seeking and information retrieval as dynamic and socially-situated activities. Seeking and retrieving information are firmly situated in everyday life experiences by information science researchers like Given (2002) and Talja (2002), who make the case that studying information behaviour in social contexts is essential for understanding the way those contexts shape action and interpretation. For more than two years, informants were observed engaged in the organisation, discovery, evaluation and retrieval of information as part of their research practices. Framing the development of the user construct of topic in this broader context is consistent with the methodological principles discussed by Denzin (1989) and Yin (1994), both of whom draw attention to the need to frame such questions interactionally, tracing activity over time and in context.

Connecting relevance judgements to the selection and retrieval of documents located in an information system without observing how a person works with information in daily life obscures the full picture of the way decisions about relevance, and understandings
of a topic, evolve during the course of a research process. Bean and Green (2001: 115-6) point out that, in terms of retrieval, searchers use relevance to trawl through vast amounts of information (thereby casting a wide net) as well as to narrow, filter and refine. Notions of the concept of relevance can thus appear contradictory or inconsistent to an observer who is not aware of the context in which relevance assessment is made. Relevance cannot be examined in isolation from the particular situation in which information is pursued, evaluated and utilised.

For the study, a central area of investigation centred on understanding how a topic is articulated and adapted during a research process. In particular, the study sought to explore:

- How searchers use informative artefacts (such as documents, citations or other representations) to identify information that is relevant to them; and
- How the meaning of relevance (topic) is communicated during a search.

The starting point was searchers using networked information systems, but the searchers' individual research interests, reactions and responses to information they encountered drove the inquiry. Informants were observed searching and evaluating both networked and print information resources (e.g., citations, abstracts and texts). They were also observed preparing documents as part of their research work. The fieldwork thus used methods of discovery that allowed the informants to shape the inquiry and for their information seeking to be observed in context; that is, as part of their own ongoing research projects.

Both informants ('Catherine' and 'John') were experienced users of networked information systems. They were selected to participate in the study because they were experienced academics at, or near, the beginning of research projects involving the use of networked information systems (e.g., bibliographic databases, digital libraries). Both were senior lecturers at the same university, but within different faculties. Both informants were involved in ongoing projects. Catherine was working part-time on a Ph.D. thesis, and was first observed when she was midway through her candidature. John was observed in the early stages of a collaborative research project.

Fieldwork involved engaging as a participant-observer with informants, observing and discussing their discovery, selection, evaluation and use of information. To more fully understand how each judged relevance during their search practices, it also involved examining expressions of their topic and the processes by which they made sense of what they found. A multi-layered
narrative was created by weaving together different ethnographic stories—impressionist tales (Van Maanen 1988) along with anecdotes and vignettes (Ely et al. 1997)—with passages from field notes, e-mail correspondence, video and audio records and other documents associated with the story of the two informants.

Ethnographic storytelling served as a tool for both the analysis and presentation of informants' experiences. The layered transcriptions created for each recorded encounter with John and Catherine are at the heart of these narrative forms. Texts of the audio and video recordings of search sessions, relevance evaluations and discussions were created through repeated listening and watching of each recording. These texts combine words and actions observed on tape with my own field notes, journal entries of the events and analysis prompted by hearing comments and watching actions in the context of the recorded event. This approach to transcriptions lies somewhere between the conventional transcript and the records of talk-in-interaction prevalent in conversation analysis (CA) (e.g., Silverman 1997; Suchman 1987) and ethnomethodologically-informed ethnography (e.g., Crabtree et al. 2000). In this way, the writing of the research narrative enabled more evocative representations of the research and became a powerful device for understanding relevance interactions.

Acknowledging the value of Vakkari's (1999) claims that information retrieval involves more than interaction within a single search session, the study sought to explore relevance assessments as part of the decision-making processes of individuals doing research. Furthermore, it was decided that the searcher must not only be the centre of investigations into relevance, but needs to be allowed to drive the exploration. The following sections of this paper discuss key features emerging from that exploration and the implications for studying relevance in context.

**Key features of relevance in the context of scholarly research: experiences from the field**

Observed experiences were varied, dynamic, and shifting. judgements of relevance were ongoing activities, embedded in the search and research processes in which the informants were engaged. Fieldwork demonstrated that, when making judgements about relevance, researchers draw on interactions with colleagues (e.g., face-to-face, e-mail and casual as well as formal encounters) and with ideas communicated in their own works as well as those of other researchers. The informants applied experience, prior
knowledge (e.g., key figures, critical issues) and intuition in these situations to personalise their judgements. Judging relevance was not only used to decide what information to pursue, select or reject, but as a strategy for managing research. In this way, relevance assessments helped informants shape individual search sessions (current and future) and the research process at large.

Informants' relevance assessment involved using many different representations of information in varied and interchangeable ways, as demonstrated by the rich and complex range of practices listed in the table presented here. These practices (and the broad categories listed in the table) are neither mutually exclusive nor exhaustive. What they do point to is the variability and richness of the ways that informants used representations of information resources to make relevance judgements. The table illustrates the many ways that searchers involved in scholarly research applied relevance judgements in the course of their individual research activities. Further analysis will tease out the distinctions and significance of the observed practices.
<table>
<thead>
<tr>
<th>Predominant focus of the practice</th>
<th>Search Session Impact</th>
<th>Research Project Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determining appropriateness of information (exploring/selecting/rejecting)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building a profile of a new item</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Articulated Clues - Looking for clues in text to relate a new item to:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: what's been collected and/or rejected</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>: ideas of interest to the searcher</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>: ideas considered important</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Intuitive Clues - Finding a way in to unfamiliar territory:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: literature</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>: conceptual</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>: less developed topics of search and/or research</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Shaping boundaries to a topic (exploring/formulating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defining boundary for a search session and/or research tasks by:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: revisiting or reviewing boundaries</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>: setting limits</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Developing frameworks for the research:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: operational; tasks</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>: conceptual</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Forming a focus for:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: the research project</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>: a particular search session</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Managing the scale of the research project and/or search session by developing priorities:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>: in relation to search tasks</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Relevance judgements in the context of
: in relation to documents or other representations ✓ ✓

: in relation to research tasks or subtasks ✓

Managing the scale of the research project and/or search session by:

: ranking representations ✓ ✓

: determining key or peripheral concepts or ideas ✓ ✓

Identifying what is out there and what is not there ✓ ✓

Making links to:

: items already known to the informant ✓ ✓

: familiar ideas, issues, concepts ✓ ✓

The practices listed here are interrelated, illustrating the varied and complex ways that relevance is experienced both at the point of a particular search and as part of the wider research process. The table lists judgements observed on a document-by-document basis. Analysis of those observations demonstrates that representations of information (e.g., bibliographic citations, abstracts, documents) are used in varied and interchangeable ways. Relevance judgements appear to be drivers of the search and research processes informants moved through during the observations. The next section will discuss significant themes in greater detail.

Discussion: relevance and the research process

Rather than focus on the qualitative richness of the material collected, which has been discussed elsewhere (especially Anderson 2000), this paper seeks to emphasise that exploring the research process in-depth and over time contributed to an understanding of the informants' interactions with various representations of information. Relevance was part of the process informants used to manage their engagement during a particular search session as well as the scope of the larger research project itself. Decisions about relevance were not necessarily related to the physical selection of items located during a search session. This interaction involved not only the notion of searcher-system communication, but also a range of encounters that informed and influenced that particular communication at the search interface.
The findings support earlier research on the use of topicality as a criterion (especially, Barry 1994; Bateman 1998; Park 1992; Wang & Soergel 1998) and the richness of topical relevance (Bean & Green 2001). While the study did not seek to identify specific criteria used during evaluations, Catherine and John seemed to refer to non-topic elements of an item as part of its content. Moreover, relevance judgements were present not only in the informants' evaluations of citations or full articles, but also in discussions about their search sessions and research activities. Criteria such as author, personal experience, currency, access and uniqueness were used in conjunction with associative triggers emerging from their personal interactions with colleagues or written texts (either their own or those prepared by others).

At the time of the study both Catherine and John were in what might be considered the early to midpoint of their individual research projects. As such, broadly speaking, the practices presented in the above table appear to be predominantly associated with focus formulation: the selection, exploration and formulation aspects of the information seeking process as discussed by Kuhlthau (1991) and Vakkari (1999b). However, it is important to acknowledge that neither informants' activities fit neatly or singularly into any such category. Furthermore, throughout the period of observation, both Catherine and John were writing research papers and preparing documents that one could argue involved moving through all stages of the information search process; implying that, in some ways, each informant had moved beyond the focus formulation stages.

As indicated in the table in the previous section, the practices informants used to move through a search and to manage their larger research projects centred upon:

- determining appropriateness of information; and
- shaping boundaries to a topic (of a search task and the research project itself).

Some interesting features emerging from the analysis of these two clusters are discussed in greater detail in the following sections.

**Determining appropriateness of information**

Catherine and John looked for clues to help them determine content, applying their knowledge of authors, journals and genres to judgements about documents (or representations of documents) they were examining. They were observed looking for and making use of *trigger words* (as they called them) in citations, documents
or referential material associated with an item. Triggers and clues emerged through engaging in search and research activities (such as document and proposal writing, e-mail communication, conference participation and personal meetings with colleagues) that enable informants to judge the appropriateness of texts and their representations.

Decisions to select or reject articles were based on a combination of factors that worked together in varied ways. Sometimes they were able to make a quick decision one way or the other, but other citations took careful reading and reflection before a decision was made. Specifically, analysis of the informants' evaluation processes builds on earlier relevance criteria research and suggests:

- Trigger words aided the selection process, helping to manage the exploration of unfamiliar literature.
- Authorship triggered relevance judgements by providing clues about the content of a document or representation.
- Titles triggered relevance judgements by providing a sense of an article's likely content, helped with selection decisions.
- Particular genres, or information types, which appeared to trigger relevance judgements at the boundaries as defined at the point of evaluation.

The dynamic nature of the interplay between criteria is illustrated throughout Catherine and John's stories. Certain words or word combinations acted as triggers to help flag the potential significance of the texts or representations under review. For example, spotting favourite writers, significant personalities in the field, familiar names in the author field or in a reference list provided some sense of what the item under review could be about and helped informants to work out how an unfamiliar item might relate to their own research or search goals. At times, mentions of a familiar or significant author or a reference contained in a respected journal were pursued, even if informants had reservations about the way the content might relate to their needs. However, there were also occasions when informants decided not to select the citation, even without full access to the content, because a sense about triggers spotted in a title, text or citation was enough to confirm a decision not to select an item.

Triggers work together to prompt very personal reactions that provide important clues to the content. Context also has a significant impact on this decision-making process by giving shape to expectations. Clues about database contents, for instance, created expectations - pre-existing knowledge of the content of
and experience with each particular database appeared to influence selection decisions. When judging a citation, abstract or document, trigger qualities are a result of the searcher's interaction with,

- ideas, people and texts encountered prior to a particular judgement;
- perceptions of the type of information to be expected (e.g., the type of article that would be contained in a particular journal helps with judgements); and
- understanding of the terminology and the concepts under discussion.

This portrayal emphasises the emergent quality of working with a topic and understandings about what constitutes a relevant piece of information in the course of dealing with information systems (human as well as mechanical), people and texts in various forms.

**Shaping boundaries to a topic**

Relevance judgements were not always associated with the selection or rejection of a citation or document. Engaging with texts, people and ideas shaped more than informants' citation or document selections. It played a pivotal role in the ways John and Catherine formulated their topics, their research themes, and the intended scope and depth of their research. Texts include databases and written representations ranging from online citations through to the full texts of retrieved material and papers prepared by colleagues, contacts and themselves. Relevance was a way for Catherine and John to work through the ideas they were interested in pursuing. Their judgements about the relevance of representations, citations and documents were an integral part of the shaping, expanding, refining and reforming of boundaries for both their search and research activities.

Experiences with texts of all kinds, tasks and topics helped the informants to frame their work, articulate their topics and respond to the content of texts they evaluated. When trying to explain what they were looking for in a specific document or representation, Catherine and John described connections to other material they had read, people they had encountered, and experiences they recalled. Catherine's need to focus on preparing course outlines for a forthcoming term and her attendance at a conference were shown to have impacts upon her awareness of the issues related to her topic, to relevant texts and to relevant terminology. Similarly, the impact of John's research proposal writing, as well as the interaction with his research assistant and his colleagues can be seen in his judgements of relevance. This dynamism is even stronger at the boundaries or margins of the topics being explored.
This notion of boundaries is connected to Hert's (1997) view of the multiple levels of information retrieval interactions. Practices observed on the document or citation level showed informants were making search as well as research decisions. Both informants appeared to make clear distinctions between their short-term and long-term research needs. These practices support Hert's (1997: 110) description of searchers in retrieval interactions operating simultaneously in two different timescales: moving through information-seeking-and-use process (macro timescale); and choosing to interact with an information retrieval system (micro timescale). Interaction associated with retrieval is only one type of information-seeking sub-process.

Whether the system in question is human (for example, a research assistant or librarian) or mechanical (for example, a database network), searcher-system interaction requires a conversion from an unarticulated understanding of what is being pursued to articulations of that understanding in a form comprehensible by the information system. Analyses of informants' experiences also offered insight into the ways unarticulated judgements shape and are shaped by experience. Such judgements appear particularly significant in two situations:

- when it is difficult to formulate a topic and fully establish trigger words. In such situations, the informants seemed to rely more on gut feeling about an item's relevance which cannot be readily articulated to either human or computer intermediaries; and
- understandings that cannot be readily articulated seem to be particularly significant when dealing with information perceived to be marginal or near the boundaries of the topic of interest to the searcher at that moment of relevance assessment.

At the margins of their understanding about their topic, informants' judgements were very different from the binary, topic-matching calculation of retrieval systems. From a searcher's perspective, retrieval is more than the physical act of selecting or rejecting a document, citation or representation. It involves identifying, locating and working with the ideas embodied in the representation under review. Positioning relevance as a process rather than as a single act draws attention to the difference between the physical acts of retrieval as they might be interpreted by a mechanised information system and this searcher experience. Building on earlier discussions about task-based information seeking (e.g., Hjørland & Christensen 2002; Vakkari 2003: 444-5; Vakkari & Hakala 2000), it also contributes to our understanding of the complexity experienced by searchers in focus formulation stages.
Conclusion: evolving concepts of a topic

The study of user-centred relevance judgements is important because effective information retrieval systems inevitably will be evaluated (explicitly or implicitly) on the basis of human relevance judgements. The study presented here demonstrates that, from the searcher's perspective, relevance is an information behaviour that extends beyond judgements made in the context of information retrieval. Neither the searcher nor the system can judge relevance in advance. It is a process that unfolds in the doing of a search, the making of judgements and the using of texts and representations of information.

Key findings from this study include:

- through interactions with people and texts, searchers generate trigger words that help them to interpret and evaluate the information they encounter;
- the complexity of human judgements of relevance is particularly evident at the margins of understanding about a topic, where researchers' judgements are radically different to the binary, topic-matching calculations made by retrieval systems; and
- separating assessments of relevance from the physical act of retrieval allows, a clearer understanding of the evolving character of a topic and its relation to the user's information need to emerge.

The complexity of relevance assessment compels us to observe the process in context over time. Acknowledging relevance as an integral part of human communication supports the need to observe the human dimensions of relevance using participant-driven methodologies. Analysis emerging out of this ethnographic exploration demonstrates how understandings of a topic evolve during the course of searching and research practices: such understandings are at the heart of any judgement of relevance.

Situating relevance assessment practices within research, rather than a single search event, has contributed to a fuller understanding of the contrast between searcher and system depictions of documents and document representations. Observing selection behaviour at the information retrieval level only scratches the surface of the richness and variability of searcher experiences. Networked information system use is effectively human-human communication mediated by information retrieval systems. For a searcher, the texts and citations that are represented within a bibliographic database, for instance, embody the ideas of other researchers. The decision to select or not select, to pursue or not pursue relates to the searcher's interests, goals, ambitions,
concerns and view of self. When interacting with networked information systems, a searcher's understanding of the content represented on the screen evolves over time. Nuances of relevance judgements made during a search relate to the creation of boundaries for a particular search as well as for the scholarly research project as a whole.

Relevance assessment is a process by which a searcher constantly shapes, defines and refines searching. We do not fully understand how these judgements take shape, nor do we fully understand how they are communicated. However, we can see, as Green (2001: 14) suggests, that understanding the ways relevance relationships are communicated and judged is an important step in the development of more responsive information systems. Furthermore, understanding judgements of relevance requires understanding the contexts in which they occur. If we are to create interfaces that can assist with navigation, it is important that we understand how searchers navigate information systems, how search tools fit into the broader task and how searchers decide the relevance of information in these contexts. Exploring relevance as it is experienced in context contributes to a richer understanding of how relevance is generated in the course of a search and in the process of research.

References


